

CSE 4125: Distributed Database Systems Chapter – 5

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

Topics to be discussed –

- Simplification of **Distributed** Query

Question type:

/*Fragments and a query is given */

Now, answer the following questions.

- Draw the *operator tree*. [2]
- Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [5]
- Transform the simplified query into fragment query by applying canonical expression based on the given fragmentation schema. [2]
- Write the equivalent query obtained from the simplified tree. [1]

Consider the following global relational schemata.

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

Corresponding fragmentation schemata:

$$EMP_1 = SL_{SAL \leq 35K} EMP$$

$$EMP_2 = SL_{SAL > 35K} EMP$$

$$DEPT_1 = SL_{MGRNUM = 375} DEPT$$

$$DEPT_2 = SL_{MGRNUM = 373} DEPT$$

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]
- iii. Transform the simplified query into fragment query by applying canonical expression based on the given fragmentation schema. [2]
- iv. Write the equivalent query obtained from the simplified tree. [1]

Pre-requisites

- Reconstruction idea of Horizontal and Vertical fragmentation.
- Union operation for Horizontal fragmentation
- Join operation for Vertical fragmentation

□ Given, EMP has 2 horizontal fragments:
 EMP_1 and EMP_2 .

EMP

Transforming Global Queries into Fragment Queries

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

Consider the following global relational schemata.

EMP (*EMPNUM*, *DEPTNUM*, *NAME*, *SAL*, *AGE*)

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

Corresponding fragmentation schemata:

$$EMP_1 = SL_{SAL \leq 35K} EMP$$

$$DEPT_1 = SL_{MGRNUM = 375} DEPT$$

$$EMP_2 = SL_{SAL > 35K} EMP$$

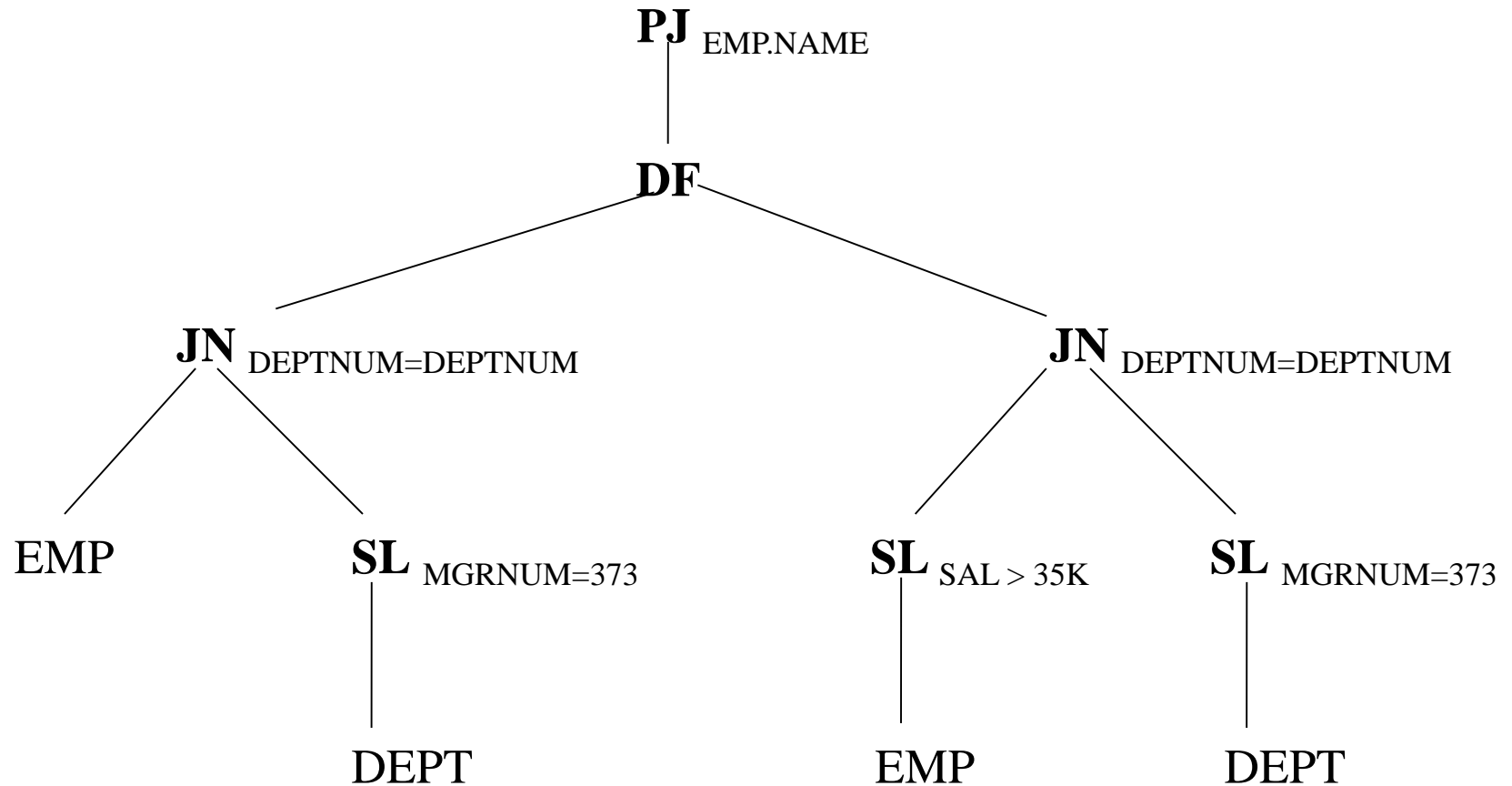
$$DEPT_2 = SL_{MGRNUM = 373} DEPT$$

Q: PJ $EMP.NAME \ ((EMP \Join_{DEPTNUM=DEPTNUM} SL_{MGRNUM=373} DEPT) \Join_{SAL > 35K} EMP \Join_{DEPTNUM=DEPTNUM} SL_{MGRNUM=373} DEPT) \Join_{SAL > 35K} EMP \Join_{DEPTNUM=DEPTNUM} SL_{MGRNUM=373} DEPT)$

