

# CSE 4125: Distributed Database Systems Chapter – 5

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

Topics to be discussed –

- Simplification of Non-distributed Query

Question type:

/\* A query is given \*/

Now, answer the following questions.

- i. Draw the *operator tree*. [2]
- ii. Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [5]

# **Non-distributed (Equivalence Transformation for Queries ):**

- ❑ Query  $\rightarrow$  Operator Tree.
- ❑ Operator Tree  $\rightarrow$  Simplified Operator Tree.

# We will follow this -

## Equivalence Query transformation steps:

1. Generate the equivalent operator tree ( $T_{\text{global}}$ ) for the given query ( $Q_{\text{global}}$ ).
2. Find the common sub-expression ( $R$ ) from  $T_{\text{global}}$ .
3. Apply rules to remove  $R$  and obtain simplified tree  $T_{\text{removed}}$ .
4. Apply criteria – 1 and 2 on  $T_{\text{removed}}$  to obtain final simplified operator tree  $T_{\text{transformed}}$ .
5. Write the query  $Q_{\text{transformed}}$  from  $T_{\text{transformed}}$ .

$$\text{So, } Q_{\text{global}} \leftrightarrow Q_{\text{transformed}}$$

# Some Rules/Properties

## Properties

- $R \text{ NJN } R \leftrightarrow R$   1
- $R \text{ UN } R \leftrightarrow R$   2
- $R \text{ DF } R \leftrightarrow 0$   3
- $R \text{ NJN } \text{SL}_F R \leftrightarrow \text{SL}_F R$   4
- $R \text{ UN } \text{SL}_F R \leftrightarrow R$   5
- $R \text{ DF } \text{SL}_F R \leftrightarrow \text{SL}_{\text{NOT } F} R$   6
- $(\text{SL}_{F1} R) \text{ NJN } (\text{SL}_{F2} R) \leftrightarrow \text{SL}_{F1 \text{ AND } F2} R$   7
- $(\text{SL}_{F1} R) \text{ UN } (\text{SL}_{F2} R) \leftrightarrow \text{SL}_{F1 \text{ OR } F2} R$   8
- $(\text{SL}_{F1} R) \text{ DF } (\text{SL}_{F2} R) \leftrightarrow \text{SL}_{F1 \text{ AND NOT } F2} R$   9

They will be used to remove common sub-expressions in the simplification of operator tree.

# Example 1

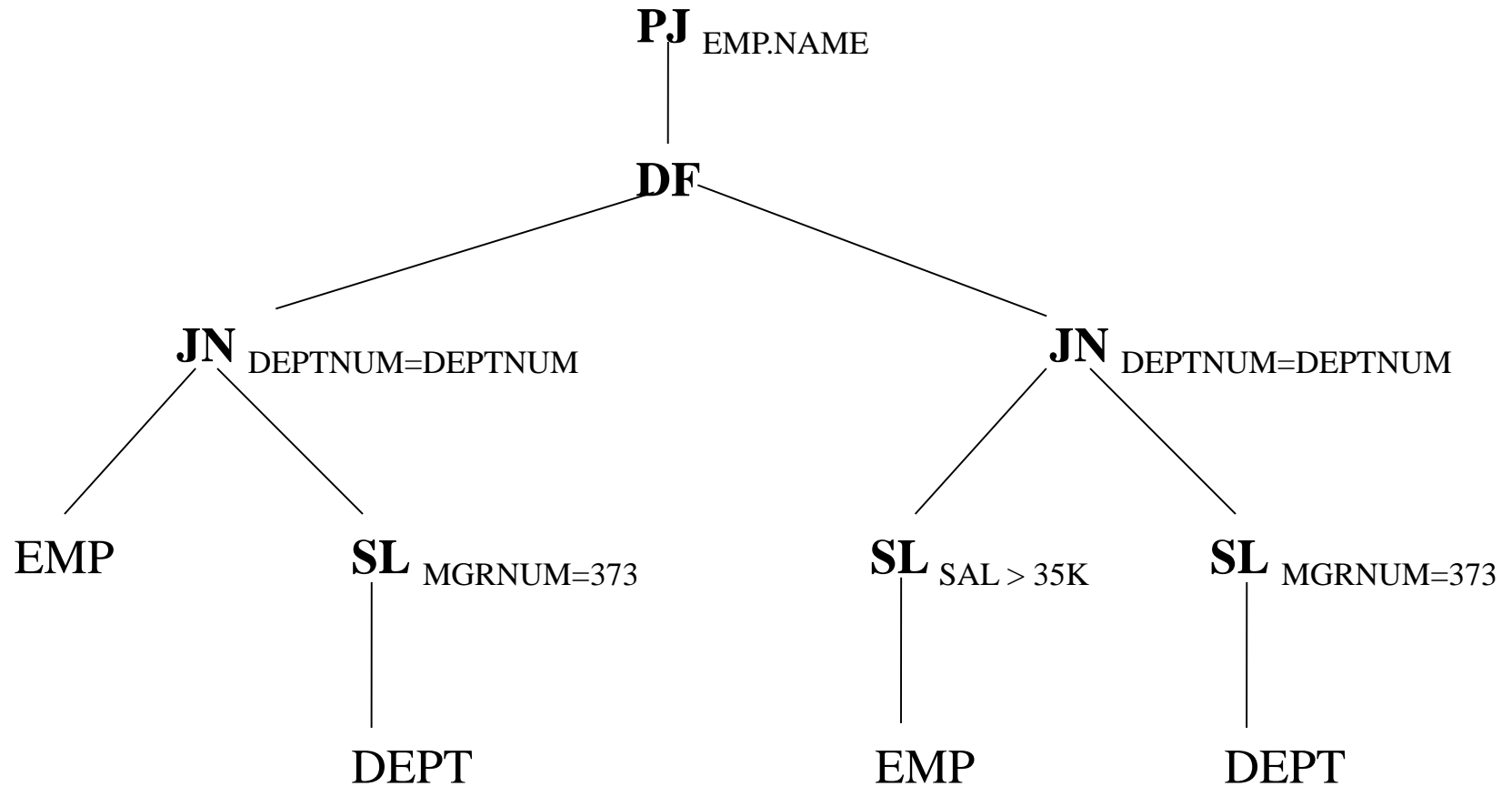
*EMP (EMPNUM, DEPTNUM, NAME, SAL, AGE)*  
*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

**Q: PJ** *EMP.NAME ((EMP JN DEPTNUM=DEPTNUM SL MGRNUM=373 DEPT ) DF (SL SAL > 35K EMP JN DEPTNUM=DEPTNUM SL MGRNUM=373 DEPT ))*

Now, answer the following questions.

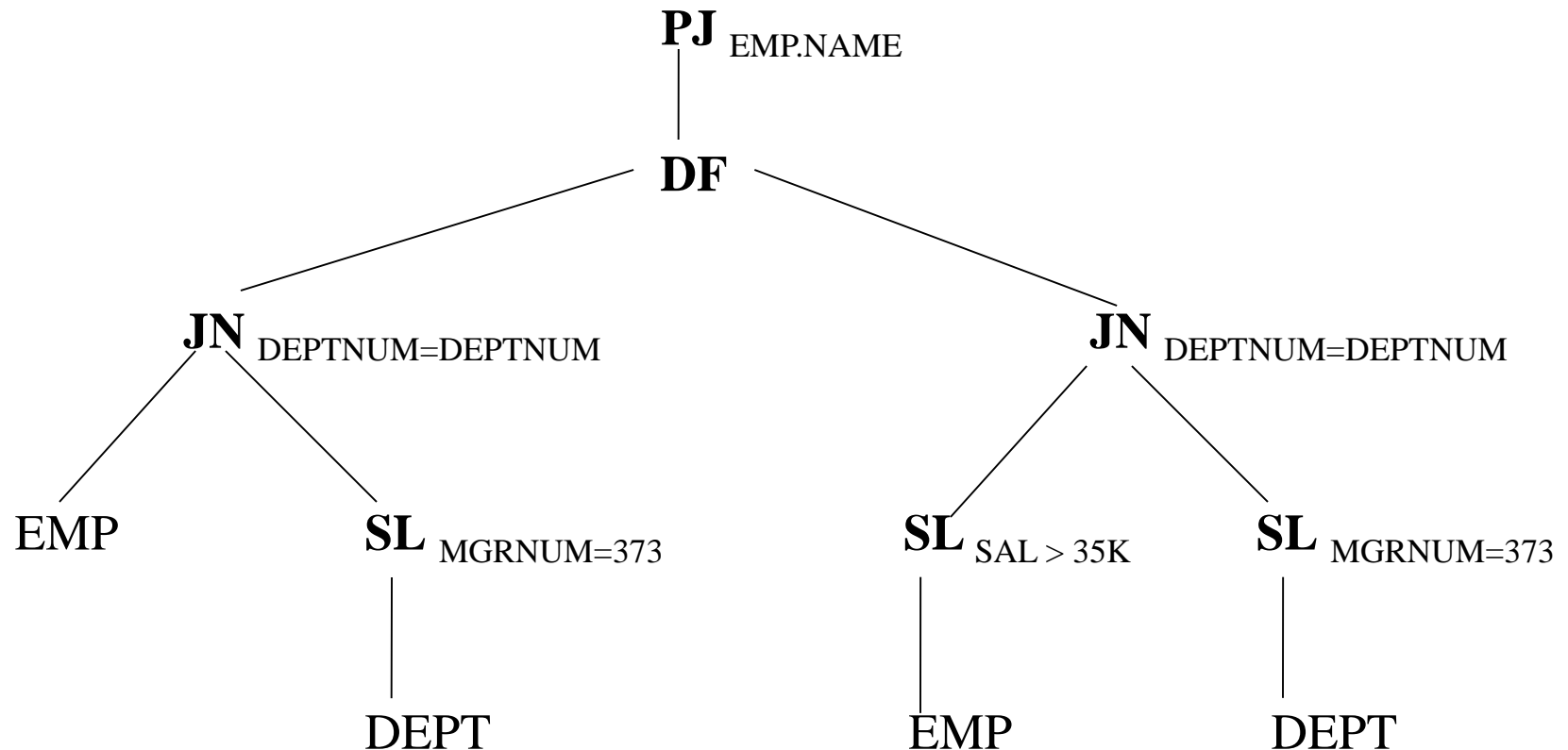
- i. Draw the *operator tree*. [2]
- ii. Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [5]

# Operator Tree



# Finding Common Sub-expression

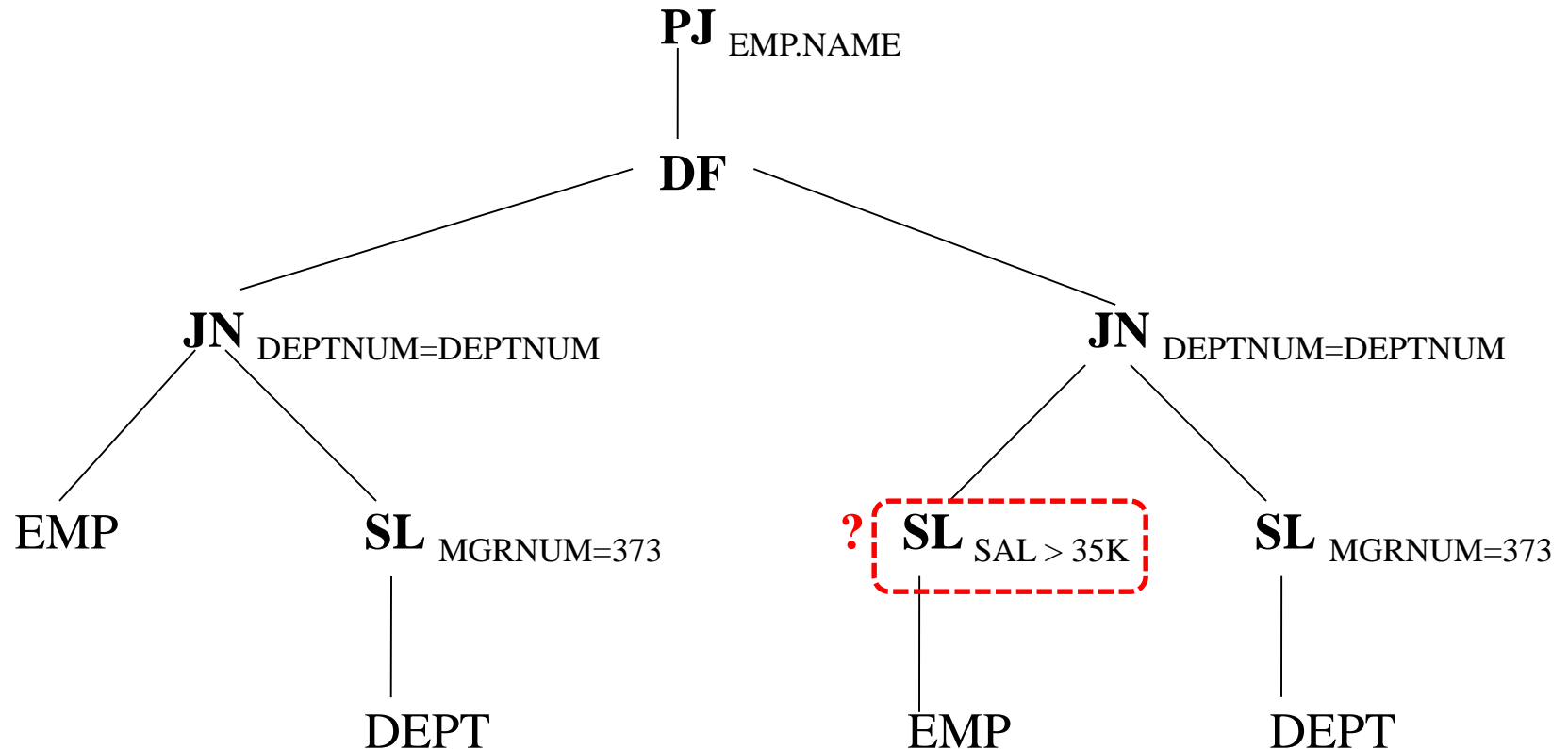
Any common portion?





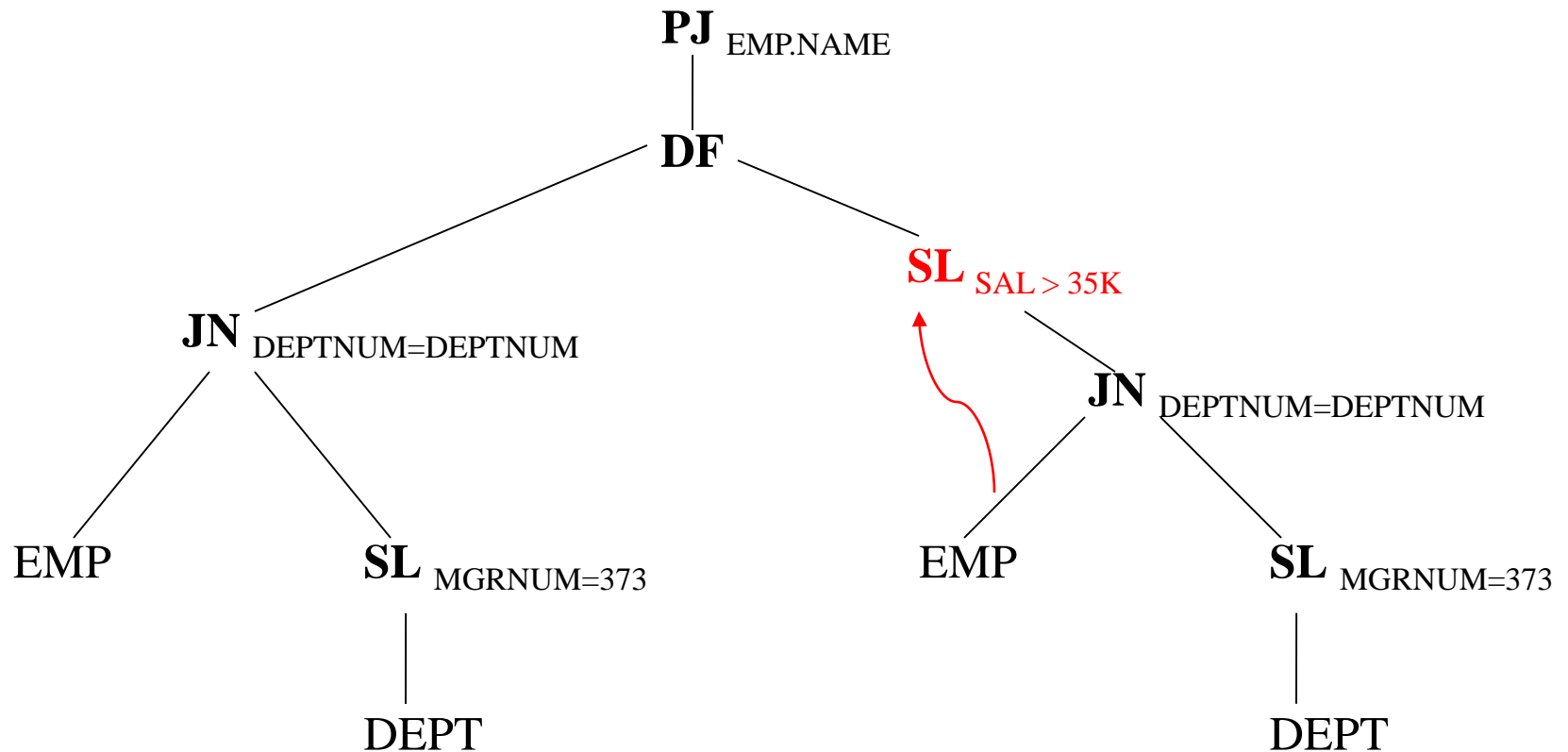
# Finding Common Sub-expression

Any common portion?

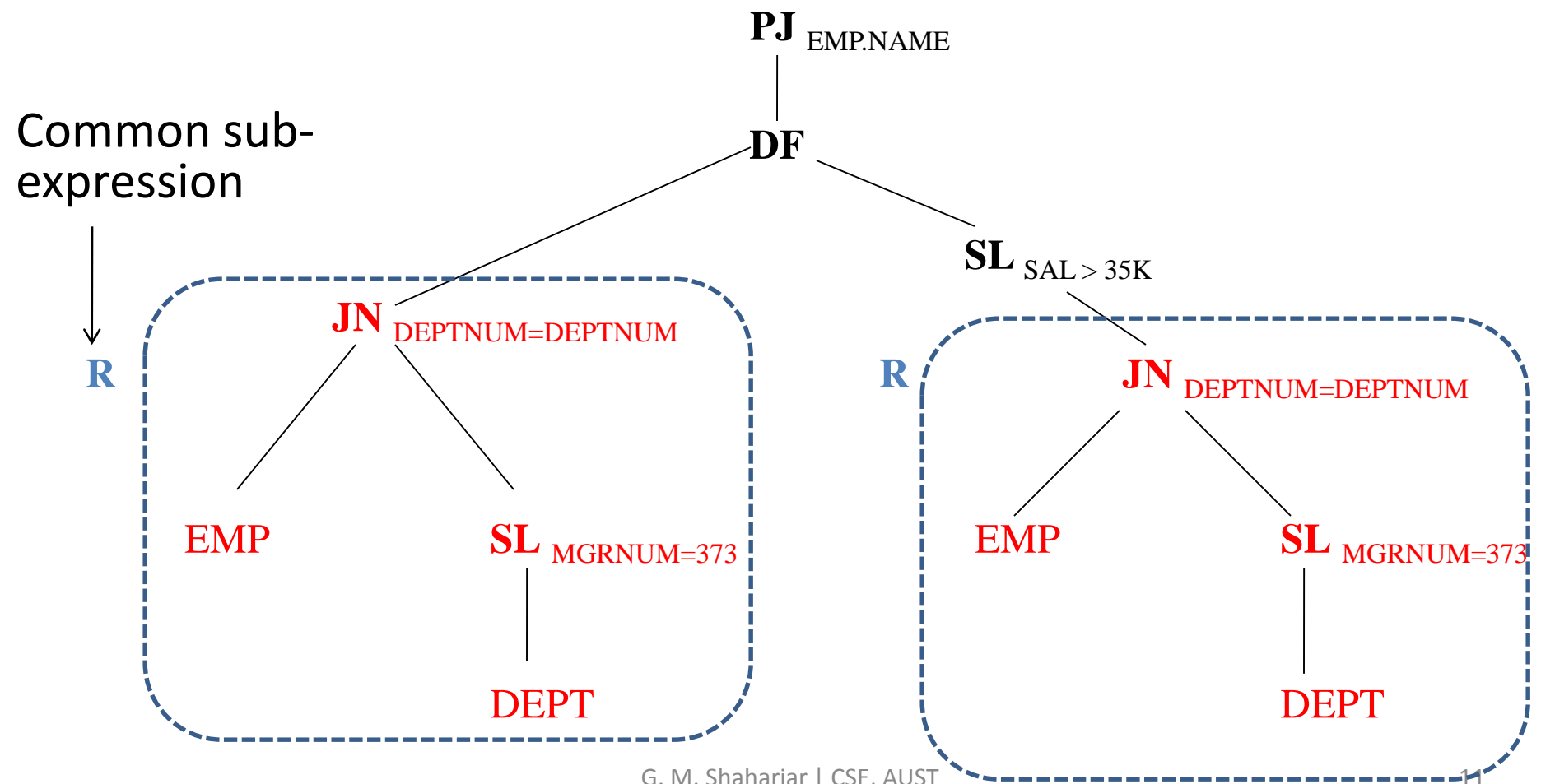


# Finding Common Sub-expression

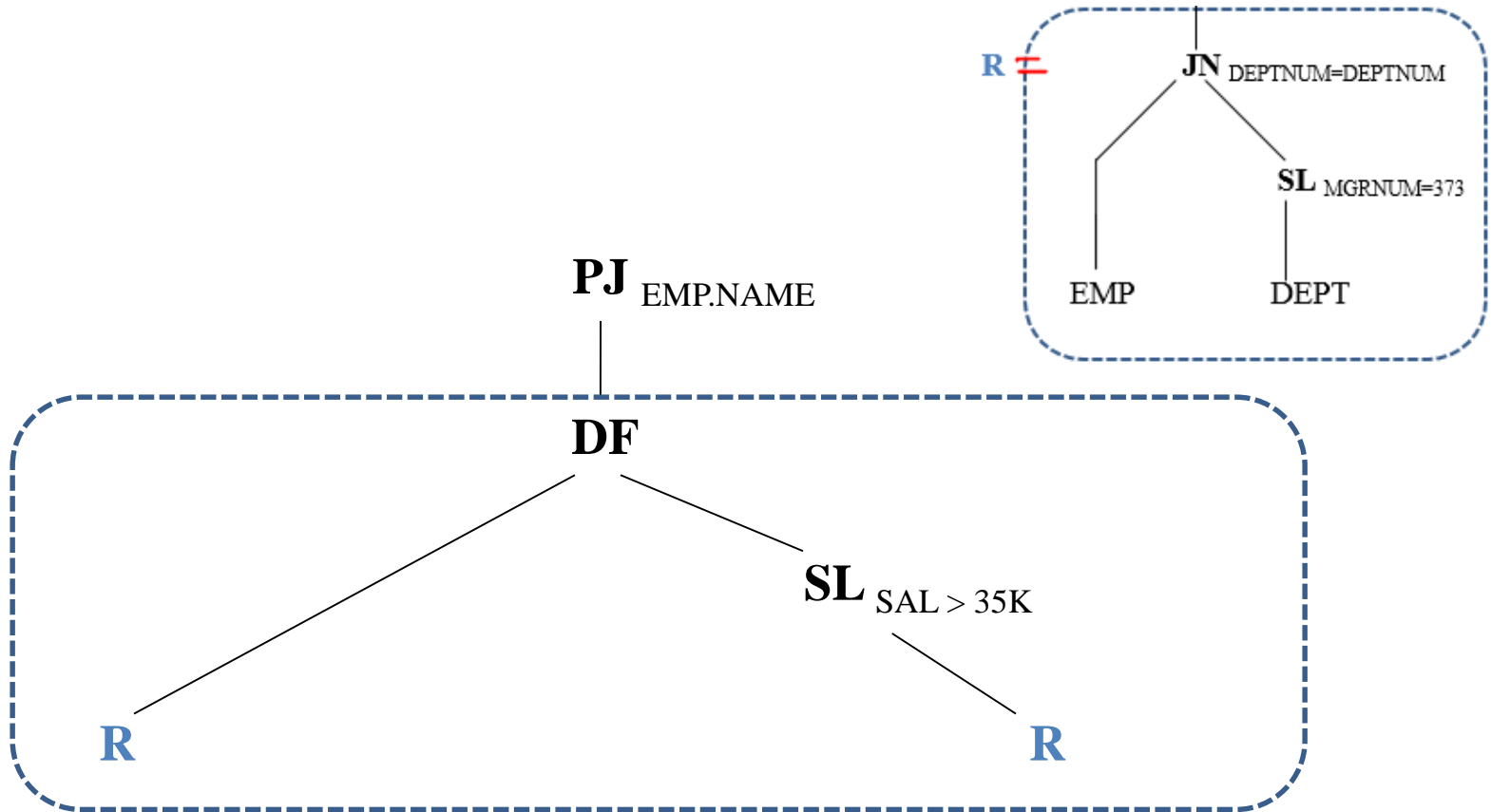
Any common portion? NOW?