Now, answer the following questions.

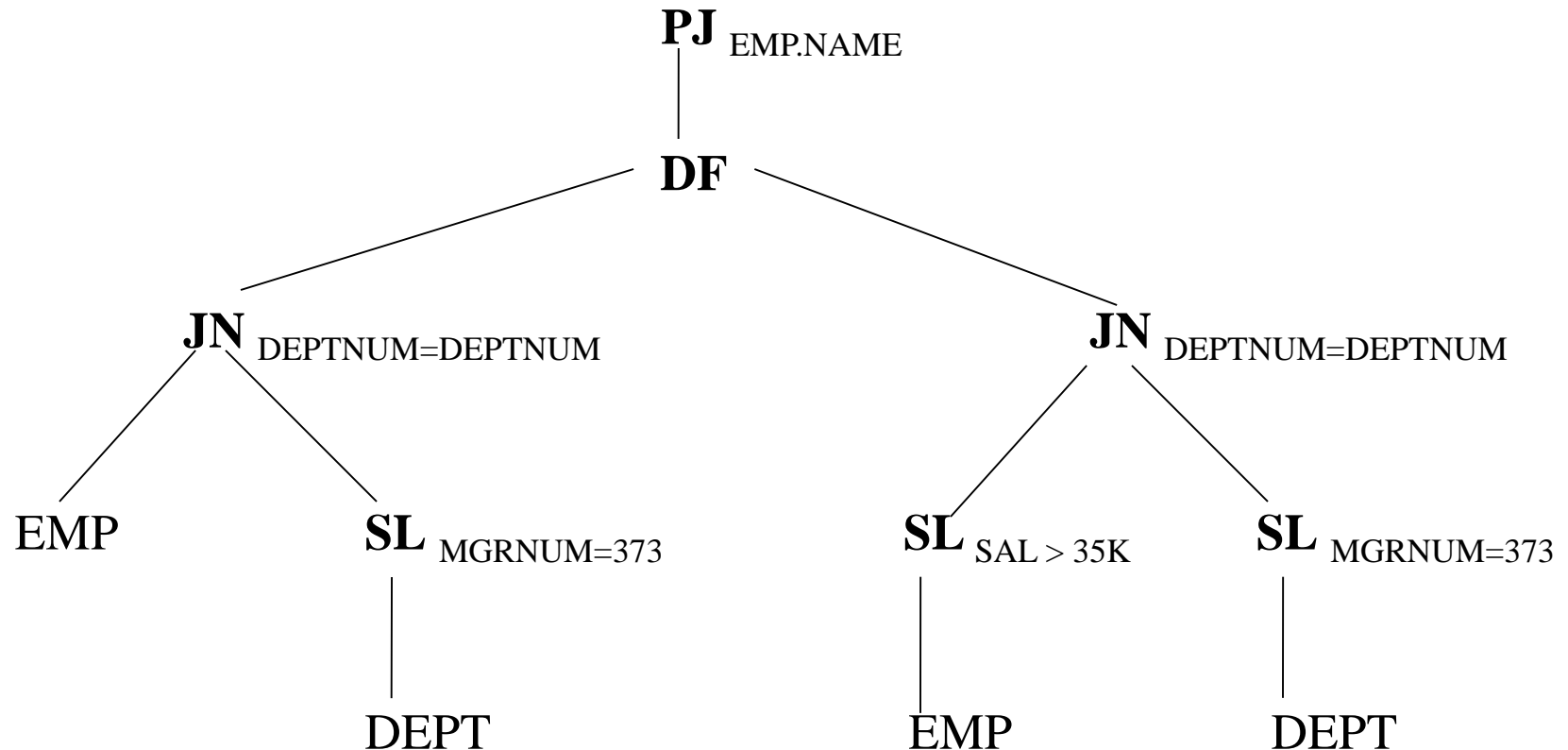
- i. Draw the *operator tree*. [2]
- ii. Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [5]
- iii. Transform the simplified query into fragment query by applying canonical expression based on the given fragmentation schema. [2]
- iv. Write the equivalent query obtained from the simplified tree. [1]