# Finding Common Sub-expression

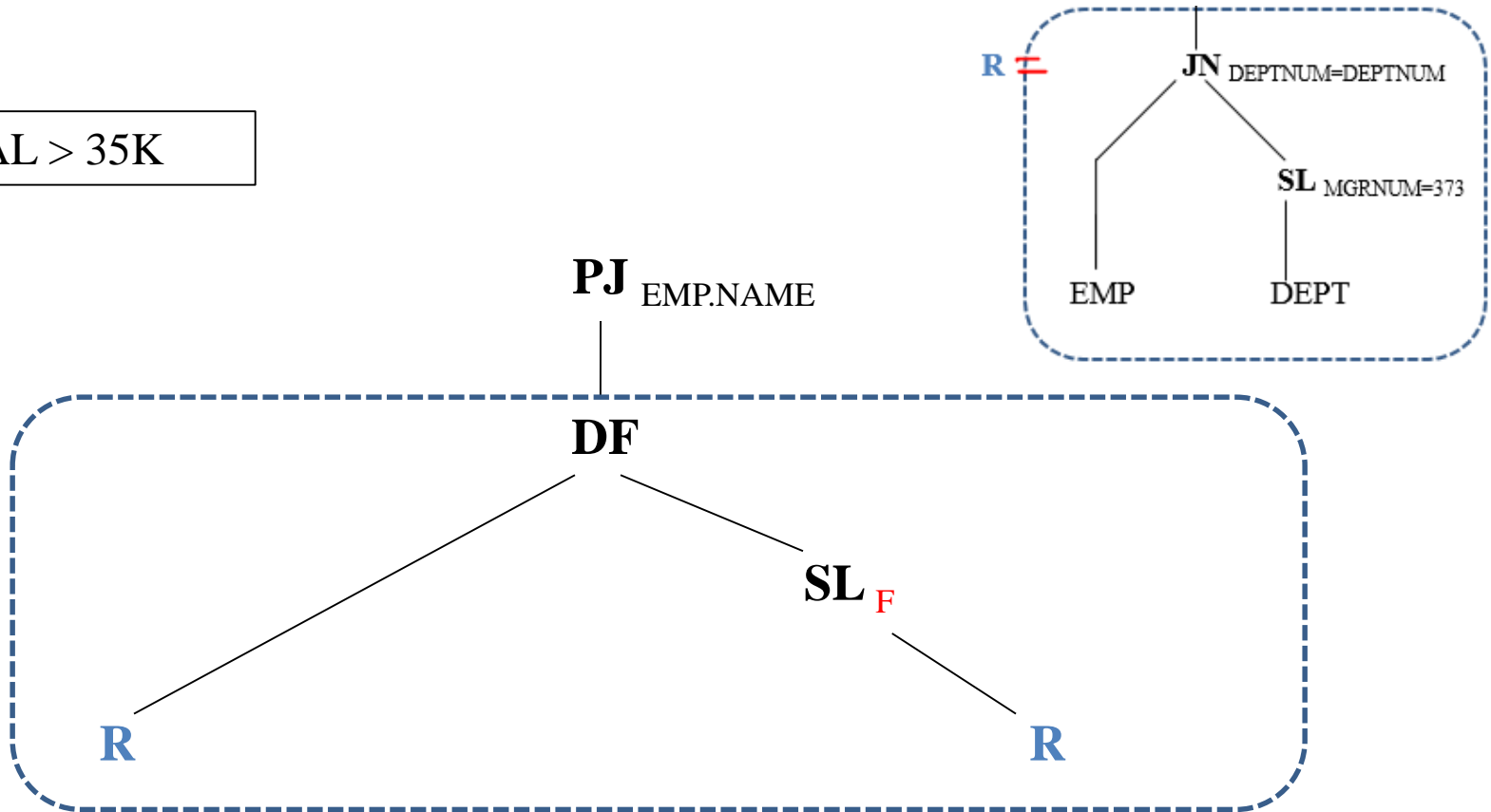


# Finding Common Sub-expression



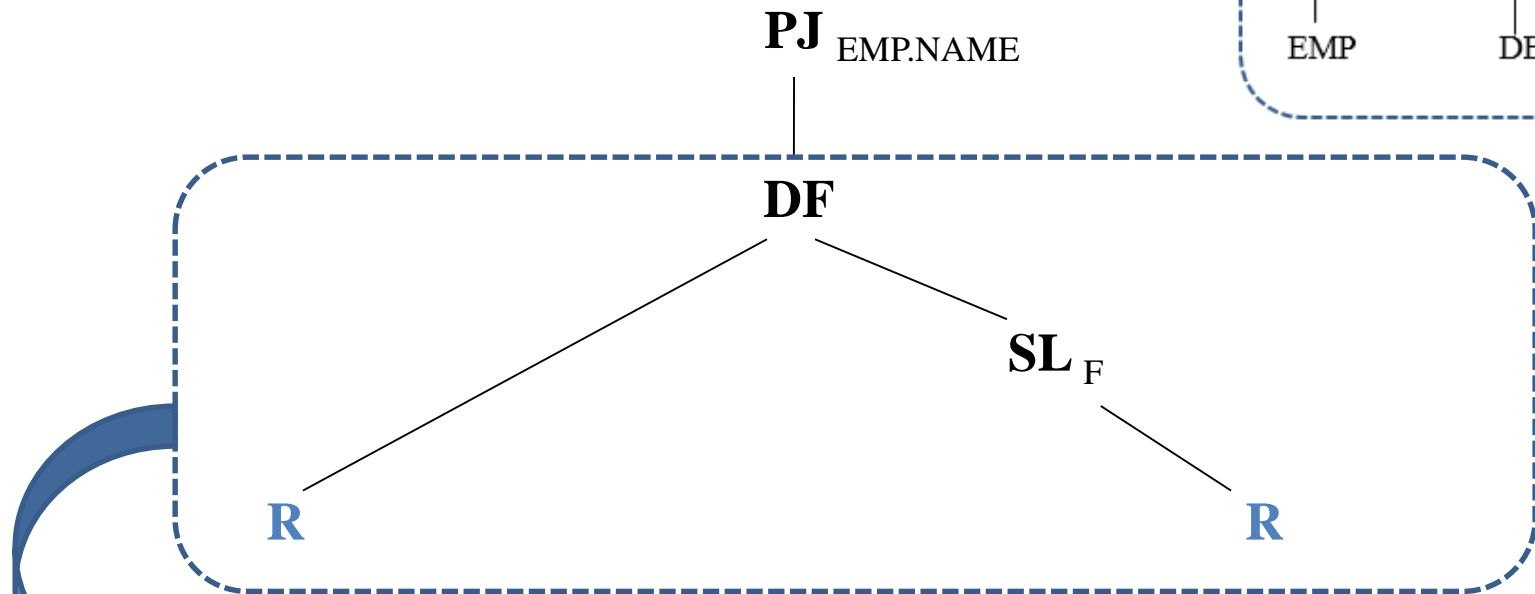
# Finding Common Sub-expression

$F = \text{SAL} > 35\text{K}$



# Finding Common Sub-expression

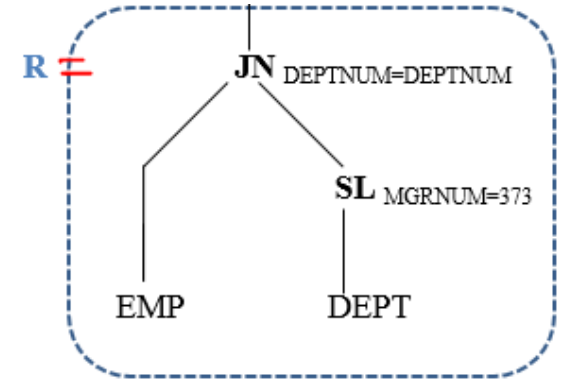
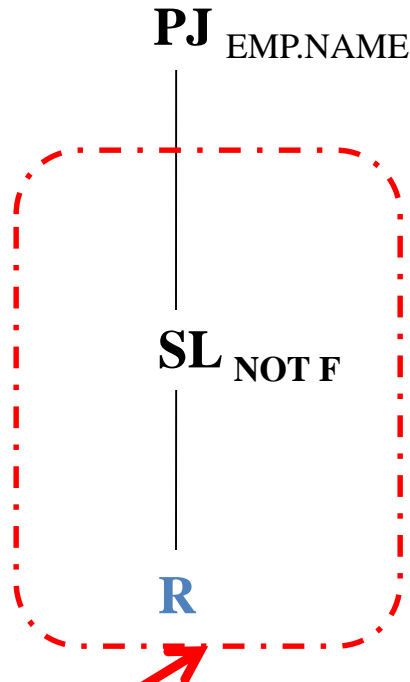
$F = \text{SAL} > 35\text{K}$



We can write it as  $\rightarrow R \text{ DF } \text{SL}_F R$  which is Rule 6 !

# Removing Common Sub-expression

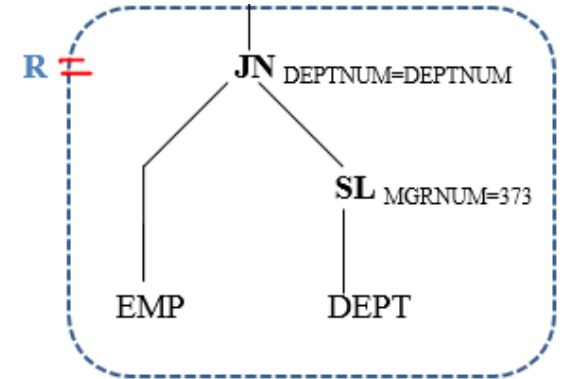
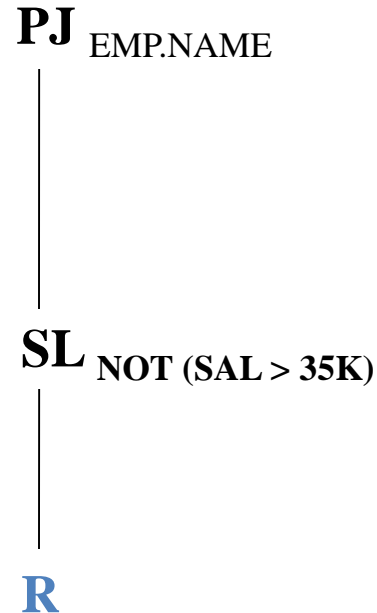
$F = \text{SAL} > 35K$



•  $R \text{ DF } SL_F R \leftrightarrow SL_{NOT F} R$

# Removing Common Sub-expression

$F = \text{SAL} > 35\text{K}$



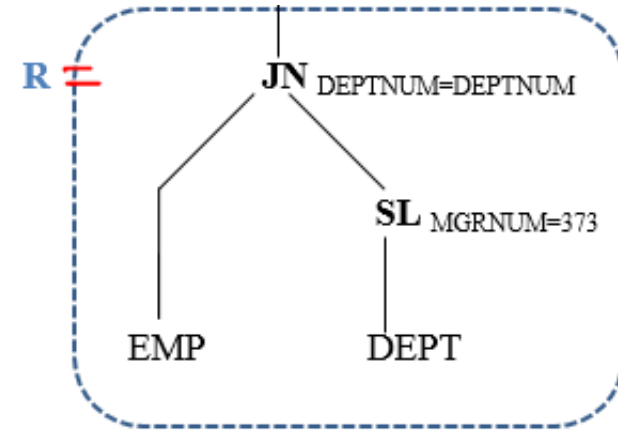


# Removing Common Sub-expression

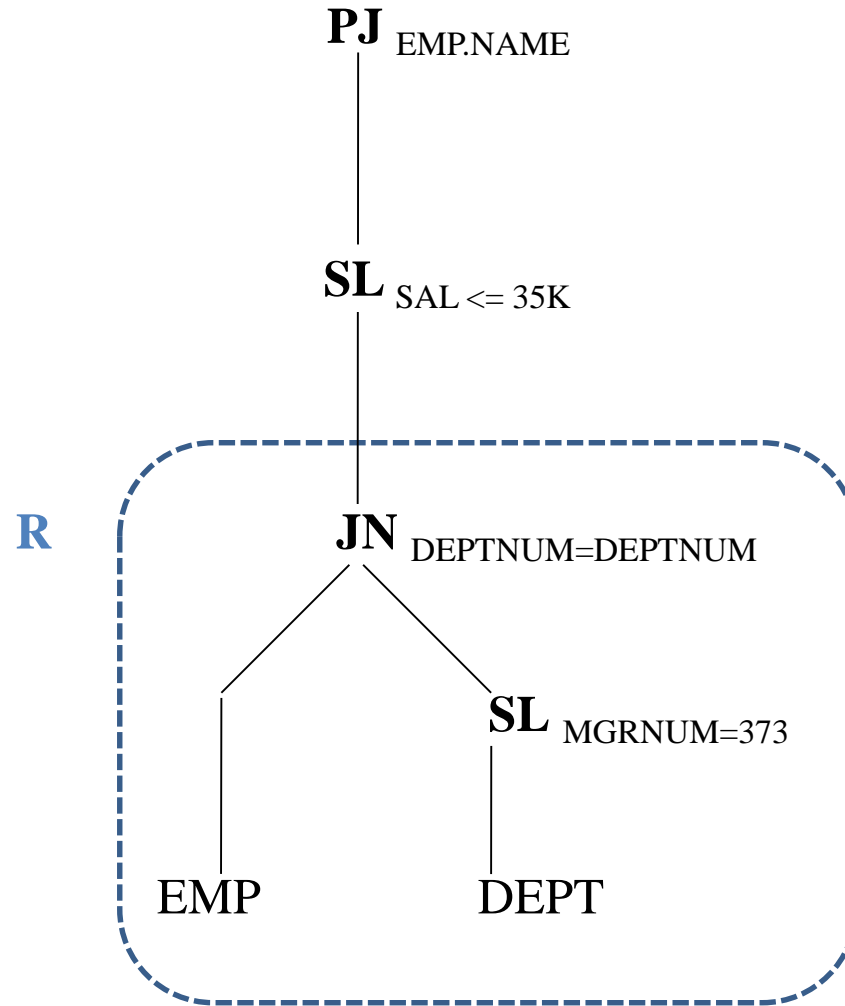
**PJ** EMP.NAME

**SL** SAL <= 35K

**R**



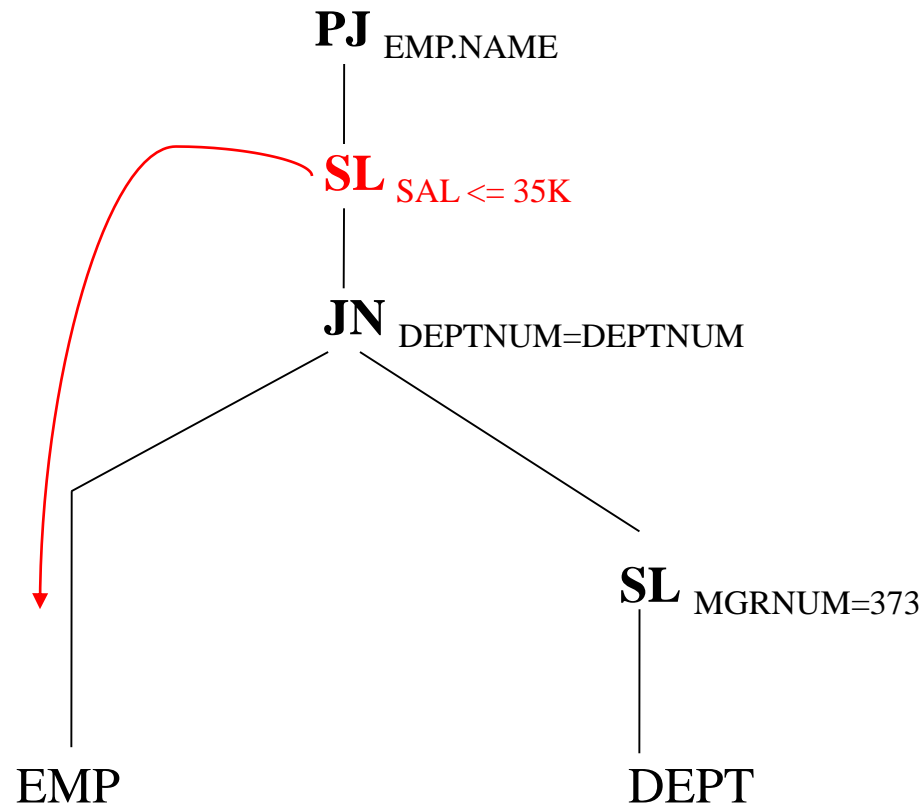
# Removing Common Sub-expression



Can you apply Criterion 1 and/or 2 on this tree?

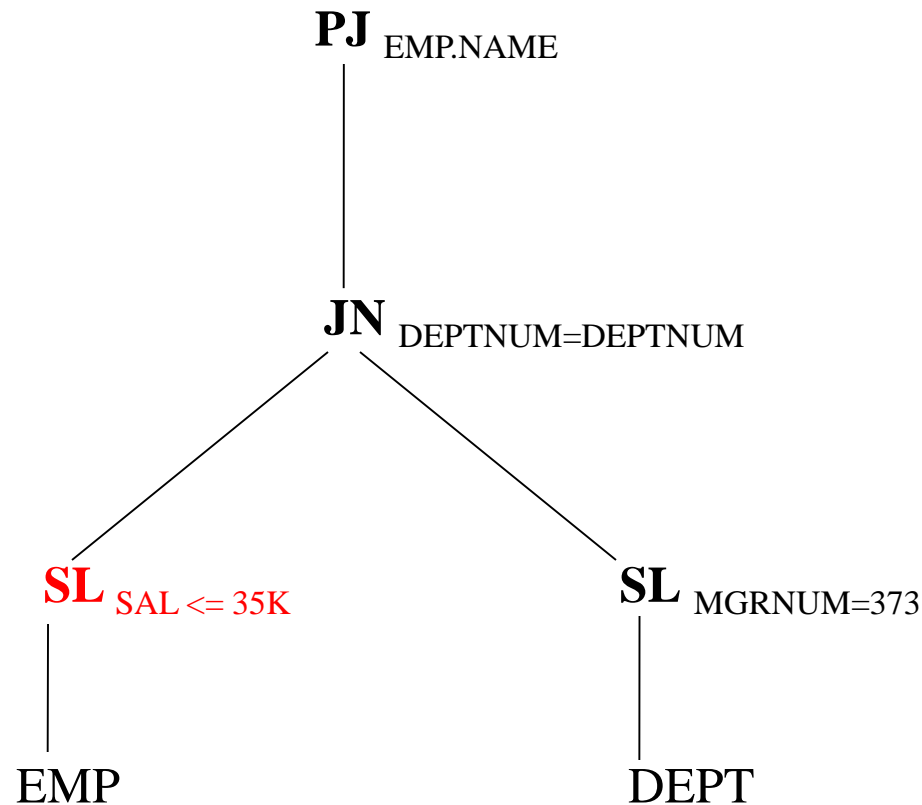
# Simplification

## Applying criterion – 2



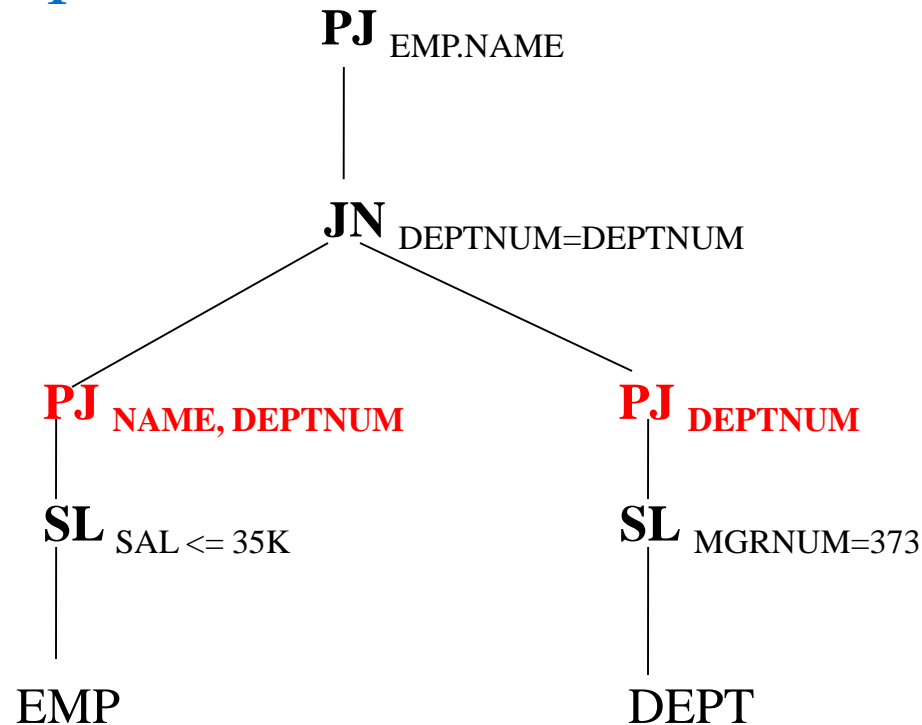
# Simplification

After Applying criterion – 2



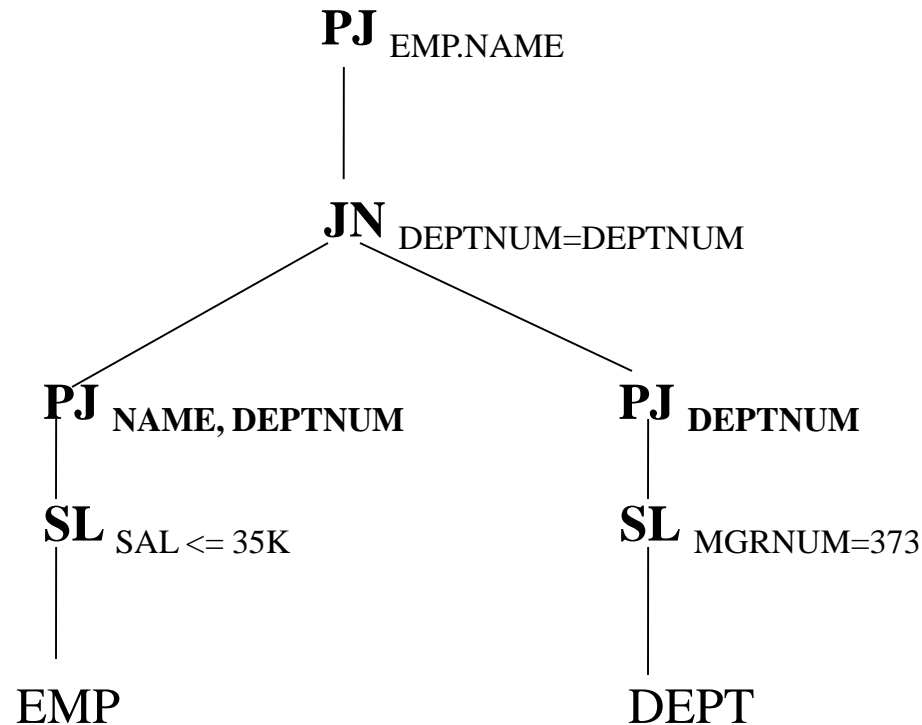
# Simplification

After Applying criterion – 1



# Transformed Query

**Q<sub>T</sub>: PJ<sub>EMP.NAME</sub> ((PJ<sub>NAME,DEPTNUM</sub> SL<sub>SAL<=35K</sub> EMP) JN<sub>DEPTNUM=DEPTNUM</sub> (PJ<sub>DEPTNUM</sub> SL<sub>MGRNUM=373</sub> DEPT))**



# Transformed Query

**Output:**

**Q<sub>T</sub>: PJ**<sub>EMP.NAME</sub> ((**PJ**<sub>NAME,DEPTNUM</sub> **SL**<sub>SAL<=35K</sub> *EMP*) **JN**<sub>DEPTNUM=DEPTNUM</sub>  
(**PJ**<sub>DEPTNUM</sub> **SL**<sub>MGRNUM=373</sub> *DEPT*))

**Input:**

**Q: PJ**<sub>EMP.NAME</sub> ((*EMP* **JN**<sub>DEPTNUM=DEPTNUM</sub> **SL**<sub>MGRNUM=373</sub> *DEPT*)  
**DF** (**SL**<sub>SAL > 35K</sub> *EMP* **JN**<sub>DEPTNUM=DEPTNUM</sub> **SL**<sub>MGRNUM=373</sub> *DEPT* ))

$$Q \longleftrightarrow Q_T$$

# Example 2.1

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

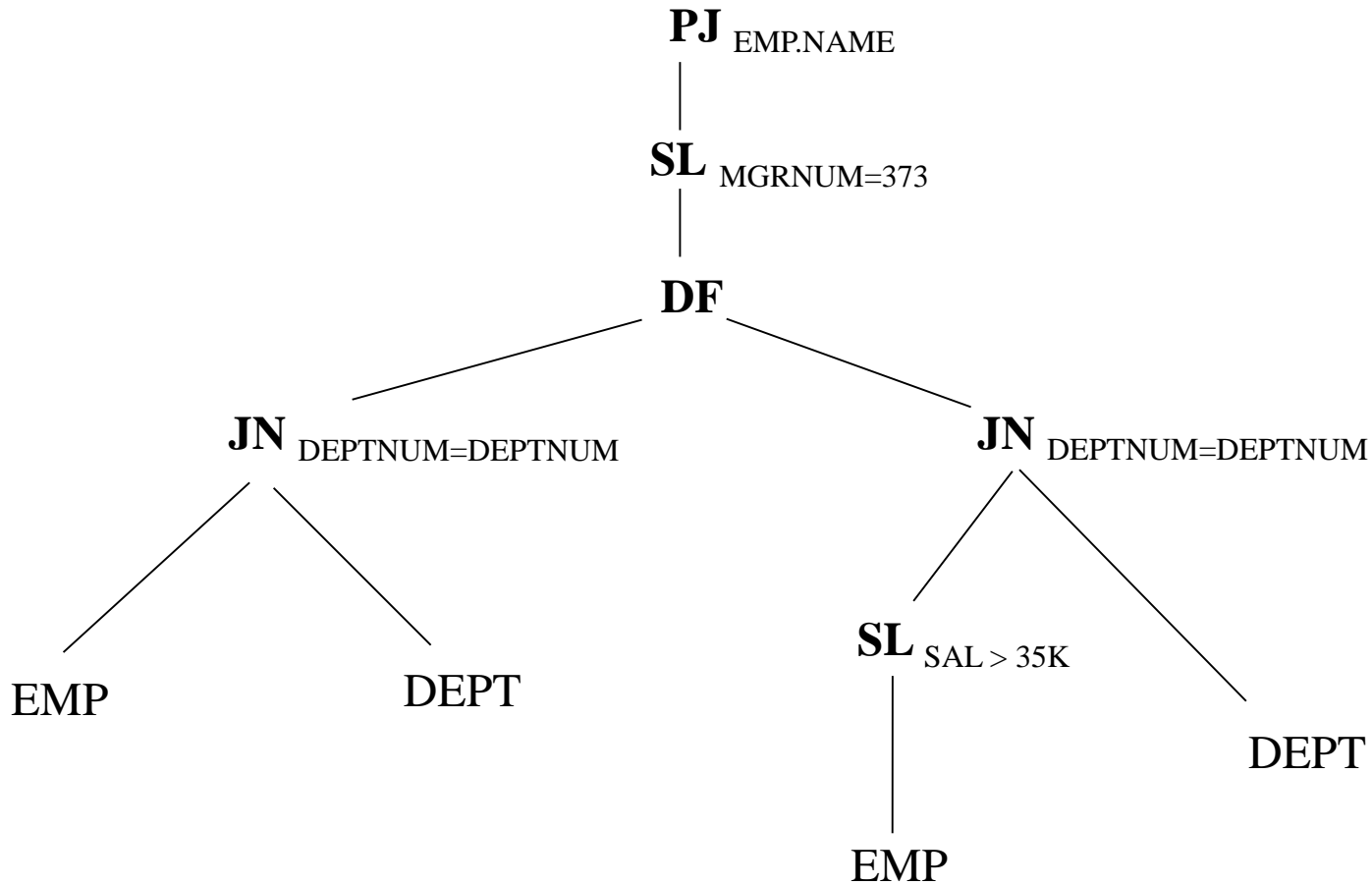
**Q: PJ** *EMP.NAME* **SL** *MGRNUM=373* ((*EMP JN* *DEPTNUM=DEPTNUM* *DEPT*) **DF** (**SL** *SAL > 35K* *EMP JN* *DEPTNUM=DEPTNUM* *DEPT*))

Now, answer the following questions.

- i. Draw the *operator tree*. [2]
- ii. Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [5]

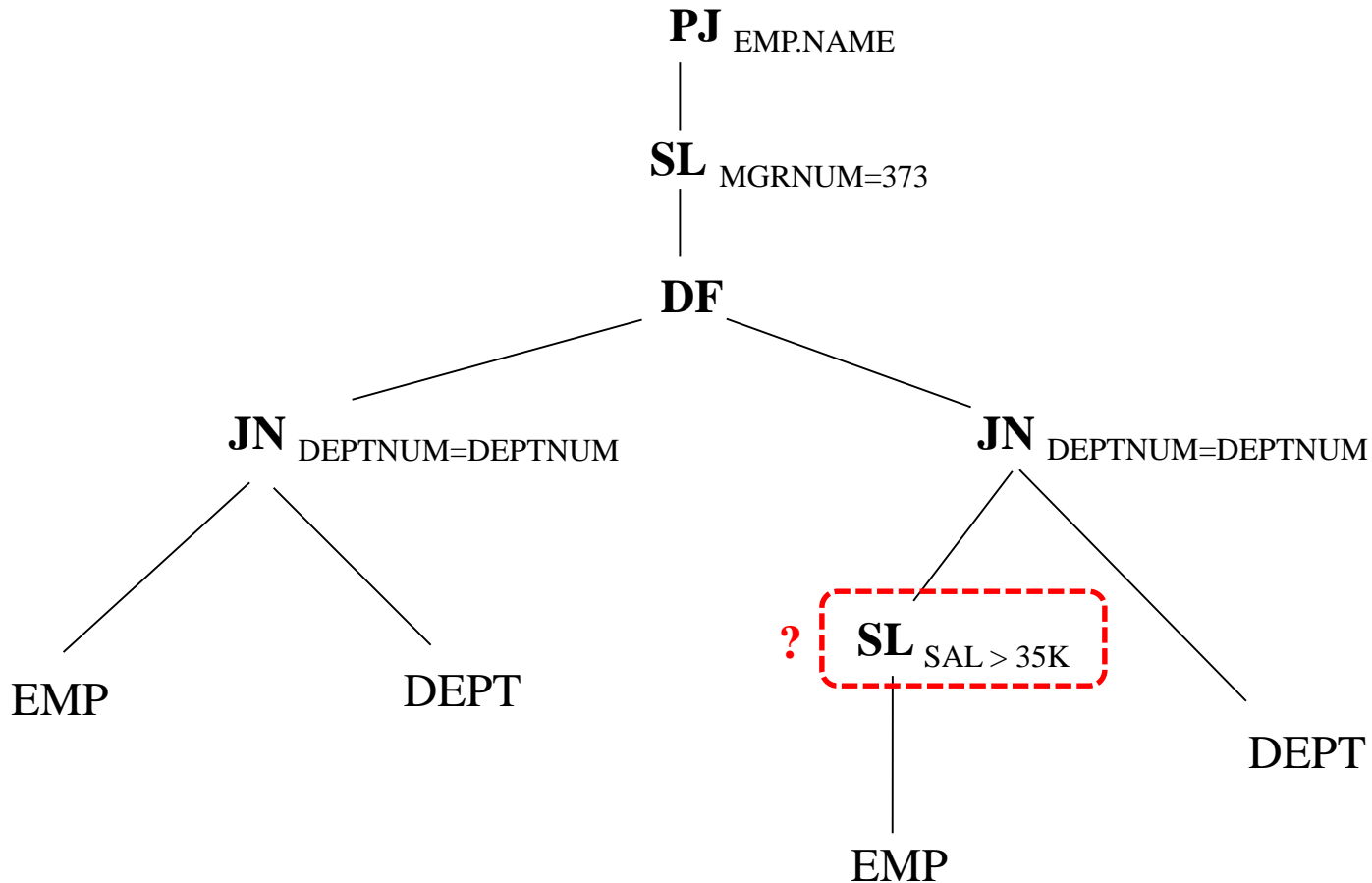


# Operator Tree



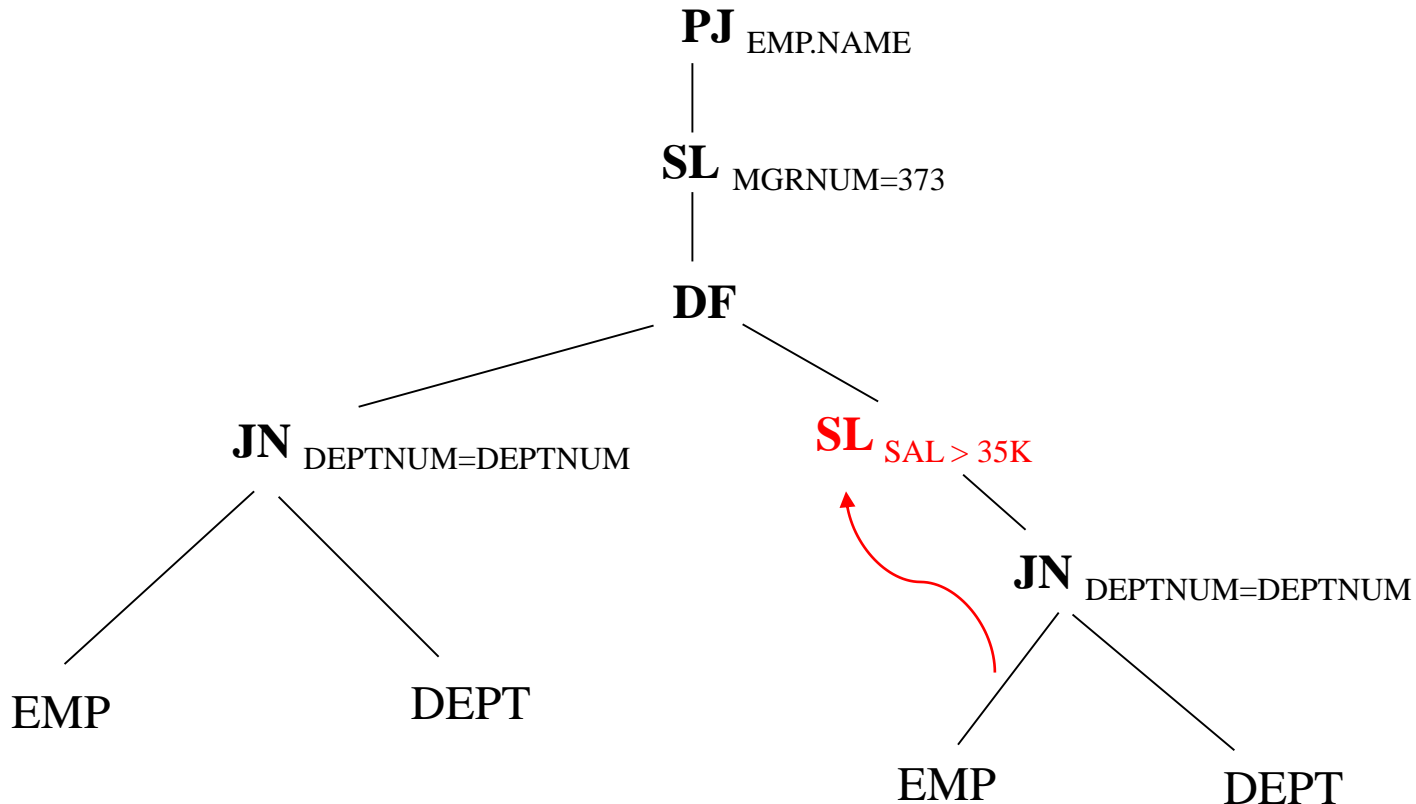
# Finding Common Sub-expression

Any common portion?



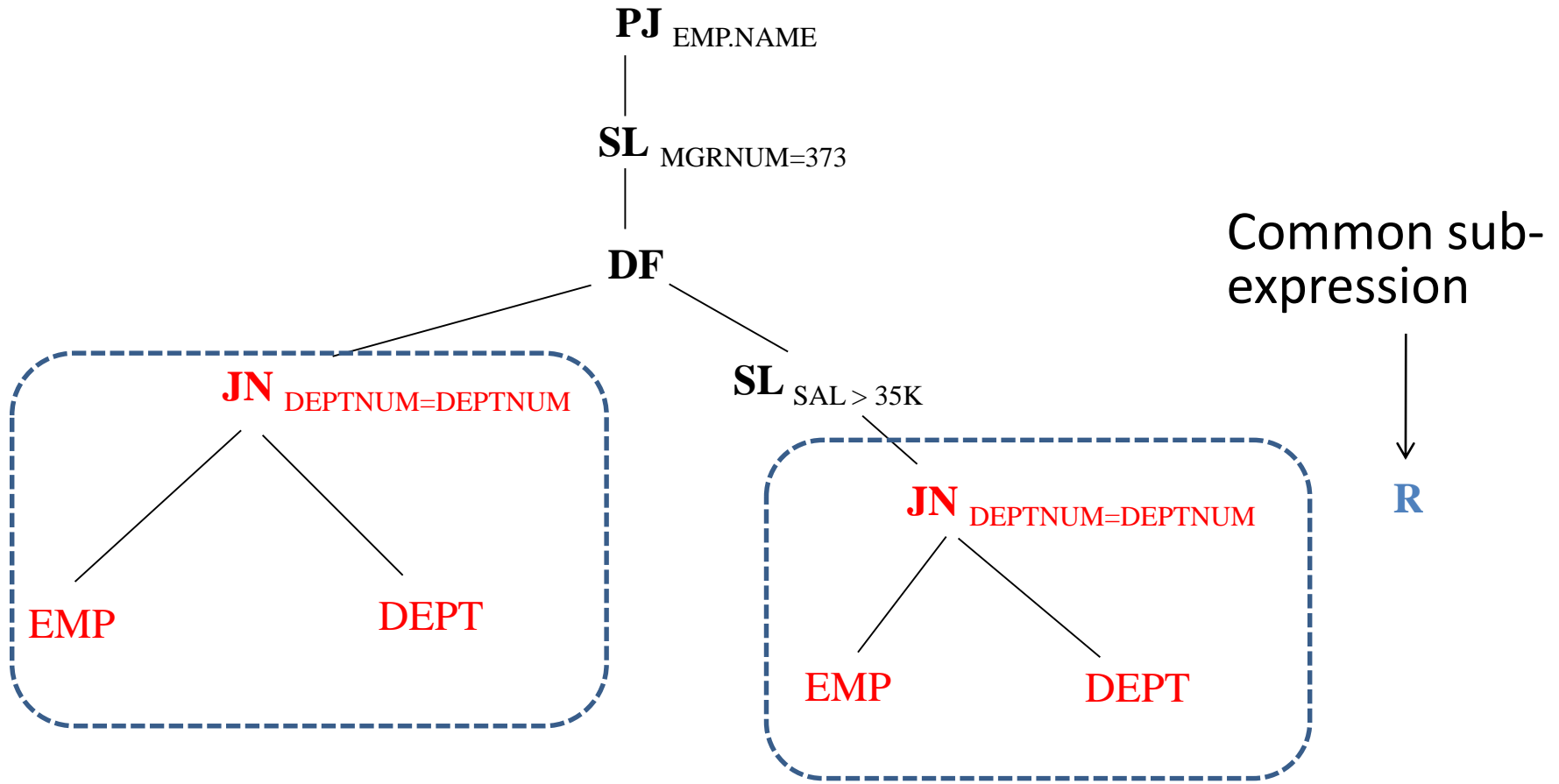
# Finding Common Sub-expression

Any common portion? NOW?

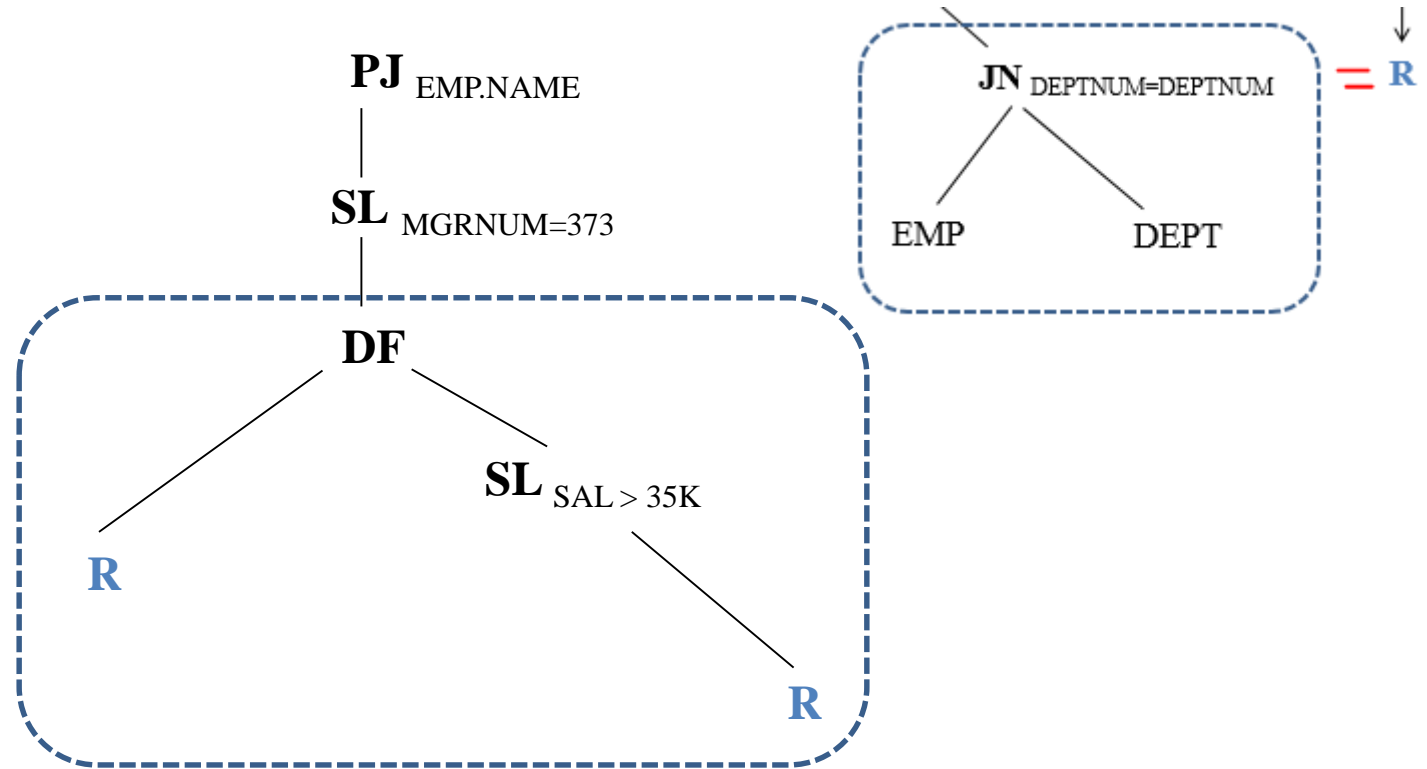


# Finding Common Sub-expression

Any common portion? NOW?

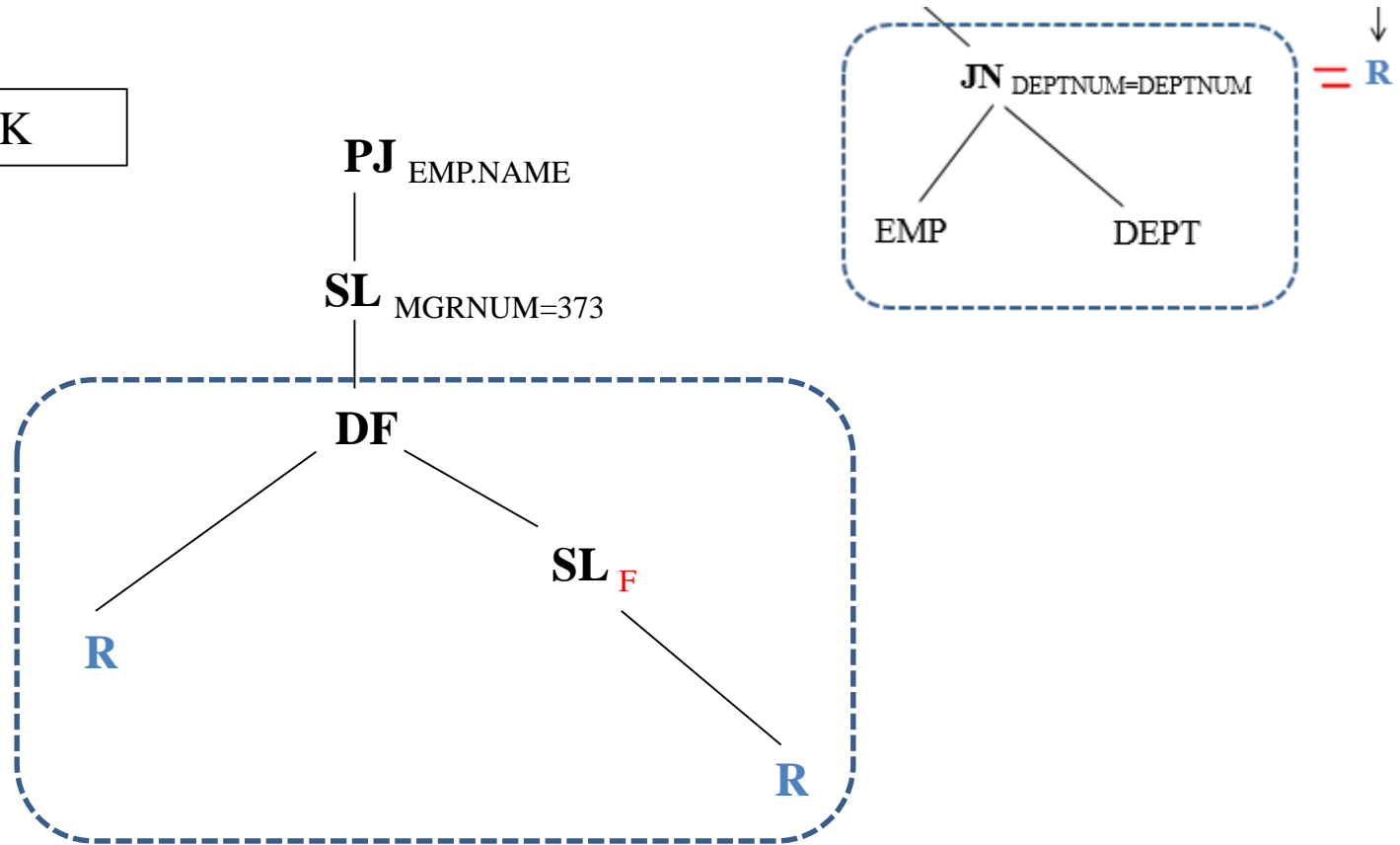


# Finding Common Sub-expression



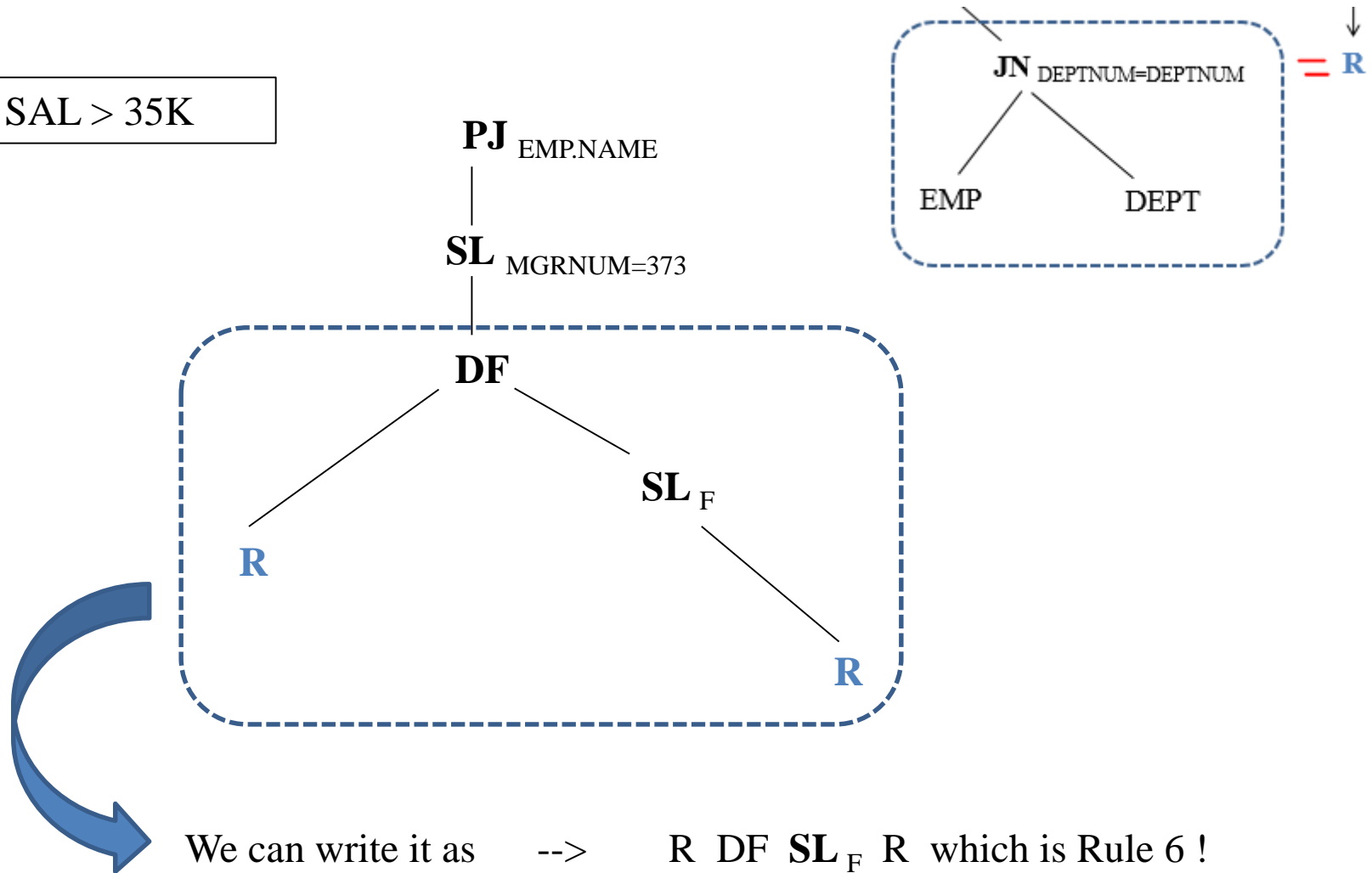
# Finding Common Sub-expression

$F = \text{SAL} > 35K$



# Finding Common Sub-expression

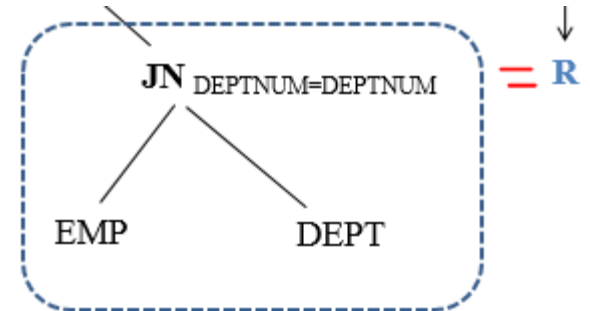
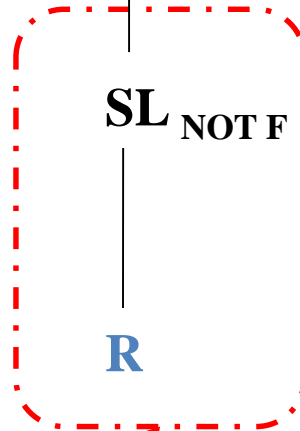
$F = \text{SAL} > 35K$



# Removing Common Sub-expression

$F = \text{SAL} > 35\text{K}$

**PJ** EMP.NAME  
|  
**SL** MGRNUM=373

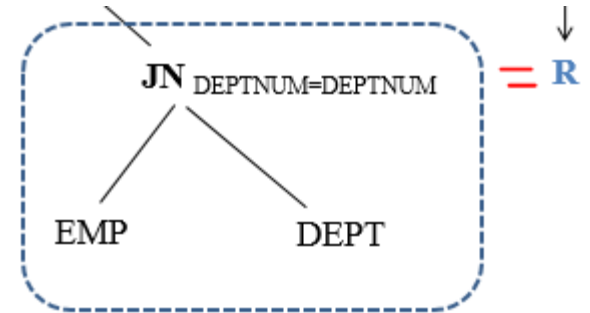
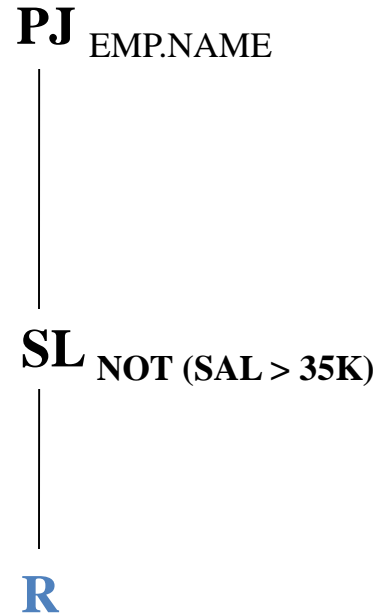


•  $R \text{ DF } SL_F R \leftrightarrow SL_{NOT F} R$

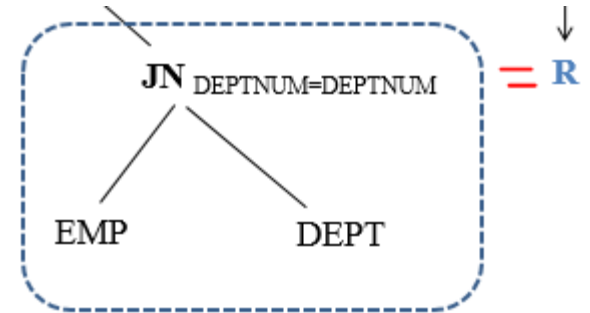
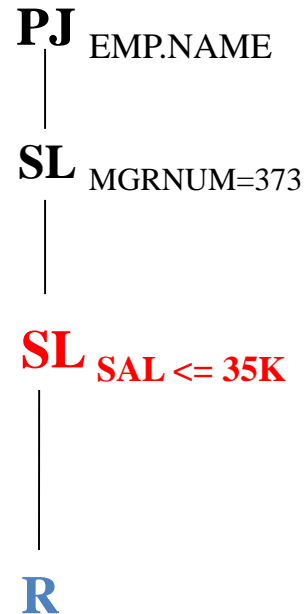