Operator Tree



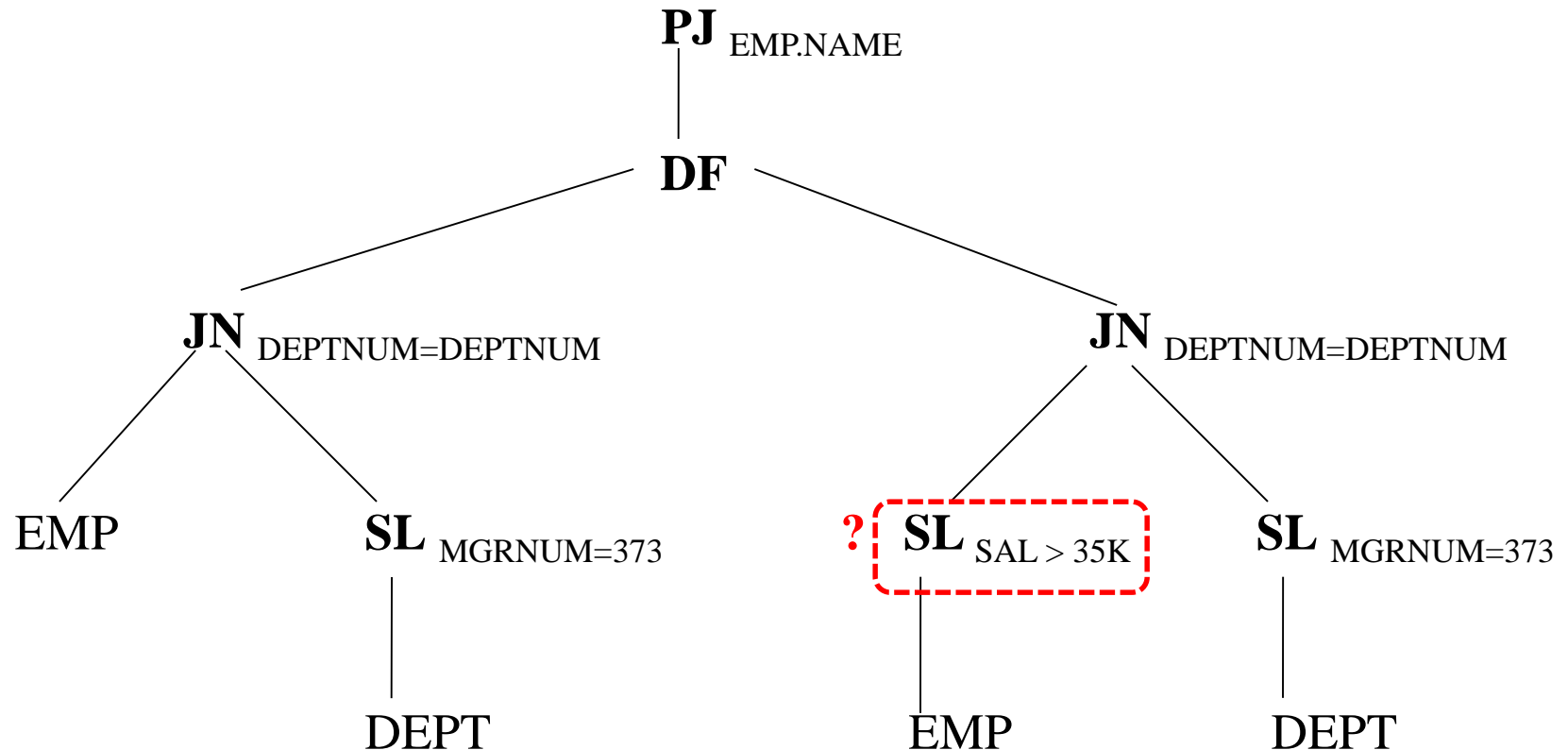
Finding Common Sub-expression

Any common portion?



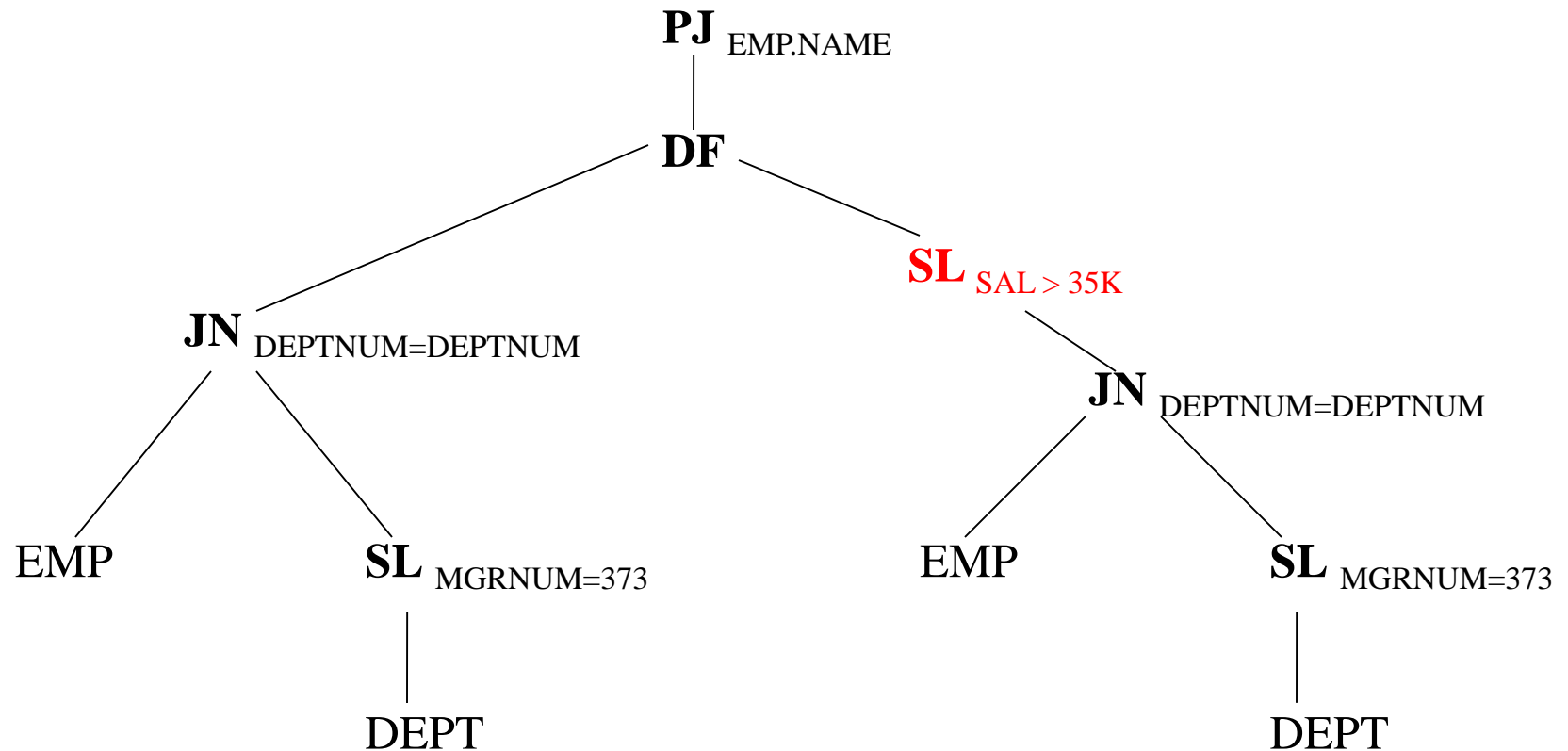
Finding Common Sub-expression

Any common portion?