# Removing Common Sub-expression

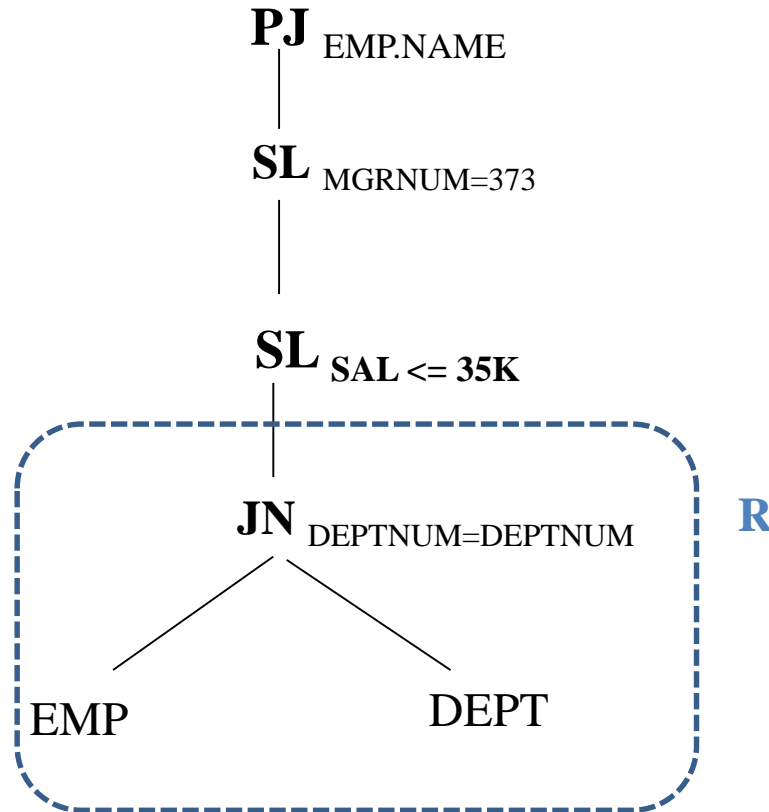
$F = \text{SAL} > 35\text{K}$



# Removing Common Sub-expression



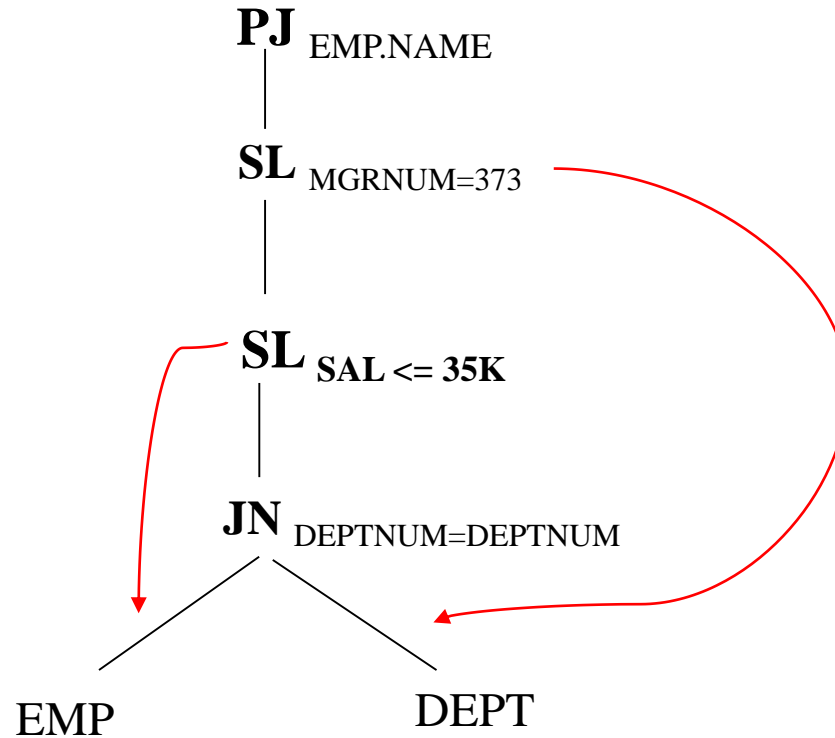
# Removing Common Sub-expression



Can you apply Criterion 1 and/or 2 on this tree?

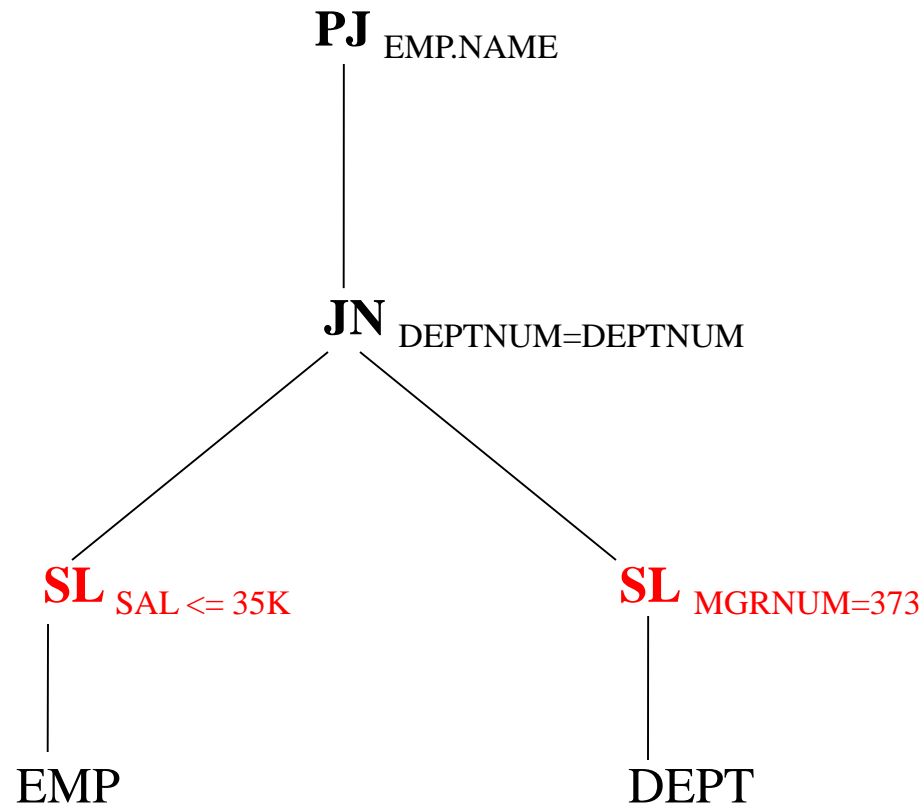
# Simplification

Applying **Criterion 2** -



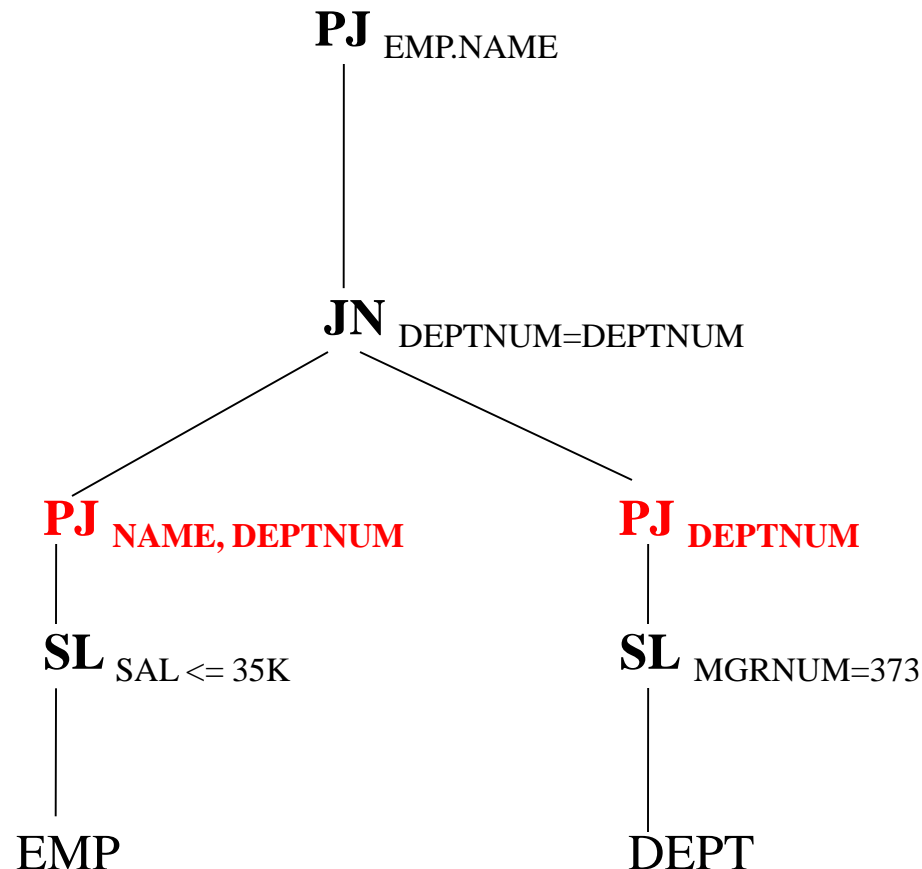
# Simplification

After Applying criterion – 2



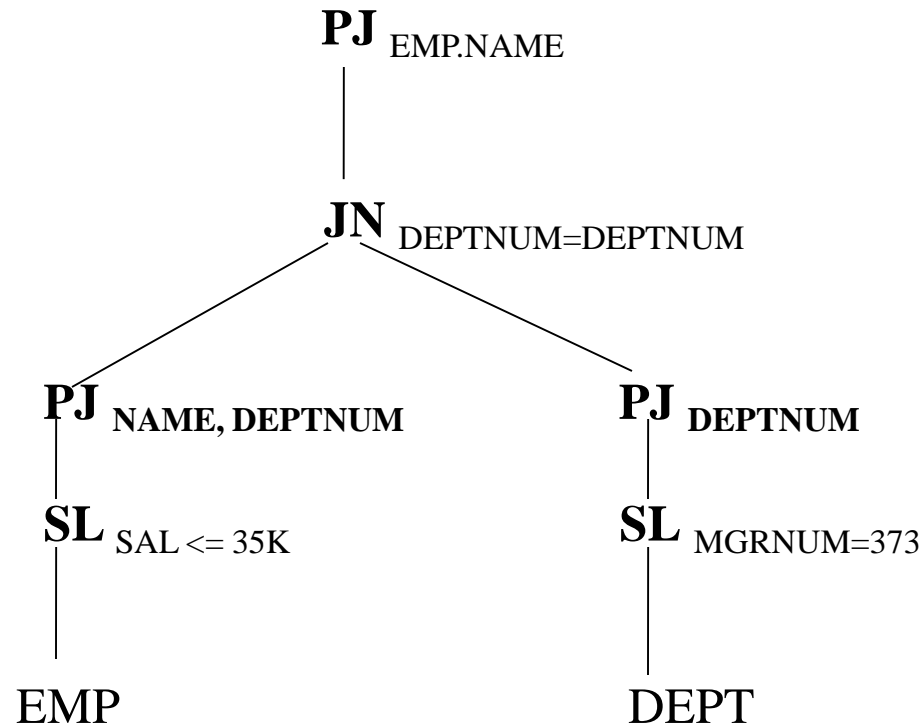
# Simplification

After Applying criterion – 1



# Transformed Query

**Q<sub>T</sub>: PJ<sub>EMP.NAME</sub> ((PJ<sub>NAME,DEPTNUM</sub> SL<sub>SAL<=35K</sub> EMP) JN<sub>DEPTNUM=DEPTNUM</sub> (PJ<sub>DEPTNUM</sub> SL<sub>MGRNUM=373</sub> DEPT))**



# Transformed Query

**Output:**

**Q<sub>T</sub>:** **PJ**<sub>EMP.NAME</sub> ((**PJ**<sub>NAME,DEPTNUM</sub> **SL**<sub>SAL≤35K</sub> *EMP*) **JN**<sub>DEPTNUM=DEPTNUM</sub>  
(**PJ**<sub>DEPTNUM</sub> **SL**<sub>MGRNUM=373</sub> *DEPT*))

**Input:**

**Q:** **PJ**<sub>EMP.NAME</sub> **SL**<sub>MGRNUM=373</sub> ((*EMP* **JN**<sub>DEPTNUM=DEPTNUM</sub> *DEPT*)  
**DF** (**SL**<sub>SAL > 35K</sub> *EMP* **JN**<sub>DEPTNUM=DEPTNUM</sub> *DEPT*))

$$Q \longleftrightarrow Q_T$$



# Example 2.2

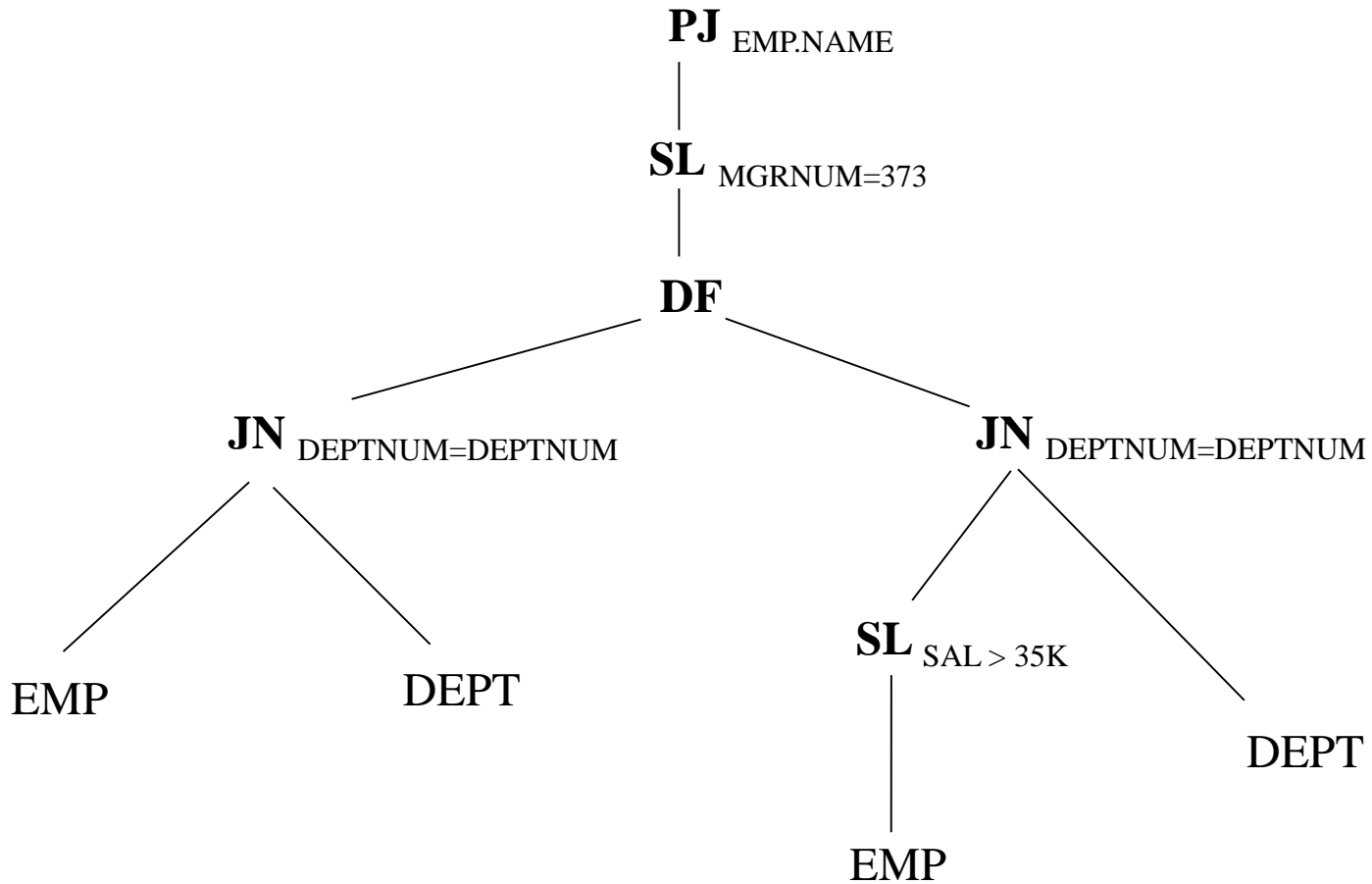
*EMP (EMPNUM, DEPTNUM, NAME, SAL, AGE)*  
*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

**Q: PJ** *EMP.NAME* **SL** *MGRNUM=373* ((*EMP JN* *DEPTNUM=DEPTNUM* *DEPT*) **DF** (*SL* *SAL > 35K* *EMP JN* *DEPTNUM=DEPTNUM* *DEPT*))

Now, answer the following questions.

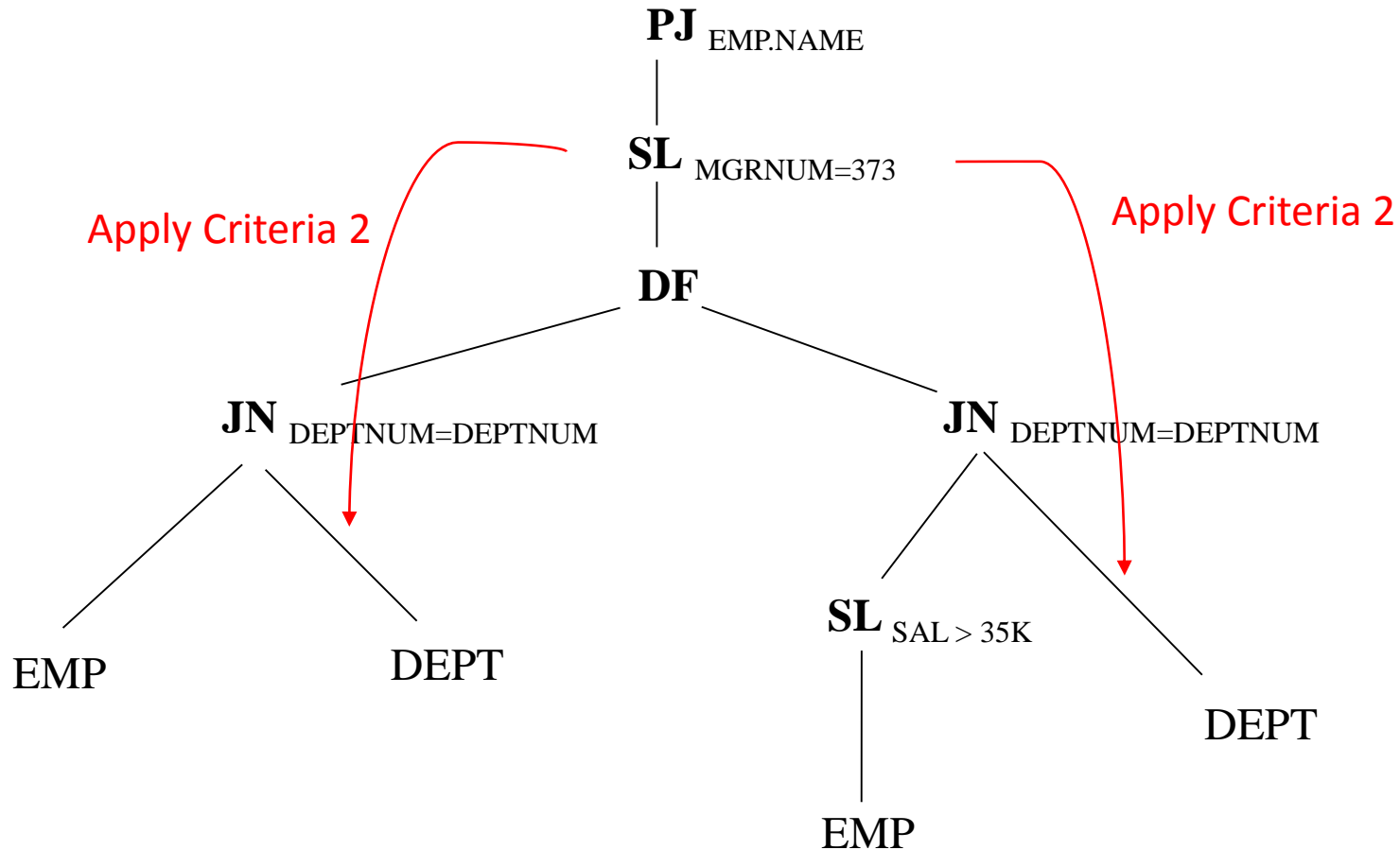
- i. Draw the *operator tree*. [2]
- ii. Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [5]

# Operator Tree



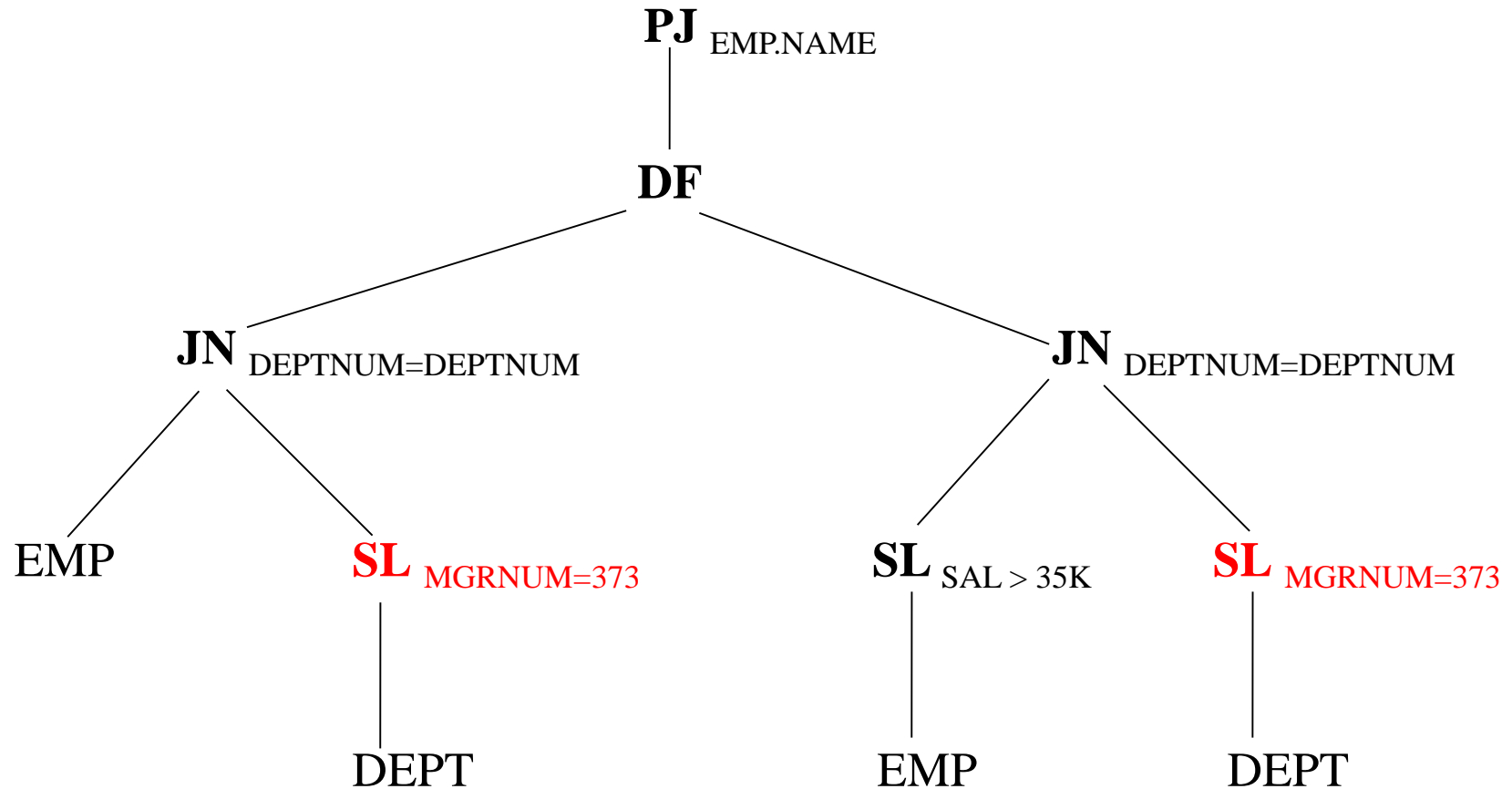
# Finding Common Sub-expression

Any common portion?



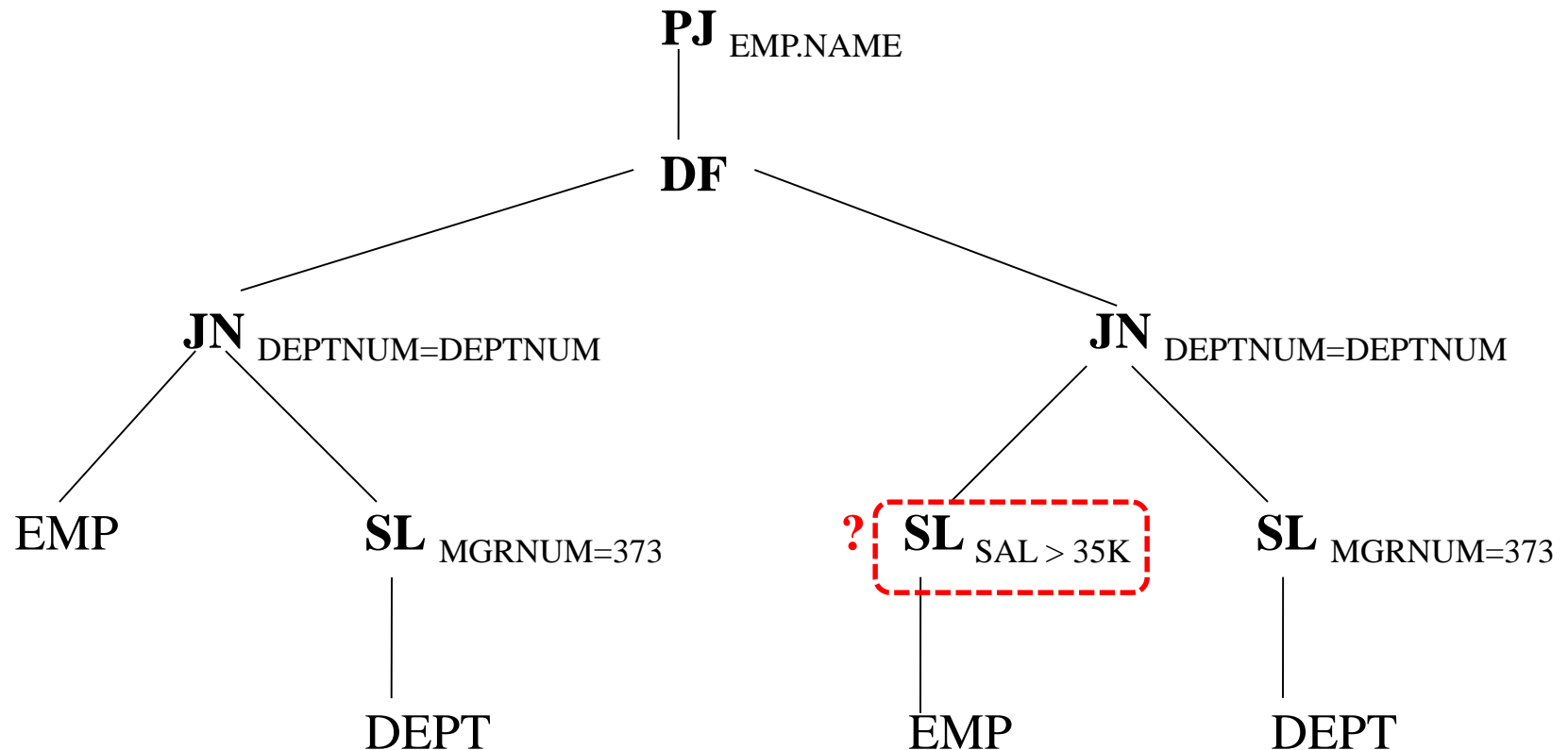
# Finding Common Sub-expression

Any common portion?



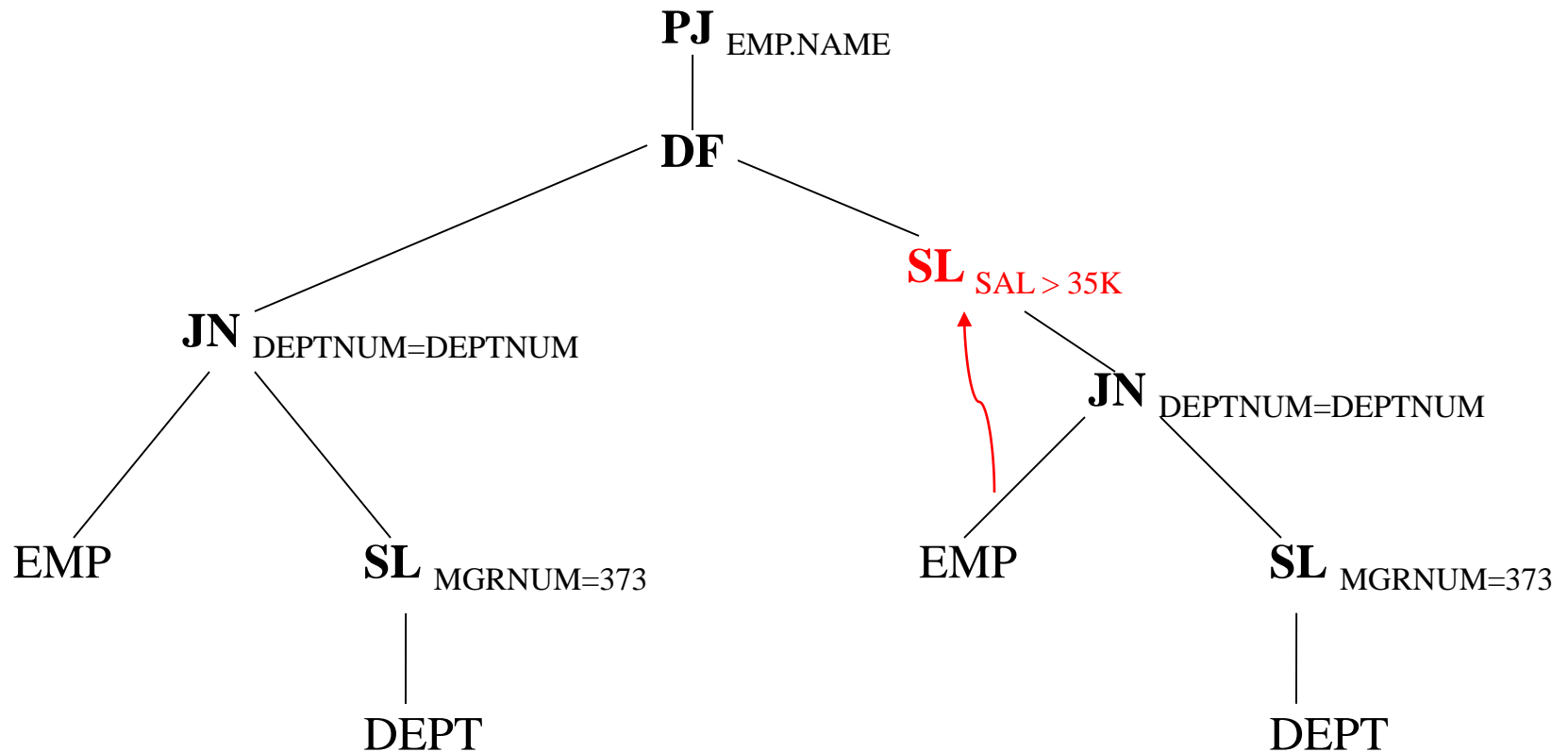
# Finding Common Sub-expression

Any common portion?

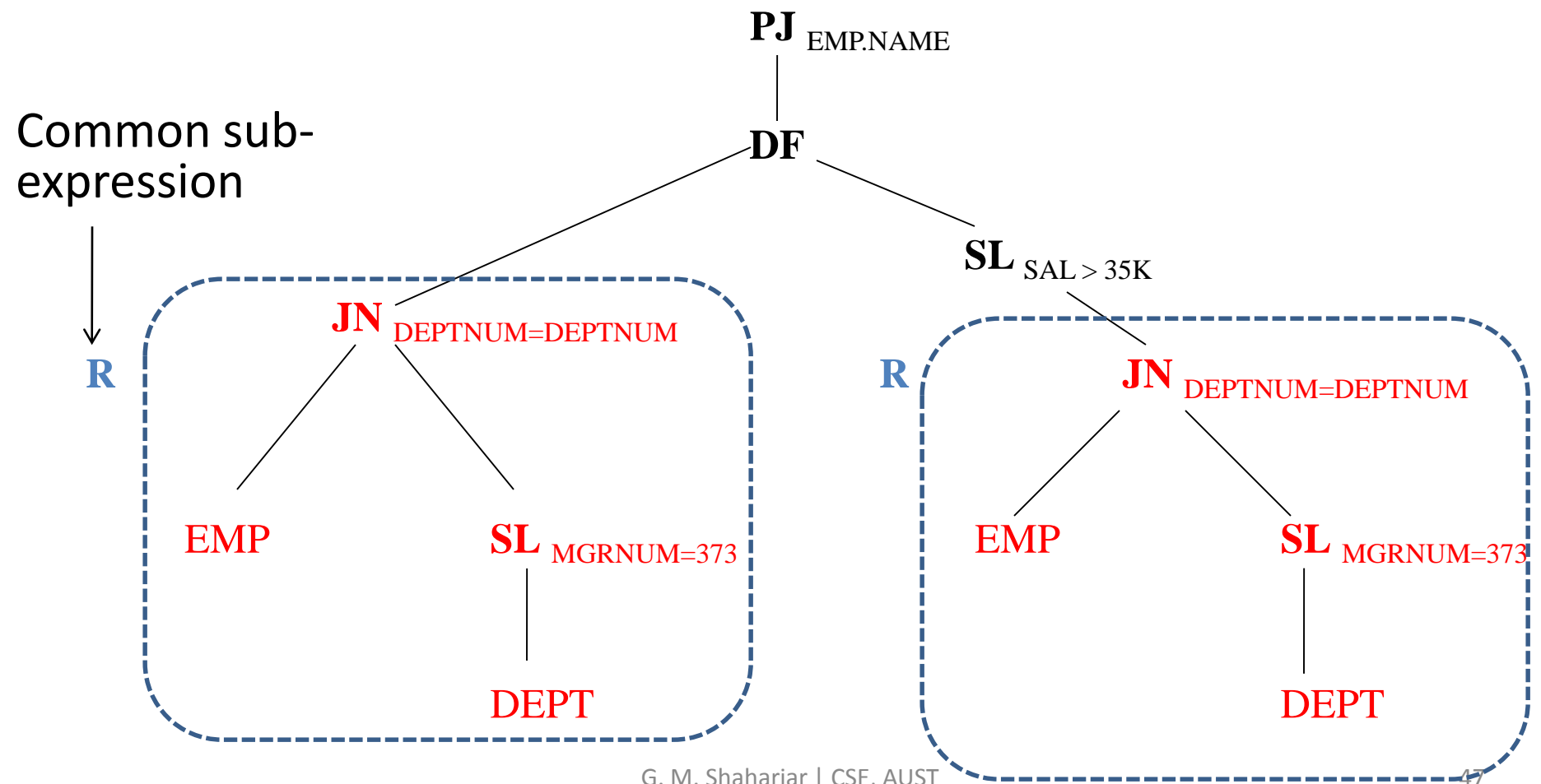


# Finding Common Sub-expression

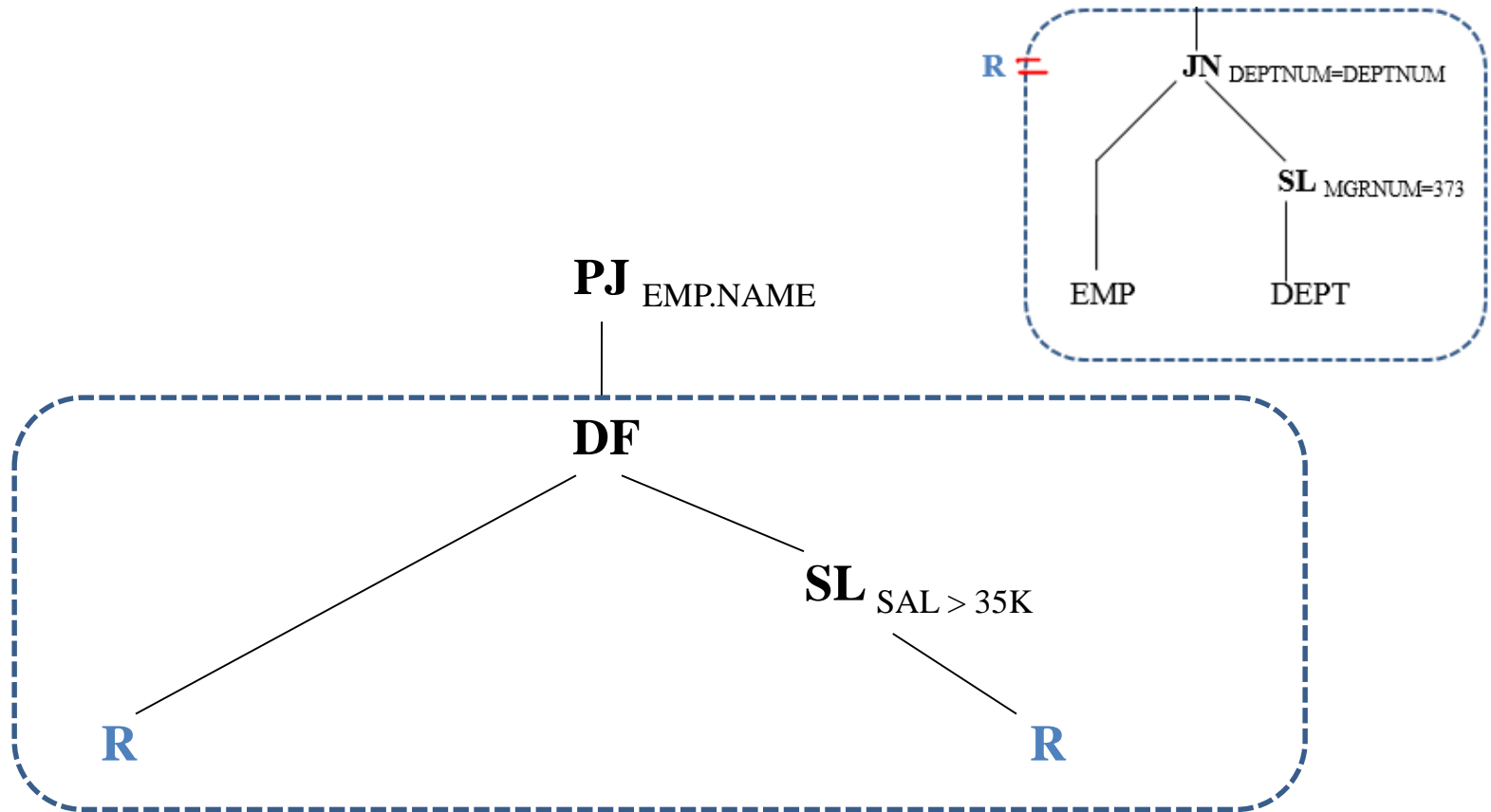
Any common portion? NOW?



# Finding Common Sub-expression



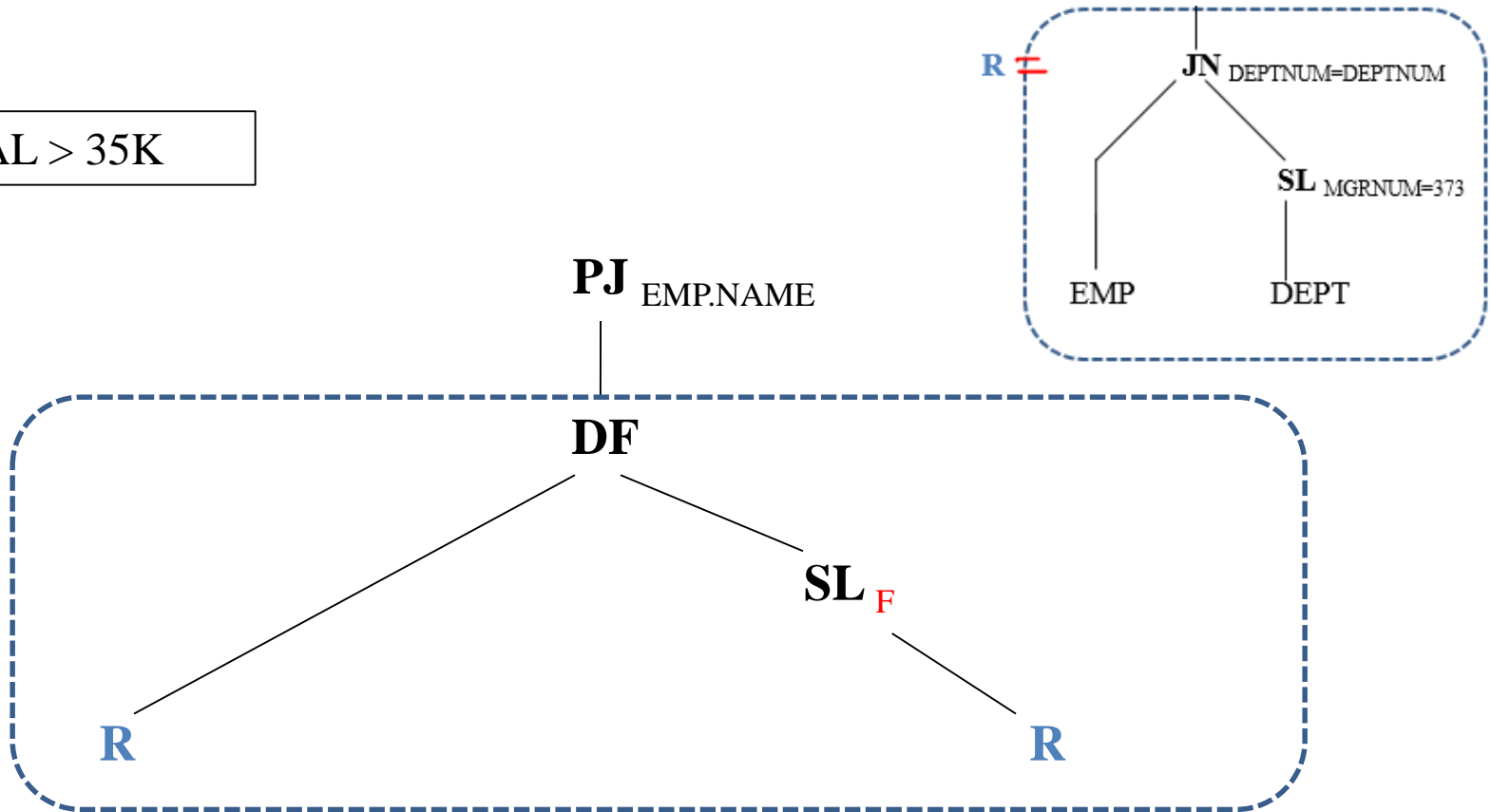
# Finding Common Sub-expression





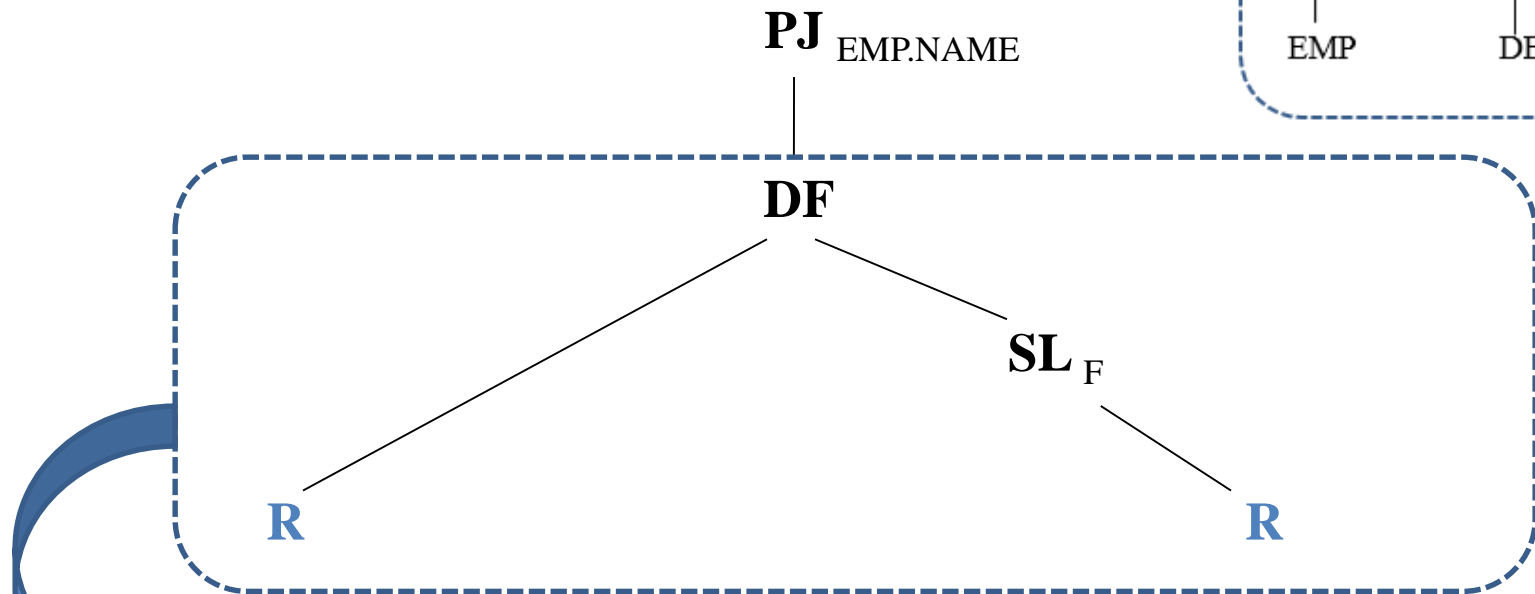
# Finding Common Sub-expression

$F = \text{SAL} > 35\text{K}$



# Finding Common Sub-expression

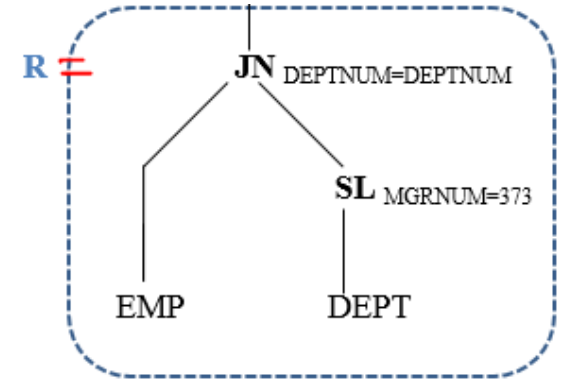
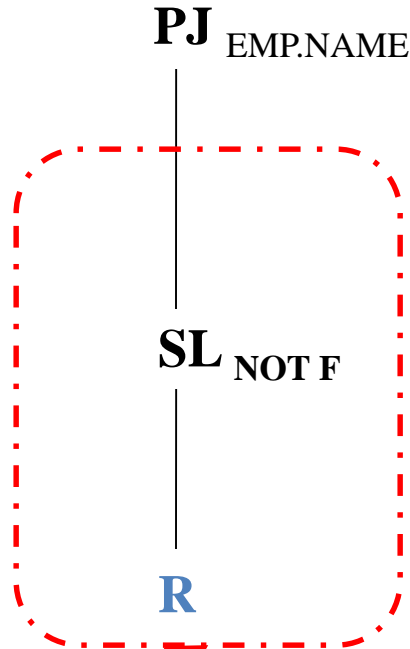
$F = \text{SAL} > 35\text{K}$



We can write it as  $--> R \text{ DF } \text{SL}_F R$  which is Rule 6 !

# Removing Common Sub-expression

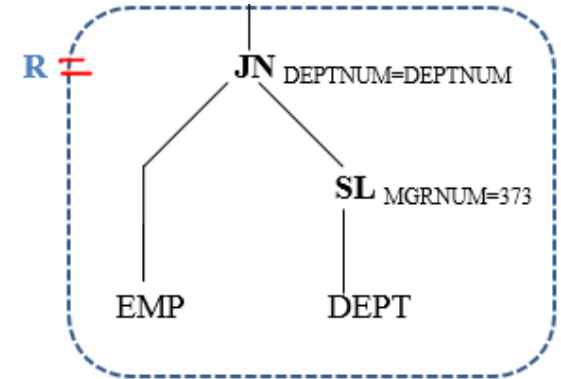
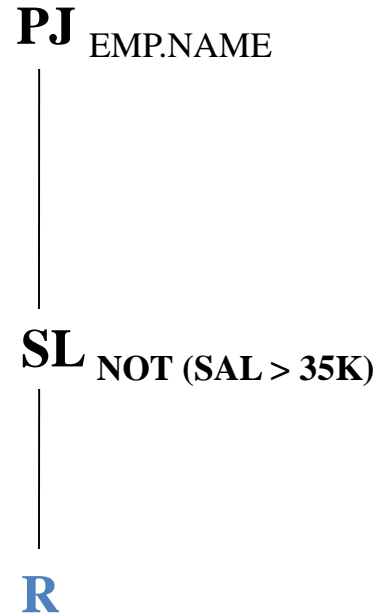
$F = \text{SAL} > 35\text{K}$



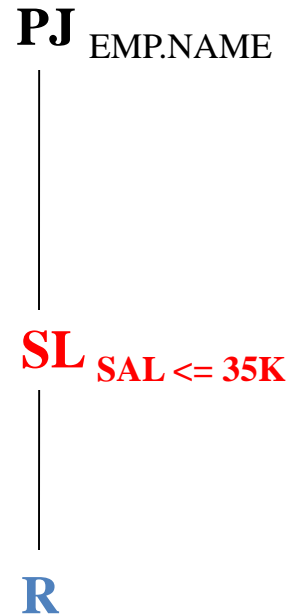
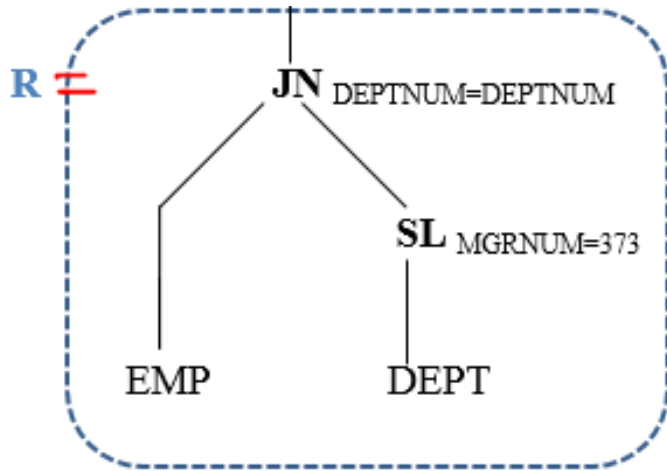
•  $R \text{ DF } SL_F R \leftrightarrow SL_{NOT F} R$

# Removing Common Sub-expression

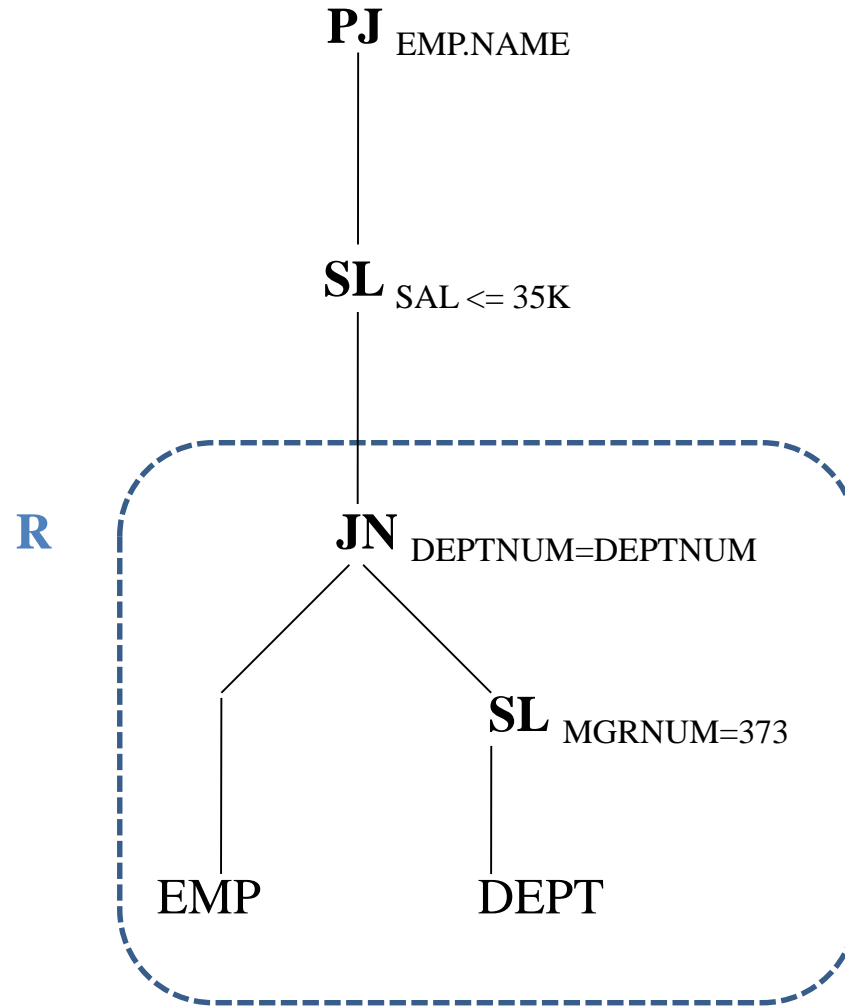
$F = \text{SAL} > 35\text{K}$



# Removing Common Sub-expression



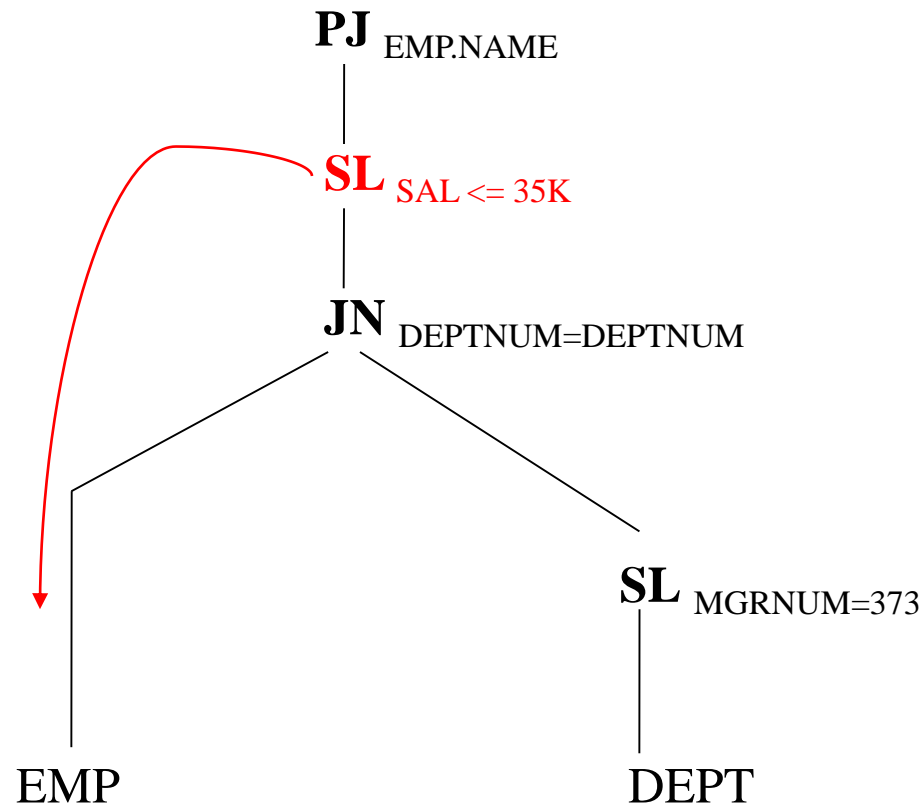
# Removing Common Sub-expression



Can you apply Criterion 1 and/or 2 on this tree?

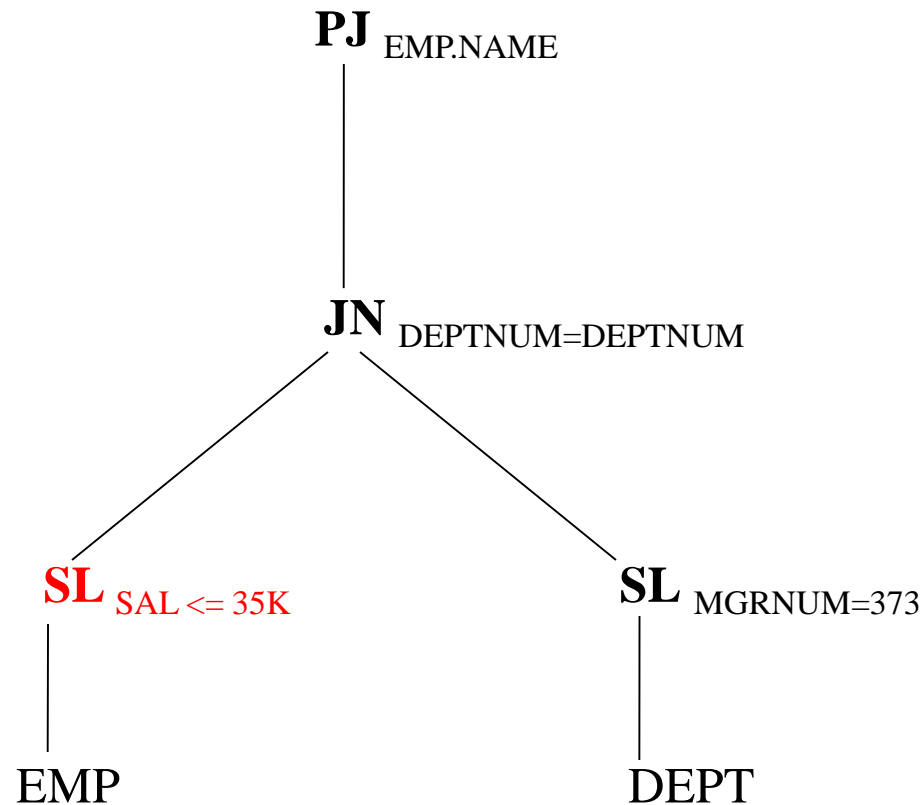
# Simplification

## Applying criterion – 2



# Simplification

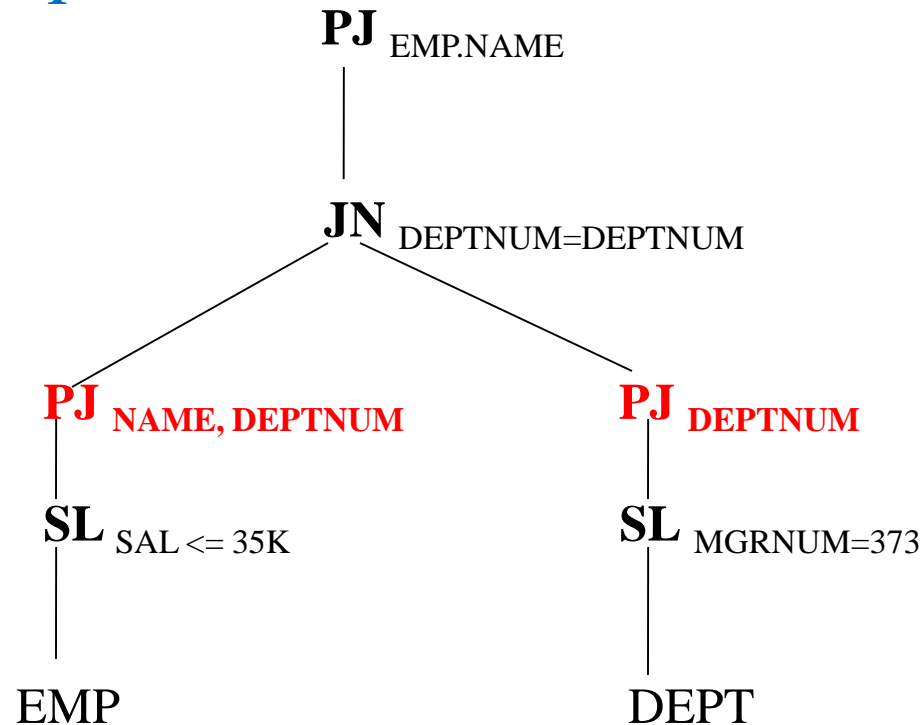
After Applying criterion – 2





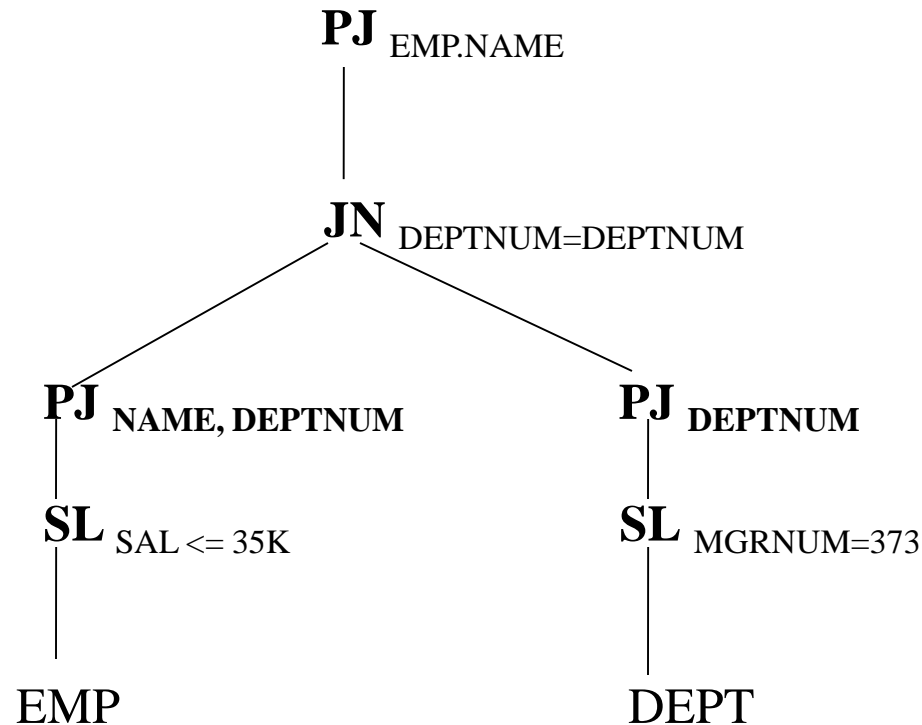
# Simplification

After Applying criterion – 1



# Transformed Query

**Q<sub>T</sub>: PJ<sub>EMP.NAME</sub> ((PJ<sub>NAME,DEPTNUM</sub> SL<sub>SAL<=35K</sub> EMP) JN<sub>DEPTNUM=DEPTNUM</sub> (PJ<sub>DEPTNUM</sub> SL<sub>MGRNUM=373</sub> DEPT))**



# Transformed Query

**Output:**

**Q<sub>T</sub>:** **PJ**<sub>EMP.NAME</sub> ((**PJ**<sub>NAME,DEPTNUM</sub> **SL**<sub>SAL≤35K</sub> *EMP*) **JN**<sub>DEPTNUM=DEPTNUM</sub>  
(**PJ**<sub>DEPTNUM</sub> **SL**<sub>MGRNUM=373</sub> *DEPT*))

**Input:**

**Q:** **PJ**<sub>EMP.NAME</sub> **SL**<sub>MGRNUM=373</sub> ((*EMP* **JN**<sub>DEPTNUM=DEPTNUM</sub> *DEPT*)  
**DF** (**SL**<sub>SAL > 35K</sub> *EMP* **JN**<sub>DEPTNUM=DEPTNUM</sub> *DEPT*))

$$Q \longleftrightarrow Q_T$$

# Example 3

## Practise

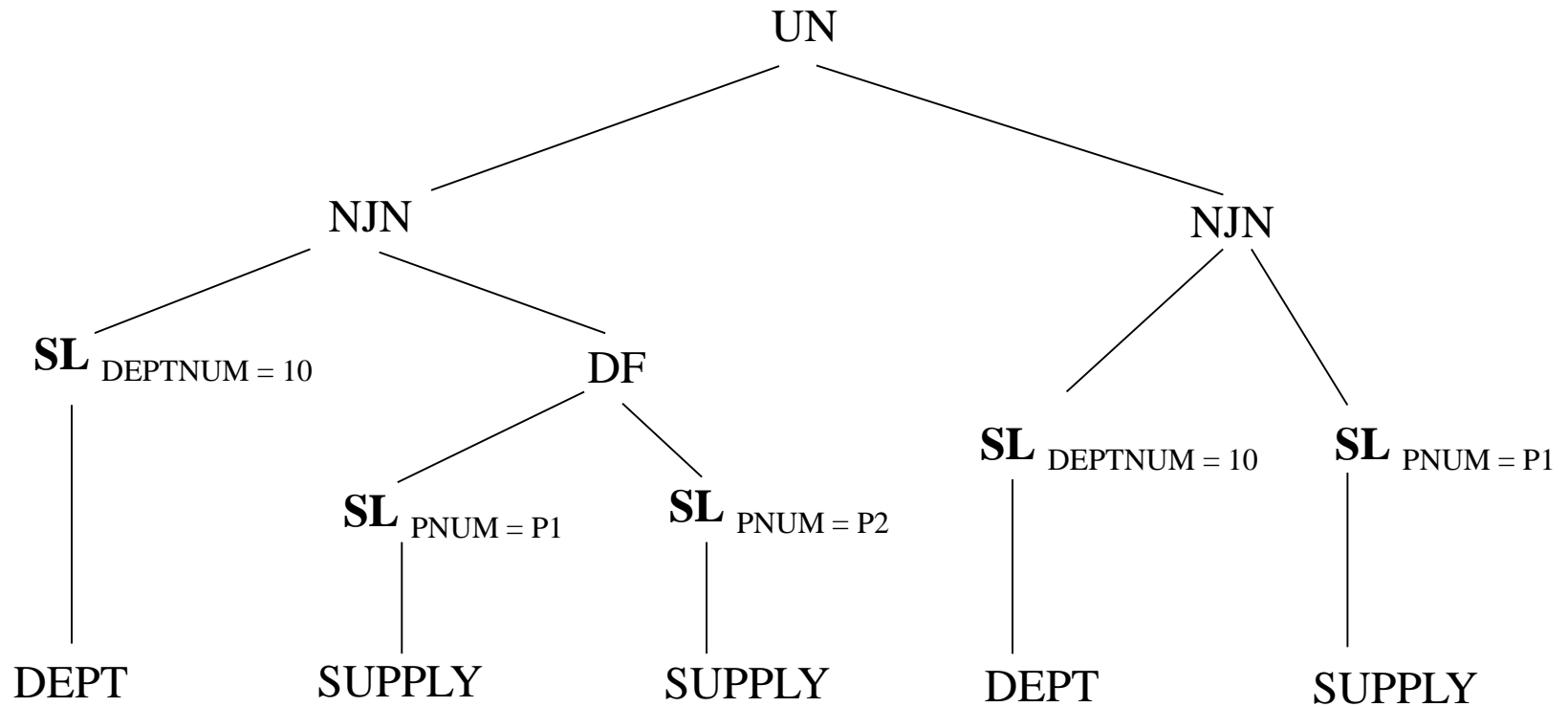
\* Draw Operator Tree for the following queries:

①  
SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)  
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Query:

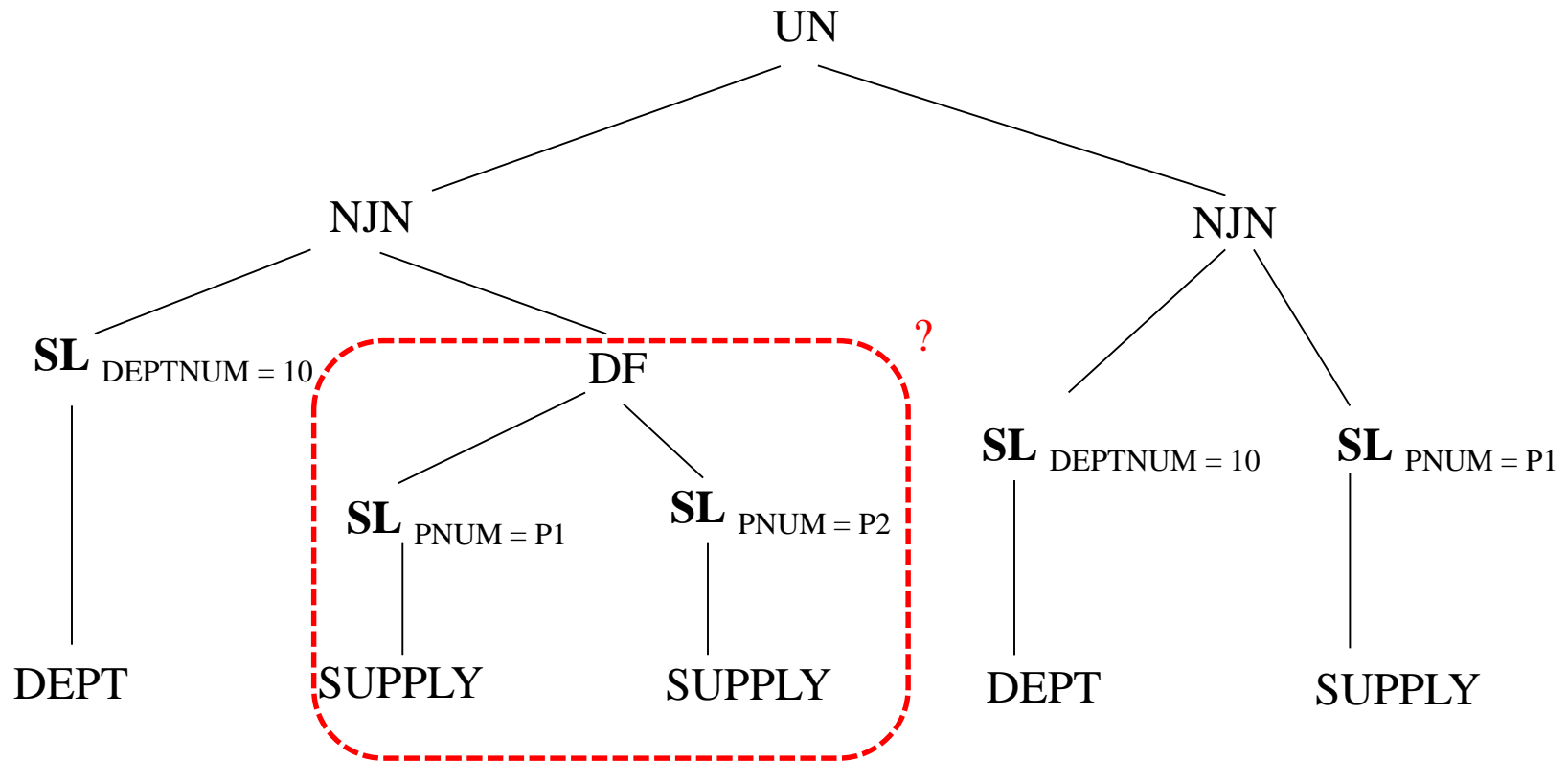
$(\sigma_{DEPTNUM=10} DEPT \bowtie (\sigma_{PNUM="P_1"} SUPPLY$   
 $\bowtie \sigma_{PNUM="P_2"} SUPPLY)) \cup (\sigma_{DEPTNUM=10} DEPT$   
 $\bowtie \sigma_{PNUM="P_1"} SUPPLY)$

# Operator Tree



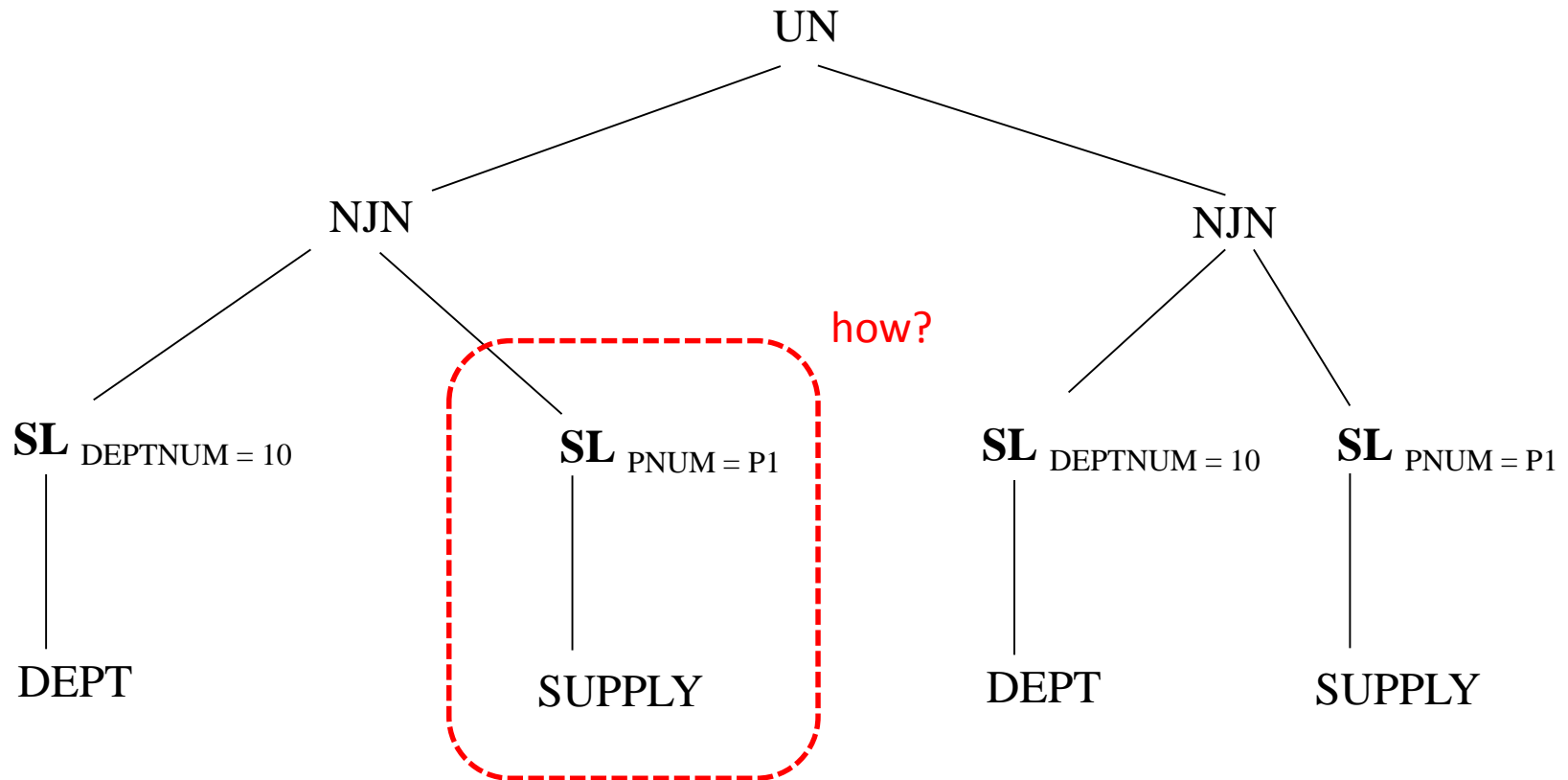
# Finding Common Sub - Expression

Any common portion?



# Finding Common Sub - Expression

Any common portion?



# How?

SUPPLY

SNUM	PNUM	DEPTNUM	QUAN
1	P1	1	10
2	P2	2	20
3	P1	1	30
4	P2	1	40
5	P1	2	50
6	P2	1	60

A

SL<sub>PNUM = P1</sub> SUPPLY

SNUM	PNUM	DEPTNUM	QUAN
1	P1	1	10
3	P1	1	30
5	P1	2	50

B

SL<sub>PNUM = P2</sub> SUPPLY

SNUM	PNUM	DEPTNUM	QUAN
2	P2	2	20
4	P2	1	40
6	P2	1	60

A

DF

B

=

A

SNUM	PNUM	DEPT NUM	QUAN
1	P1	1	10
3	P1	1	30
5	P1	2	50

-

SNUM	PNUM	DEPT NUM	QUAN
2	P2	2	20
4	P2	1	40
6	P2	1	60

=

SNUM	PNUM	DEPT NUM	QUAN
1	P1	1	10
3	P1	1	30
5	P1	2	50

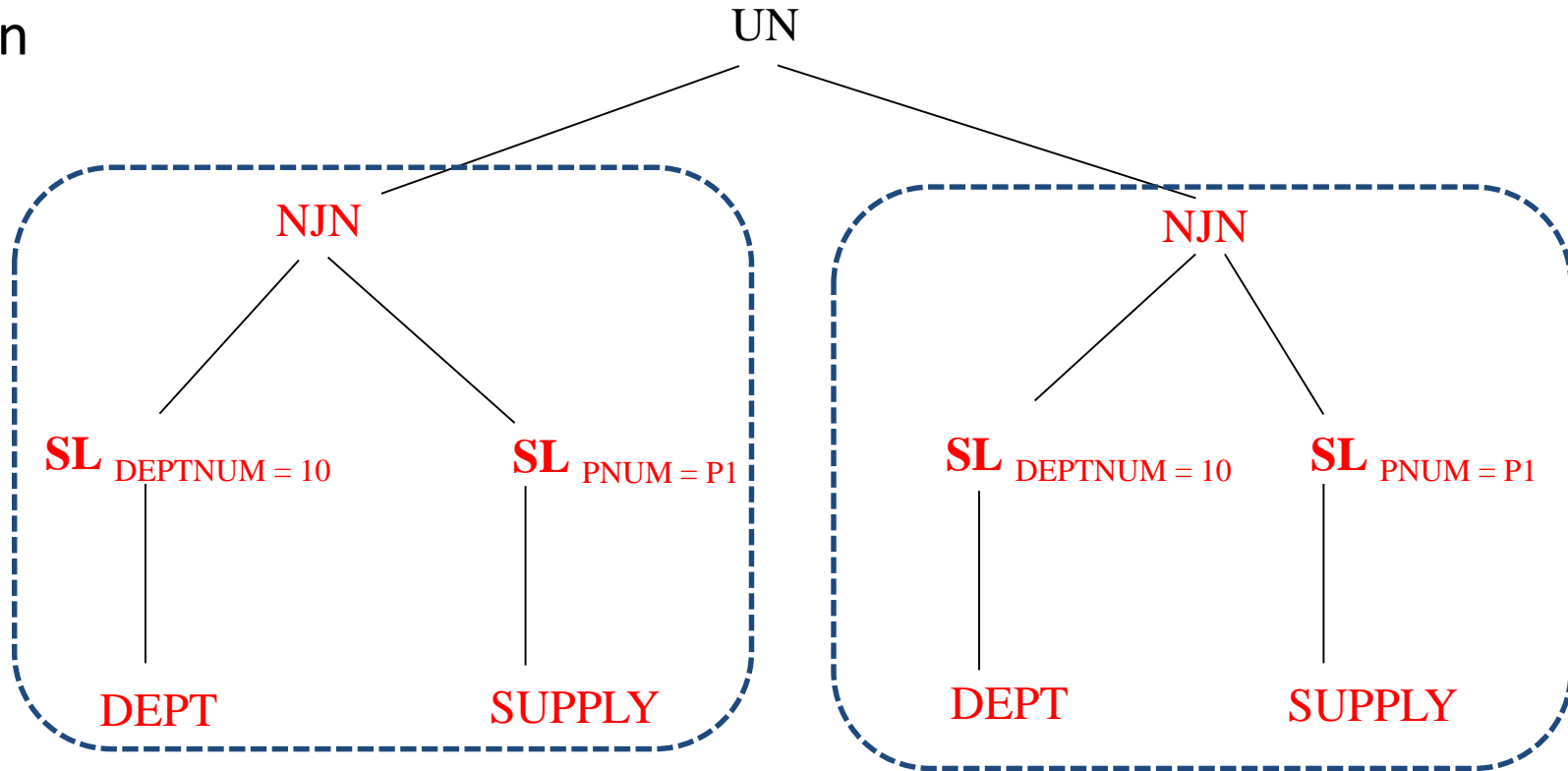


# Finding Common Sub - Expression

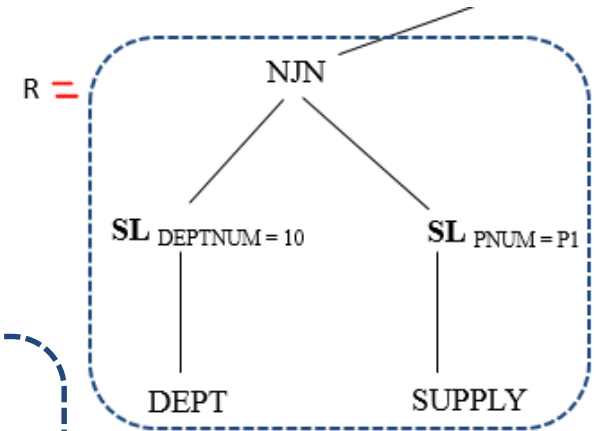
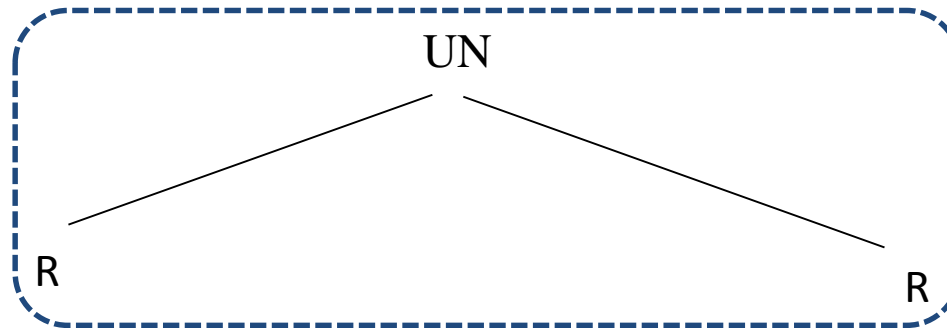
Any common portion? NOW?

Common sub-expression

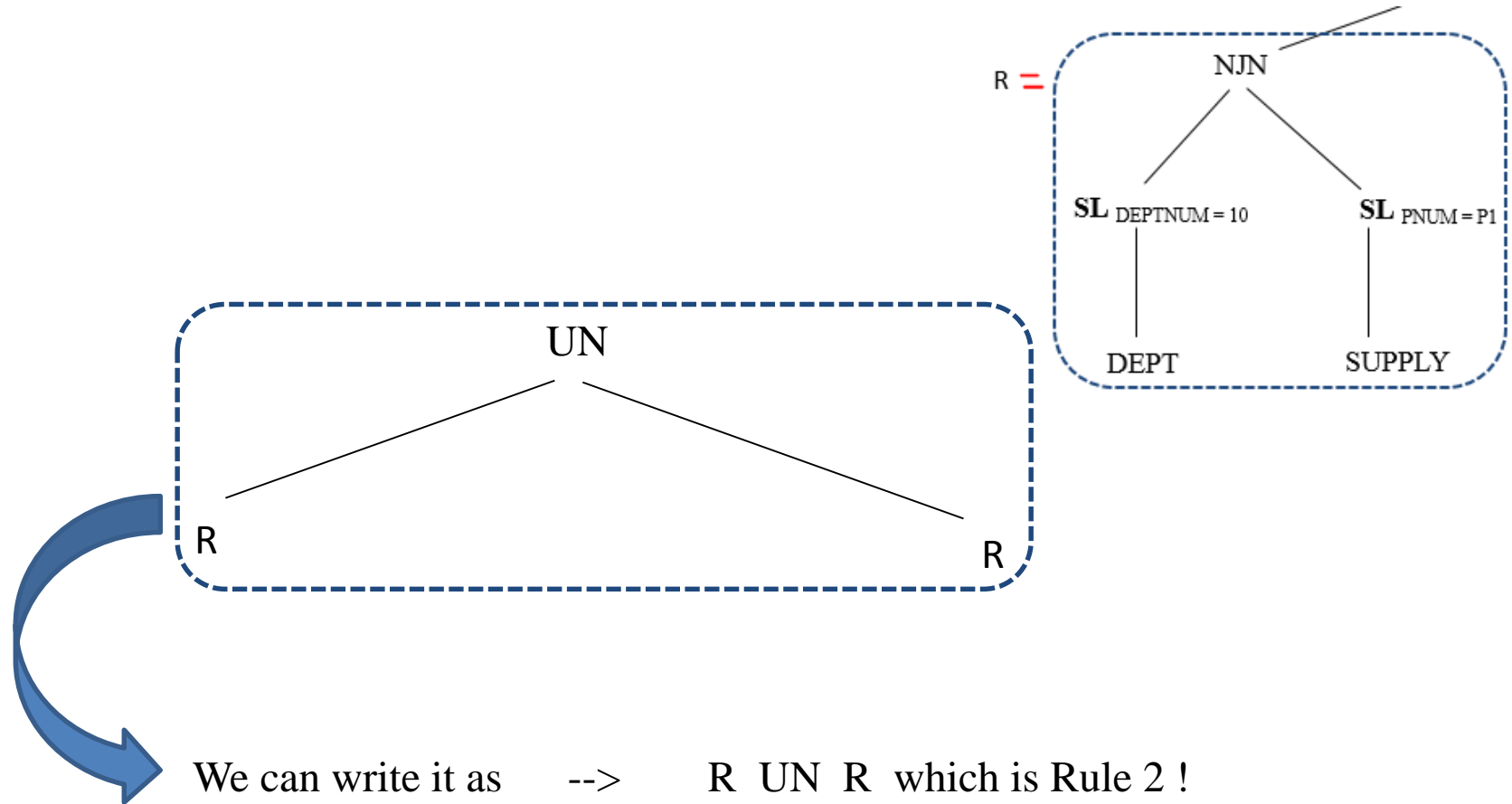
R



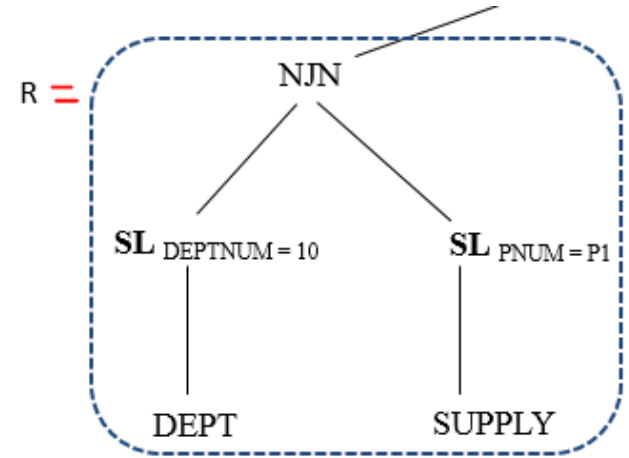
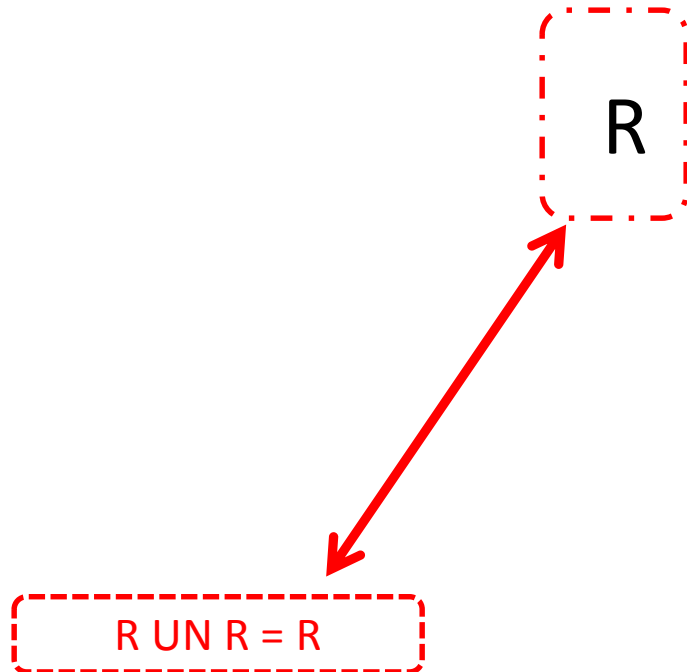
# Finding Common Sub - Expression



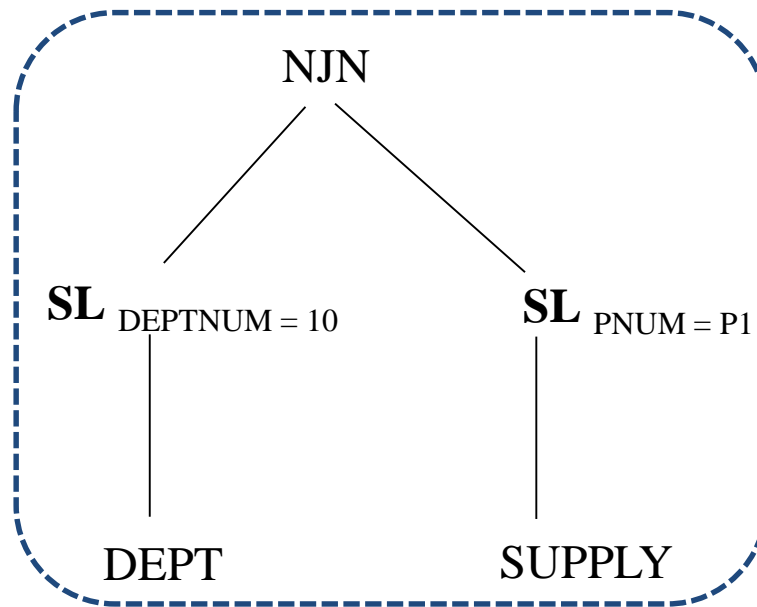
# Finding Common Sub - Expression



# Removing Common Sub - Expression



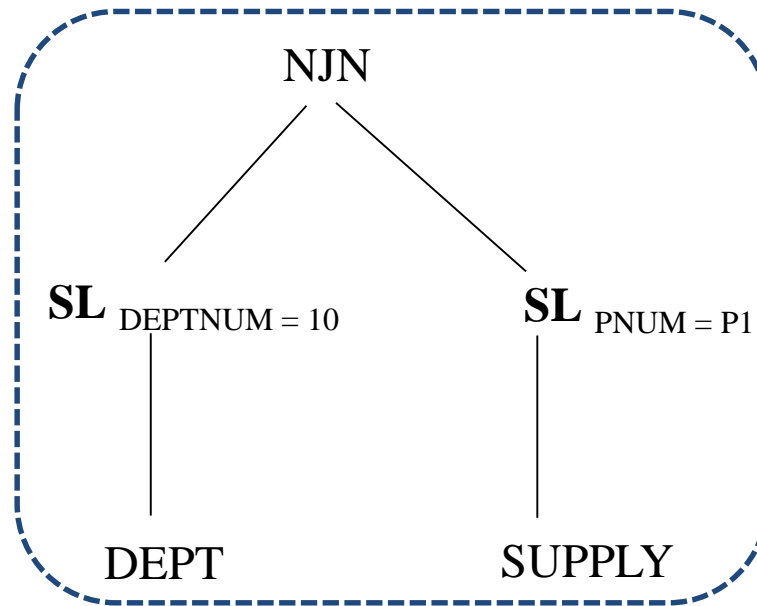
# Removing Common Sub - Expression



Do we need to apply criteria 1 and/or 2? No, already simplified.

# Transformed Query

**Q<sub>T</sub>:** **SL**<sub>DEPTNUM = 10</sub> *DEPT* **NJN** **SL**<sub>PNUM = P1</sub> *SUPPLY*



# Transformed Query

**Output:**

**Q<sub>T</sub>:** **SL**<sub>DEPTNUM = 10</sub> *DEPT* **NJN** **SL**<sub>PNUM = P1</sub> *SUPPLY*

**Input:**

**Q:** (**SL**<sub>DEPTNUM = 10</sub> *DEPT* **NJN** (**SL**<sub>PNUM = P1</sub> *SUPPLY* **DF** **SL**<sub>PNUM = P2</sub> *SUPPLY*))

**UN** (**SL**<sub>DEPTNUM = 10</sub> *DEPT* **NJN** **SL**<sub>PNUM = P1</sub> *SUPPLY*)

$$Q \longleftrightarrow Q_T$$

# Last Example

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

Consider the following global query:

$((SL_{F1} \text{ EMP } \Join_{A=B} \text{ DEPT}) \text{ UN } (SL_{F2} \text{ EMP } \Join_{A=B} \text{ DEPT})) \text{ DF } (SL_{F3} \text{ EMP } \Join_{A=B} \text{ DEPT})$

Here,

F1, F2, F3 can represent any condition. In this example consider none of them are same.

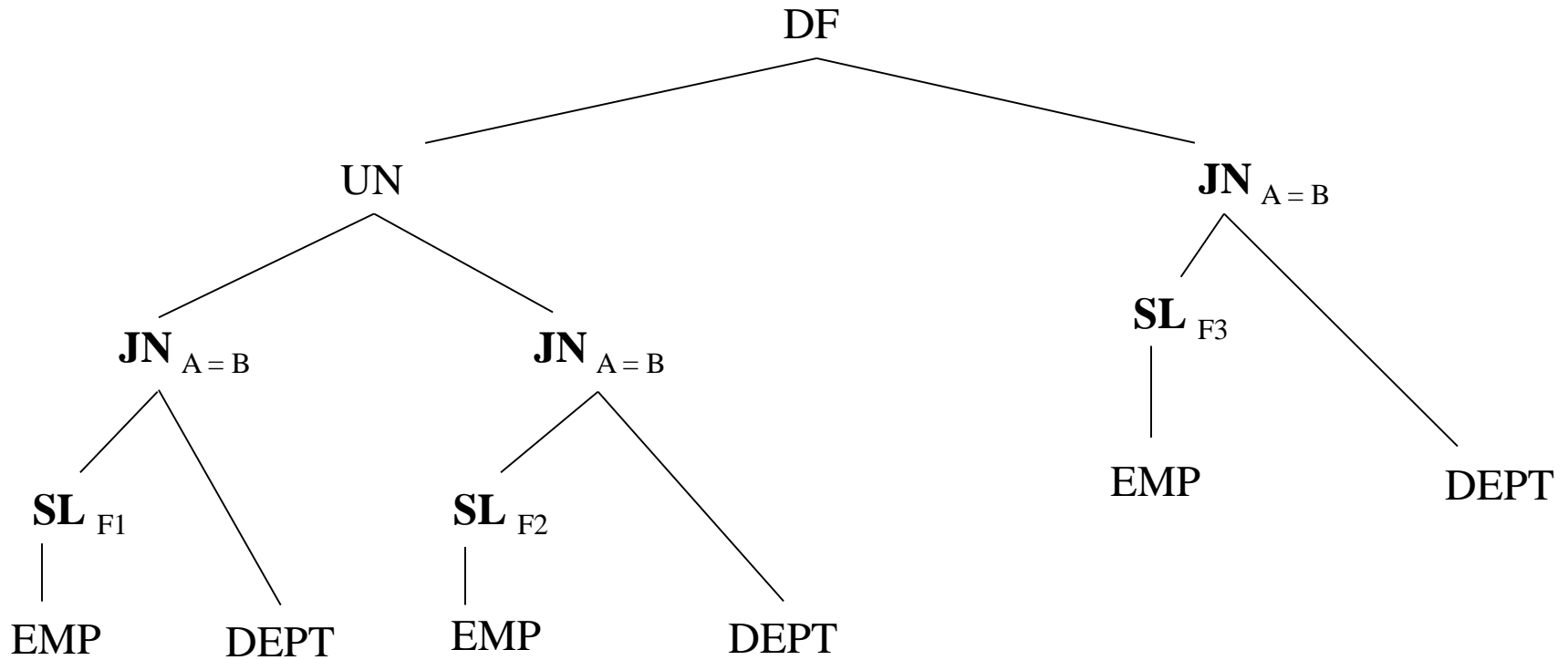
Imagine,  $A = B = \text{DEPTNUM}$

Now, answer the following questions.

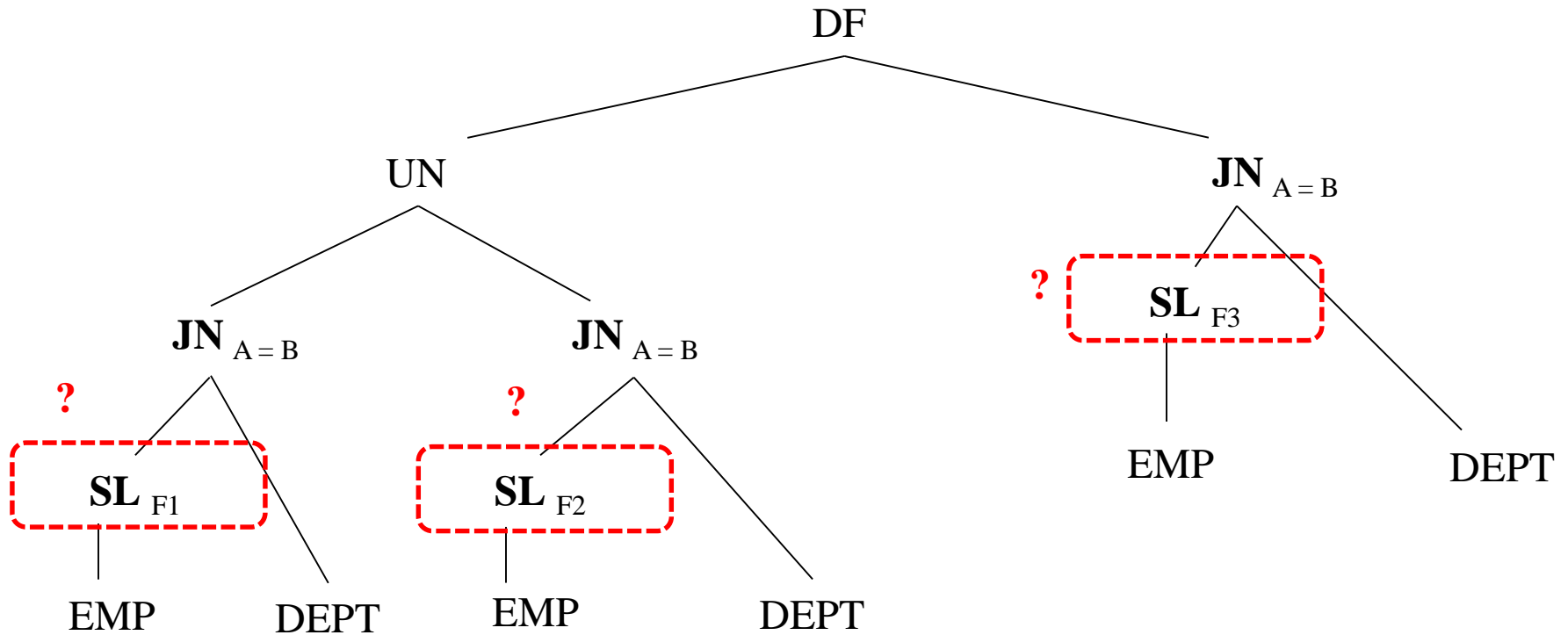
- i. Draw the *operator tree*. [2]
- ii. Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [5]



# Operator Tree

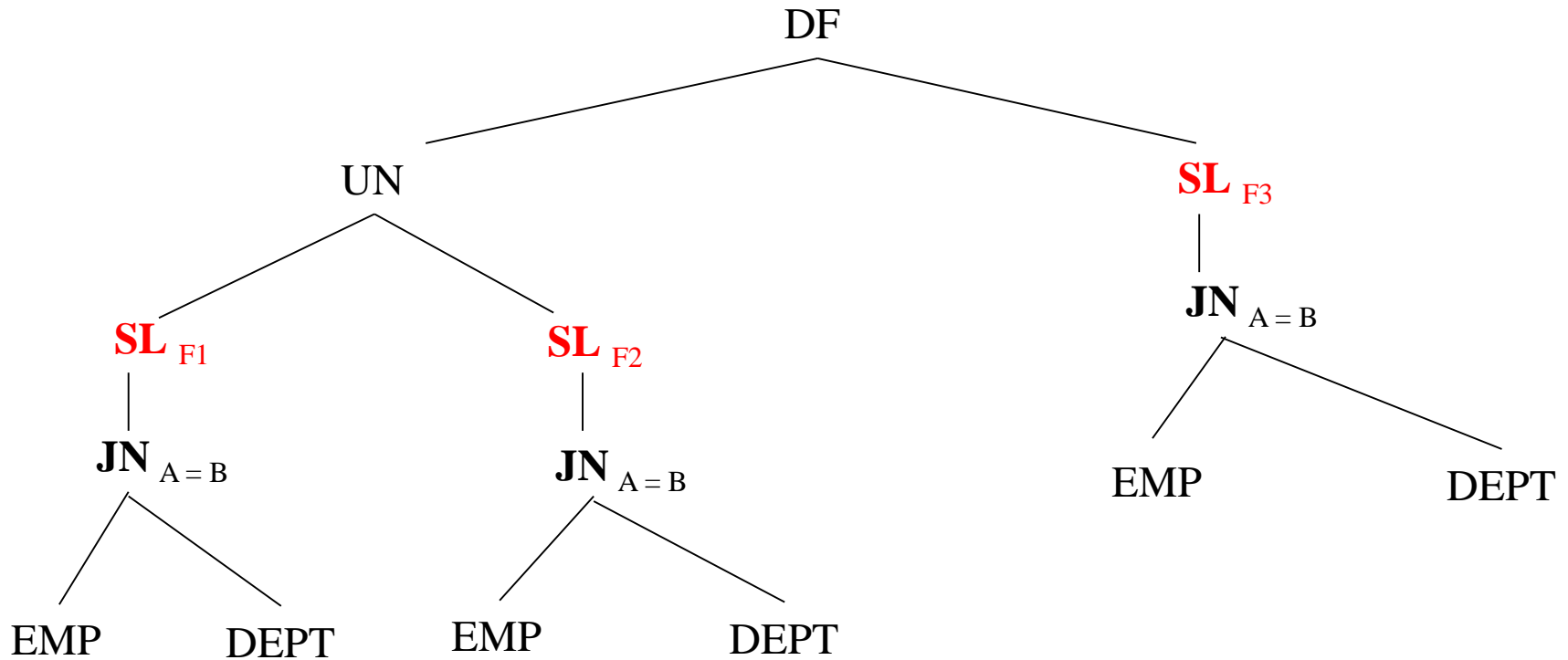


# Finding Common Sub-Expression



# Finding Common Sub-expression

Any common portion?

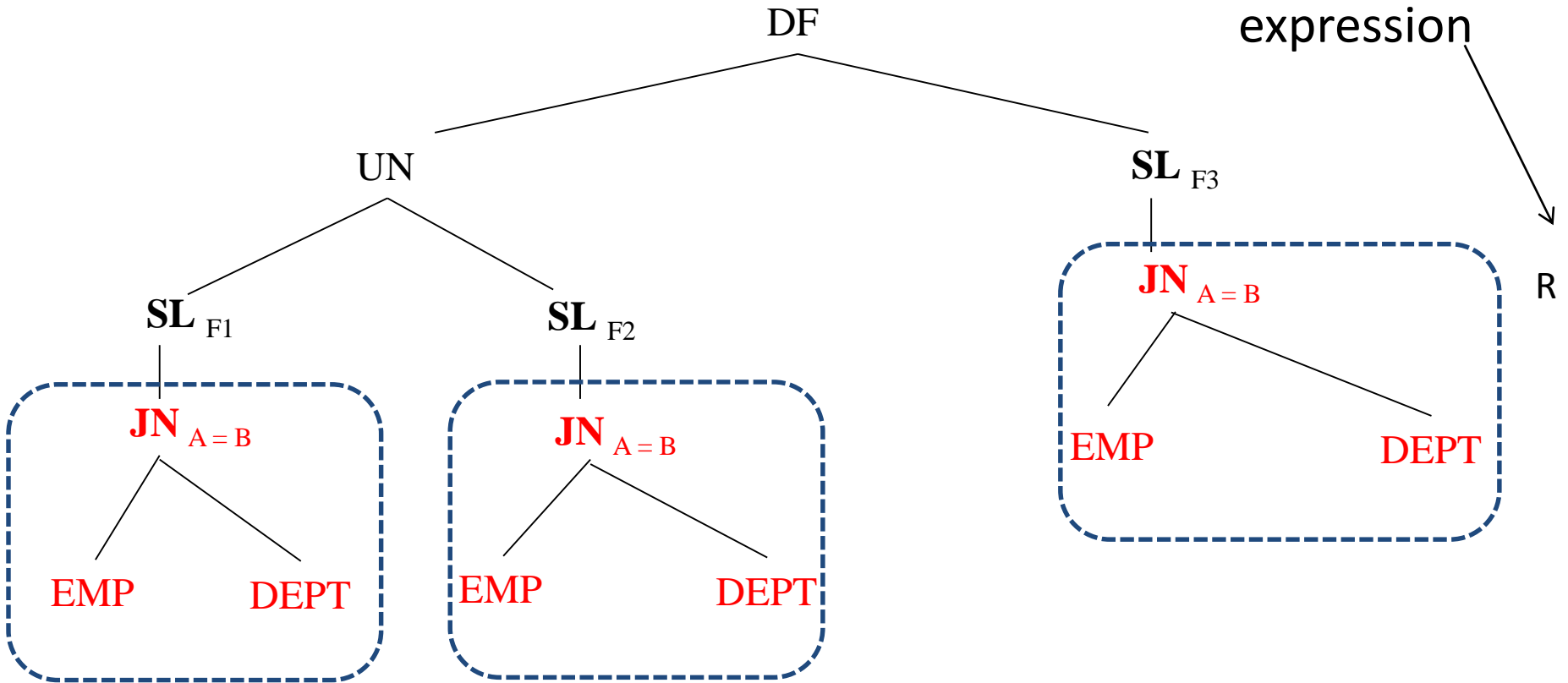


# Finding Common Sub-expression

Any common portion?

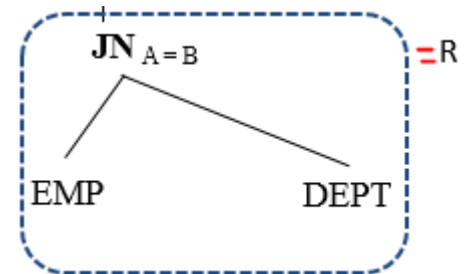
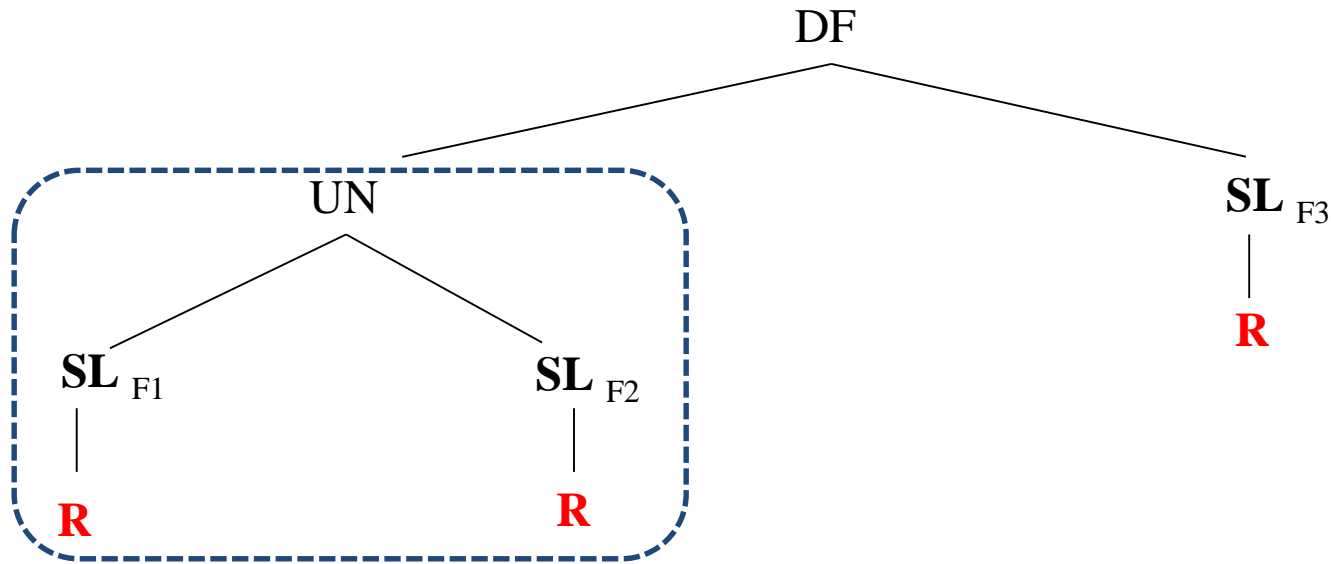
Common sub-expression

R



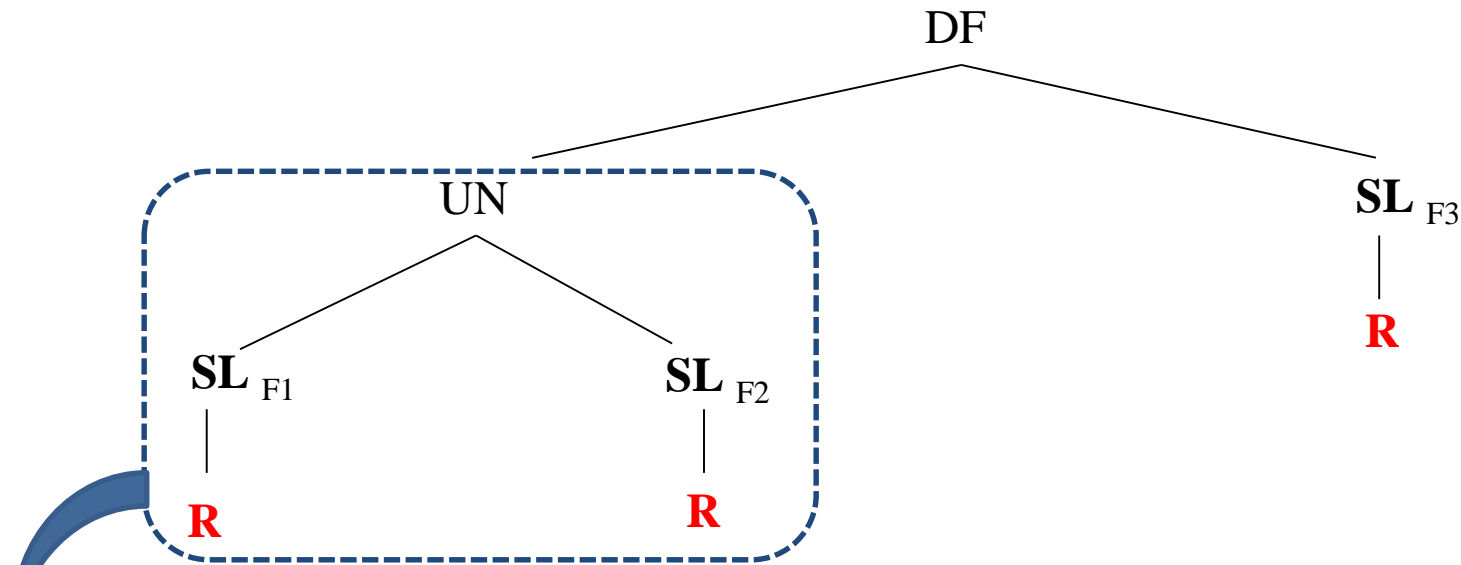
# Finding Common Sub-expression

Any common portion?

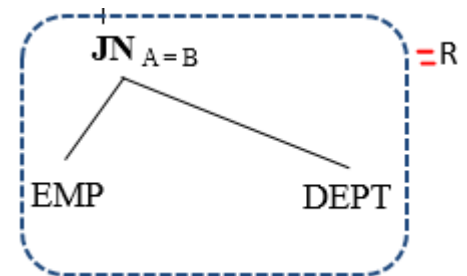


# Finding Common Sub-expression

Any common portion?

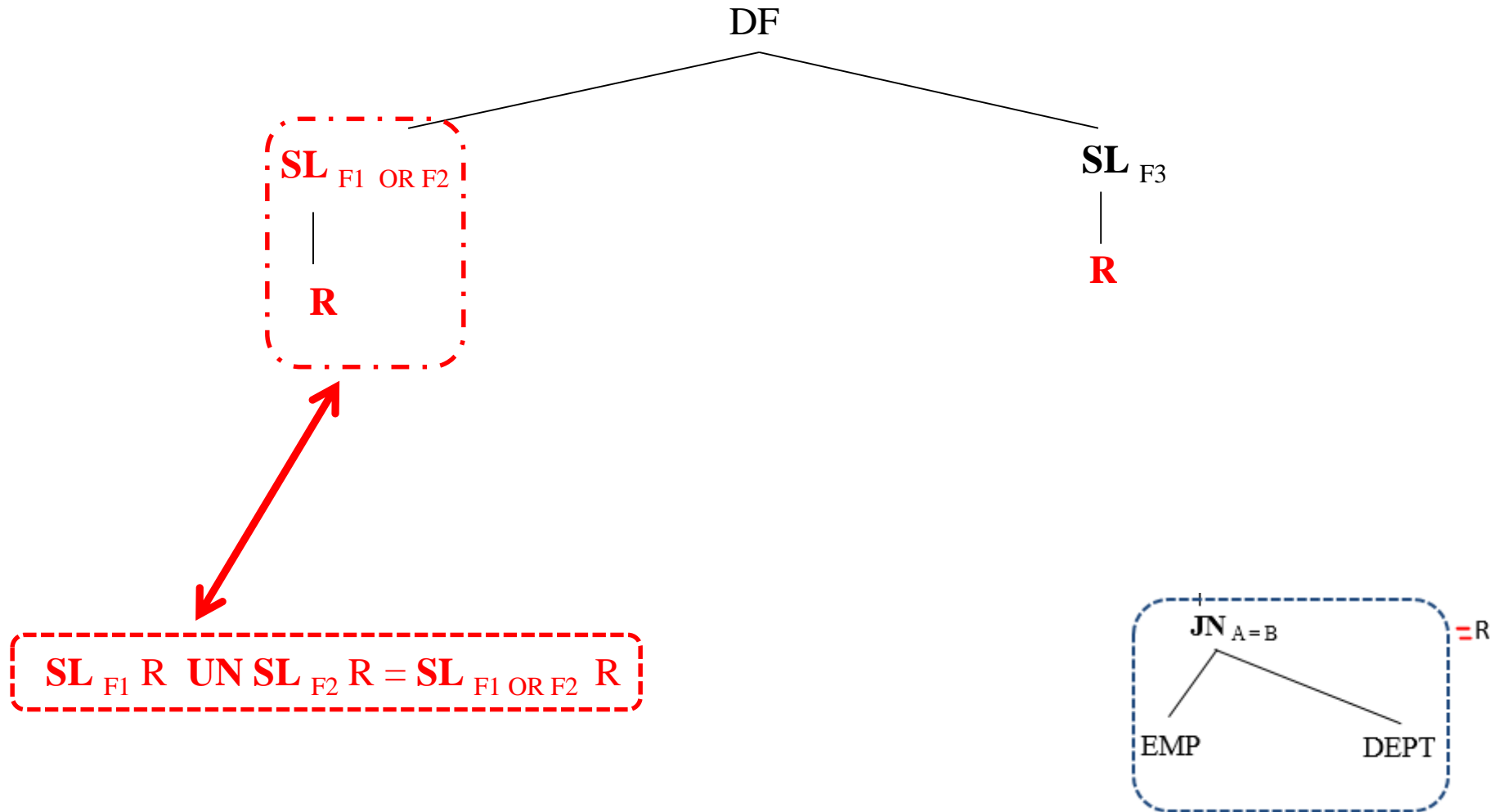


We can write it as  $SL_{F1} R \text{ UN } SL_{F2} R$  which is Rule 8 !



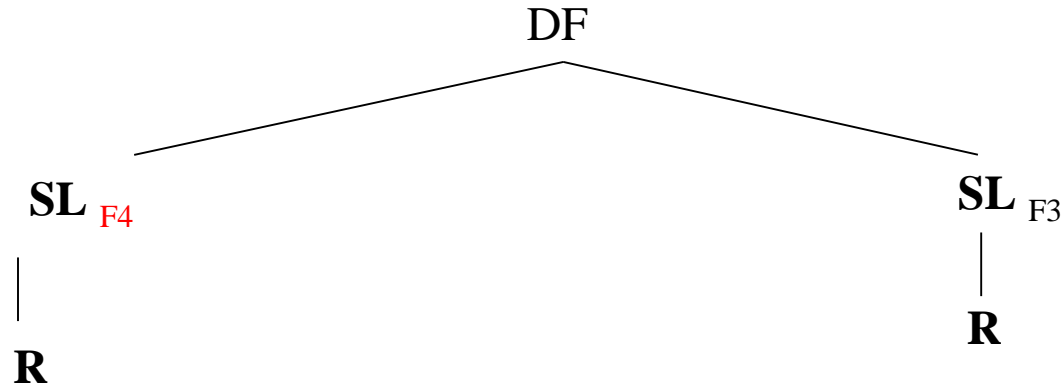
# Removing Common Sub-expression

Any common portion?

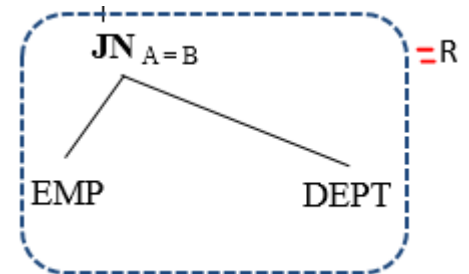


# Finding Common Sub-expression

Any common portion?



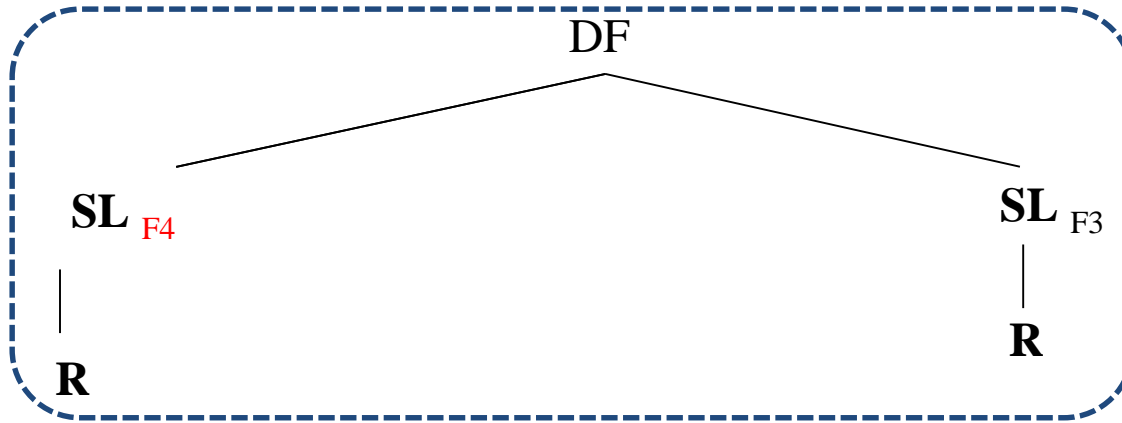
Let,  $F4 = F1 \text{ OR } F2$





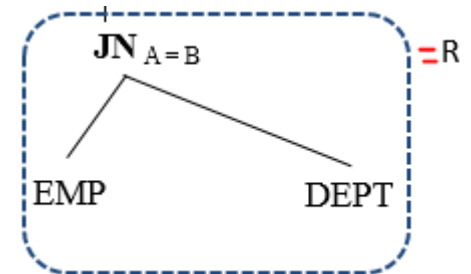
# Finding Common Sub-expression

Any common portion?



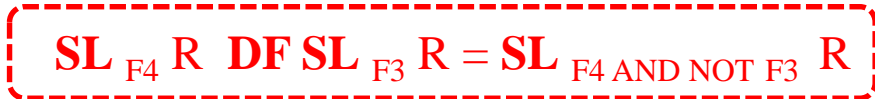
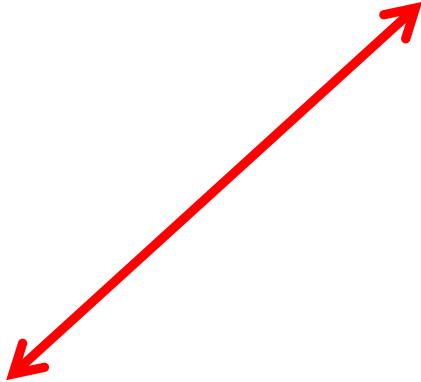
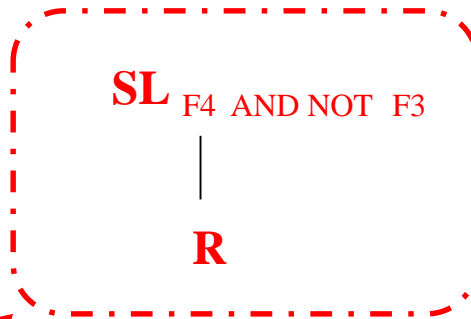
We can write it as  $SL_{F4} R \quad DF \quad SL_{F3} R$  which is Rule 9 !

Let,  $F4 = F1 \text{ OR } F2$

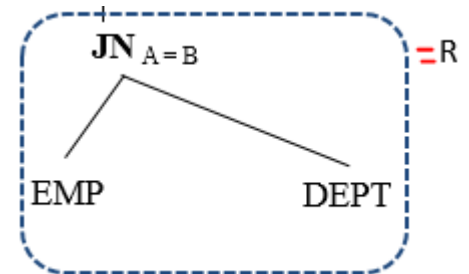


# Removing Common Sub-expression

Any common portion?



Let,  $F4 = F1 \text{ OR } F2$

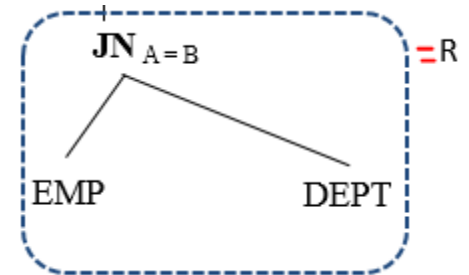


# Removing Common Sub-expression

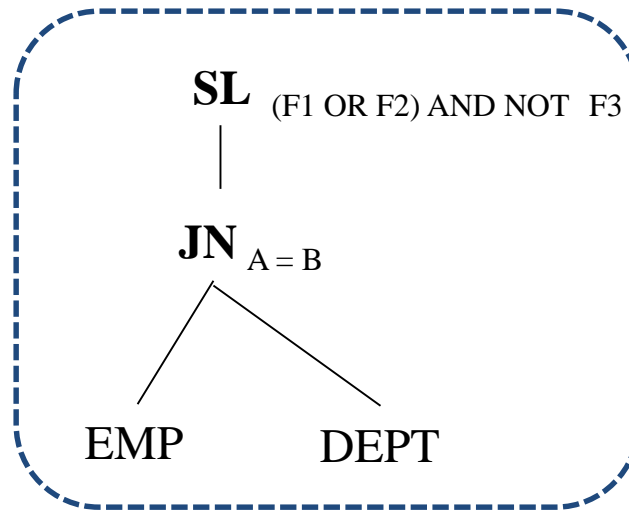
**SL** (F1 OR F2) AND NOT F3

|

**R**



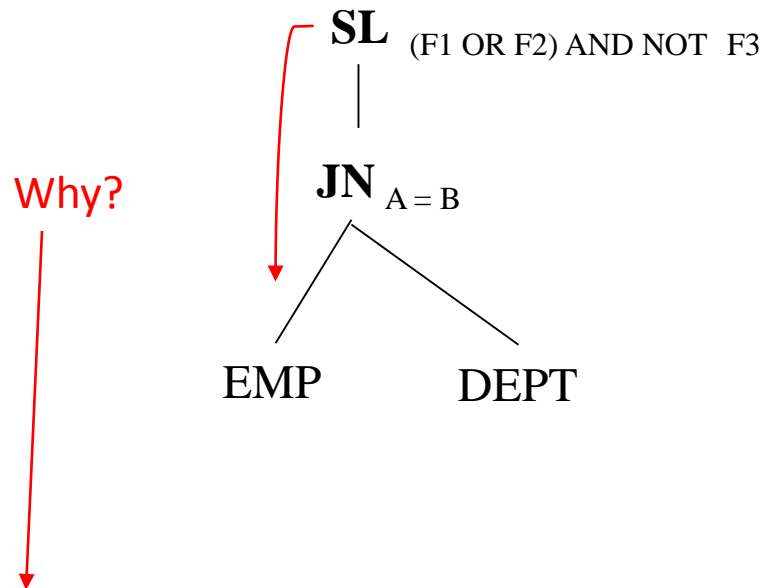
# Removing Common Sub-expression



Can we apply Criterion 1 and/or 2?

# Simplification

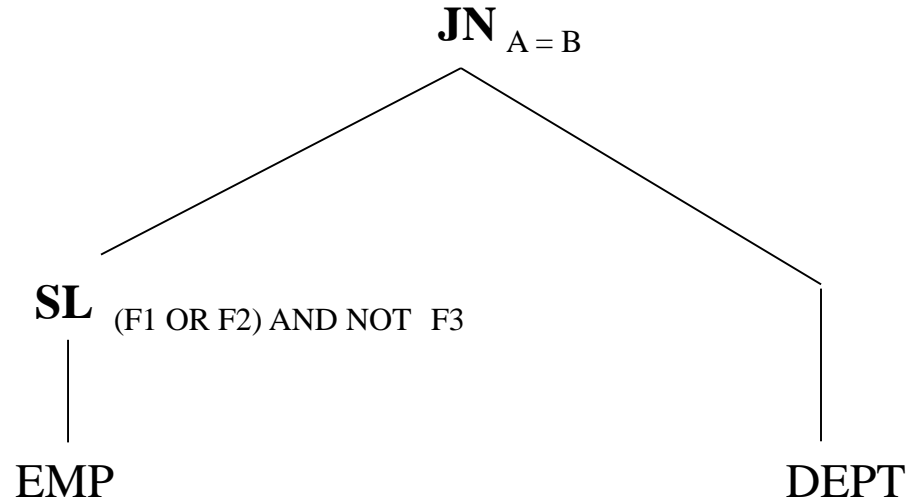
Applying Criterion 2 -



Because in the original query, all **SL**<sub>F1</sub>, **SL**<sub>F2</sub>, **SL**<sub>F3</sub> were applied on EMP relation.

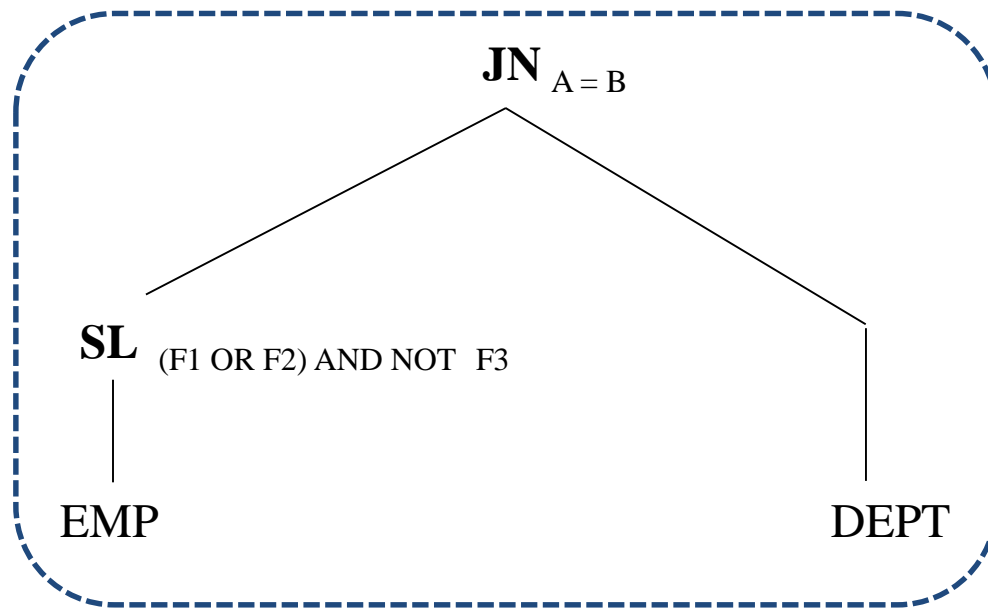
# Simplification

After Applying Criterion 2 -



# Transformed Query

**Q<sub>T</sub>:** **SL**<sub>(F1 OR F2) AND NOT F3</sub> *EMP* **JN**<sub>A=B</sub> *DEPT*



# Transformed Query

**Output:**

**Q<sub>T</sub>:** **SL** <sub>(F1 OR F2) AND NOT F3</sub> *EMP* **JN** <sub>A=B</sub> *DEPT*

**Input:**

**Q:**

$((SL_{F1} \text{ EMP } JN_{A=B} \text{ DEPT}) \text{ UN } (SL_{F2} \text{ EMP } JN_{A=B} \text{ DEPT})) \text{ DF } (SL_{F3} \text{ EMP } JN_{A=B} \text{ DEPT})$

$$Q \longleftrightarrow Q_T$$



# Exercise 1

② EMP (EMPNUM, DEPTNUM, NAME, SAL, AGE)  
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Query: PJ NAME, AGE ((EMP JN DEPTNUM=DEPTNUM SL AREA="North"  
DEPT) DF (EMP JN DEPTNUM=DEPTNUM SL DEPTNUM < 10 DEPT))

# Exercise 2

4. Consider the following global relational schemata.

*EMP* (*ID*, *NAME*, *SAL*, *AGE*, *MGRNUM*, *DEPTNUM*)

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

Corresponding fragmentation schemata:

$EMP_1 = SL_{SAL \leq 25K} EMP$

$EMP_2 = SL_{SAL > 25K} EMP$

$DEPT_1 = SL_{AREA = "North"} DEPT$

$DEPT_2 = SL_{AREA = "South"} DEPT$

Also consider the following global query.

$PJ_{NAME, AREA}(((SL_{SAL > 25K} EMP \Join_{ID=ID} SL_{AREA = "North"} DEPT) \Join_{ID=ID} SL_{SAL \leq 25K} EMP \Join_{ID=ID} SL_{AREA = "North"} DEPT)) \Join_{ID=ID} (SL_{AREA = "North"} (EMP \Join_{ID=ID} DEPT)))$

# Exercise 3

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

Consider the following global query:

$$\left( (SL_{F1} \text{ EMP } JN_{A=B} \text{ DEPT}) \text{ DF } (SL_{F2} \text{ EMP } JN_{A=B} \text{ DEPT}) \right) \text{ NJN } \\ \left( (\text{EMP } JN_{A=B} \text{ DEPT}) \text{ UN } (SL_{F3} \text{ EMP } JN_{A=B} \text{ DEPT}) \right)$$

Here,  
F1, F2, F3 can represent any condition. In this example consider none of them are same.  
Imagine,  $A = B = \text{DEPTNUM}$

# Exercise 4

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

Consider the following global query:

$$\left( (SL_{F_1} \text{ EMP } JN_{A=B} \text{ DEPT}) \text{ UN } (SL_{F_2} \text{ EMP } JN_{A=B} \text{ DEPT}) \right) \text{ NJN } \\ \left( (EMP \text{ } JN_{A=B} \text{ DEPT}) \text{ DF } (SL_{F_3} \text{ EMP } JN_{A=B} \text{ DEPT}) \right)$$

Here,  
F1, F2, F3 can represent any condition. In this example consider none of them are same.  
Imagine,  $A = B = \text{DEPTNUM}$

# Exercise 5

Consider the following global query and answer the questions from (i) to (iii).

$$\begin{aligned} &(((SL_{F1} R JN_{A=B} S) DF (SL_{F2} R JN_{A=B} S)) NJN (R JN_{A=B} S)) \\ &UN (SL_{F1 AND NOT F2} R JN_{A=B} S) \end{aligned}$$

- i) Draw the operator tree. [2]
- ii) Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [5]
- iii) Write the query from the obtained simplified tree. [2]

Here,  
F1, F2, F3 can represent any condition. In this example consider none of them are same.  
Imagine, A and B represents the same attribute of two different relation R and S.

# Exercise 6

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

(a) Consider the following global query and answer the questions (i) and (ii).

$$\left( \left( (SL_{SAL} > 25K \text{ EMP } \Join_{ID=ID} \text{ DEPT}) \text{ DF } (SL_{AGE} \leq 25 \text{ EMP } \Join_{ID=ID} \text{ DEPT}) \right) \text{ NJN } (\text{EMP } \Join_{ID=ID} \text{ DEPT}) \right)$$
  
$$\text{DF } (SL_{SAL} > 25K \text{ AND } AGE > 25 \text{ EMP } \Join_{ID=ID} \text{ DEPT})$$

- i. Draw the operator tree. [2]
- ii. Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [6]