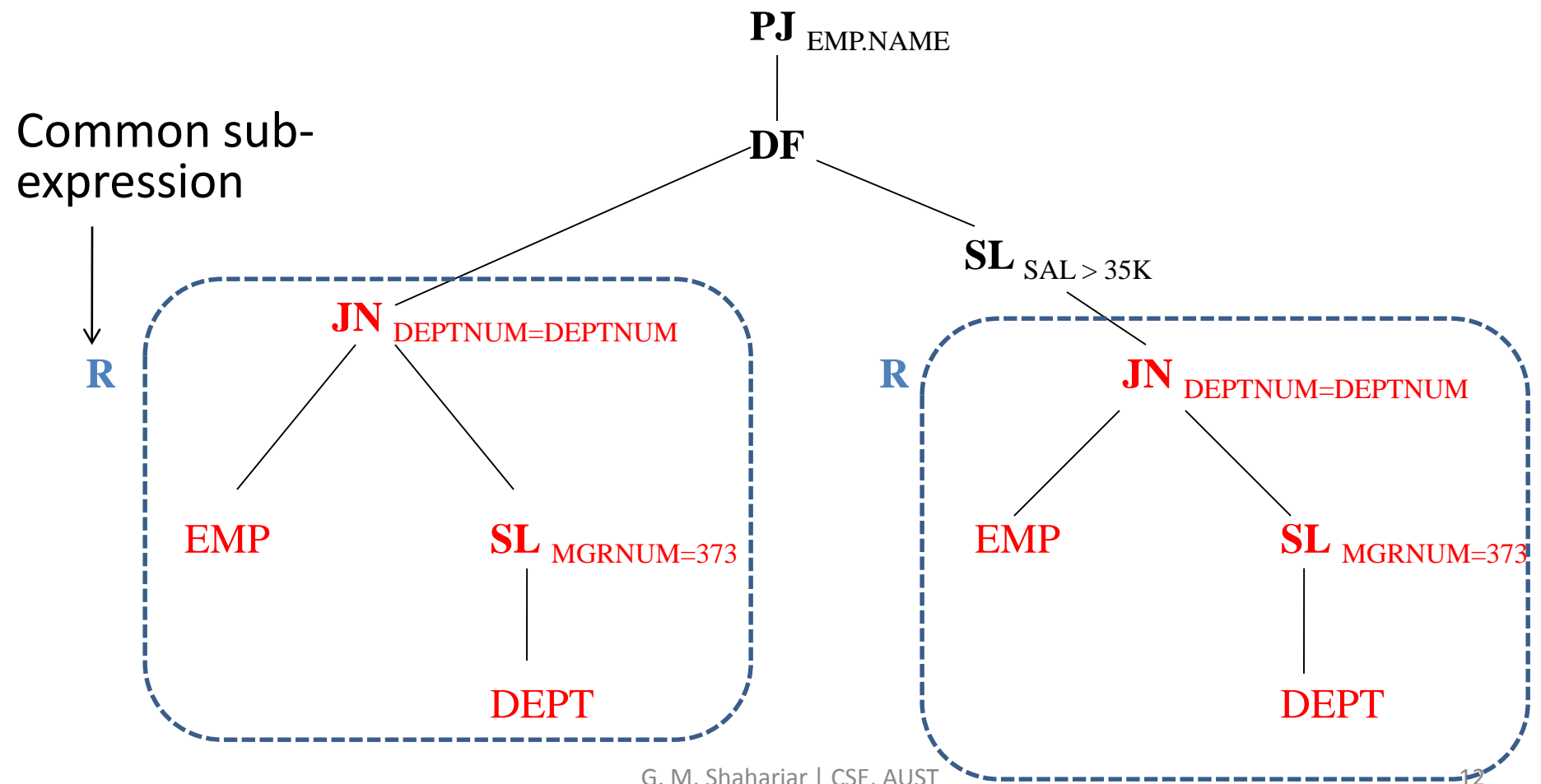


Finding Common Sub-expression

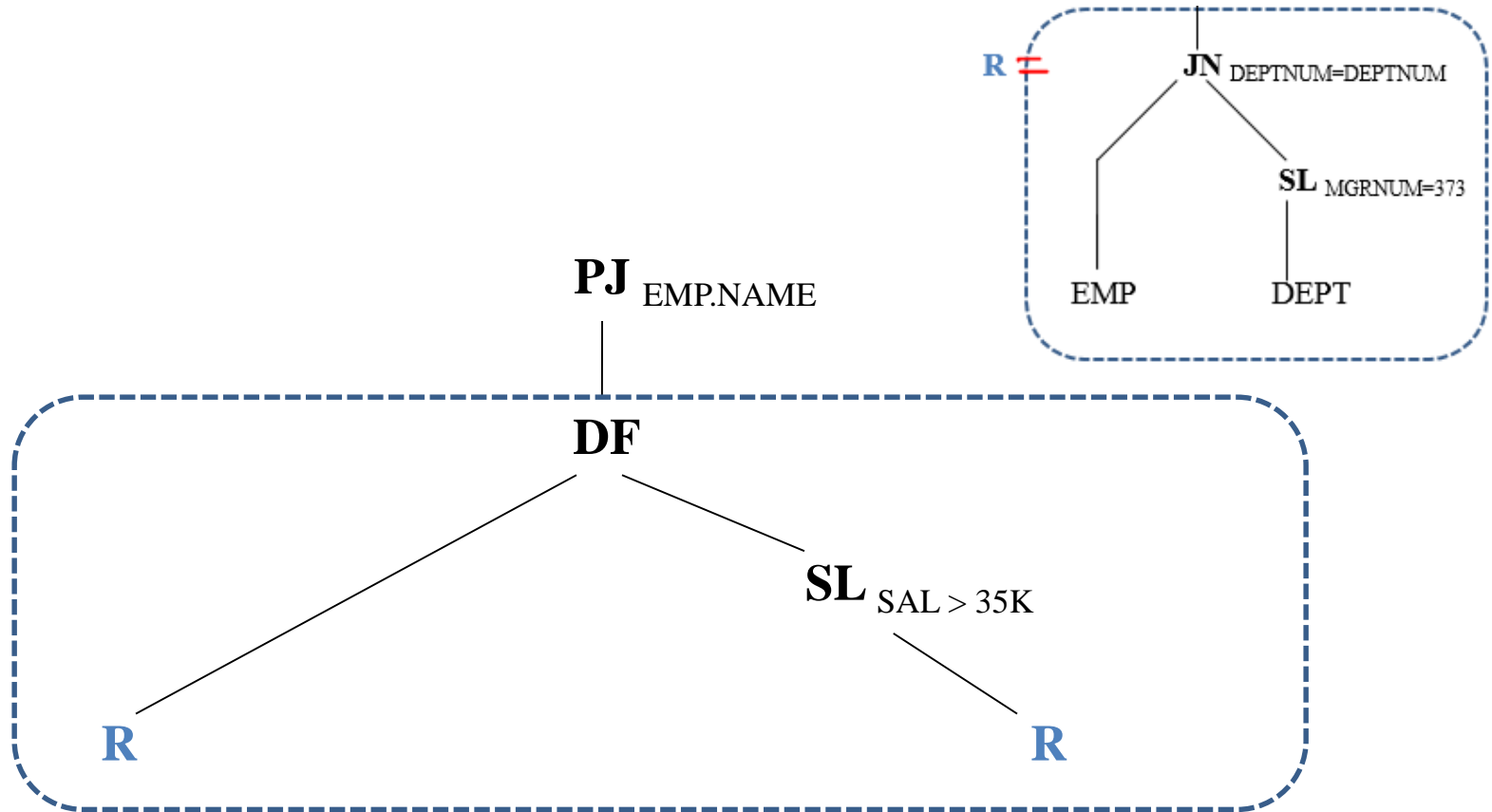
Any common portion? NOW?



Finding Common Sub-expression

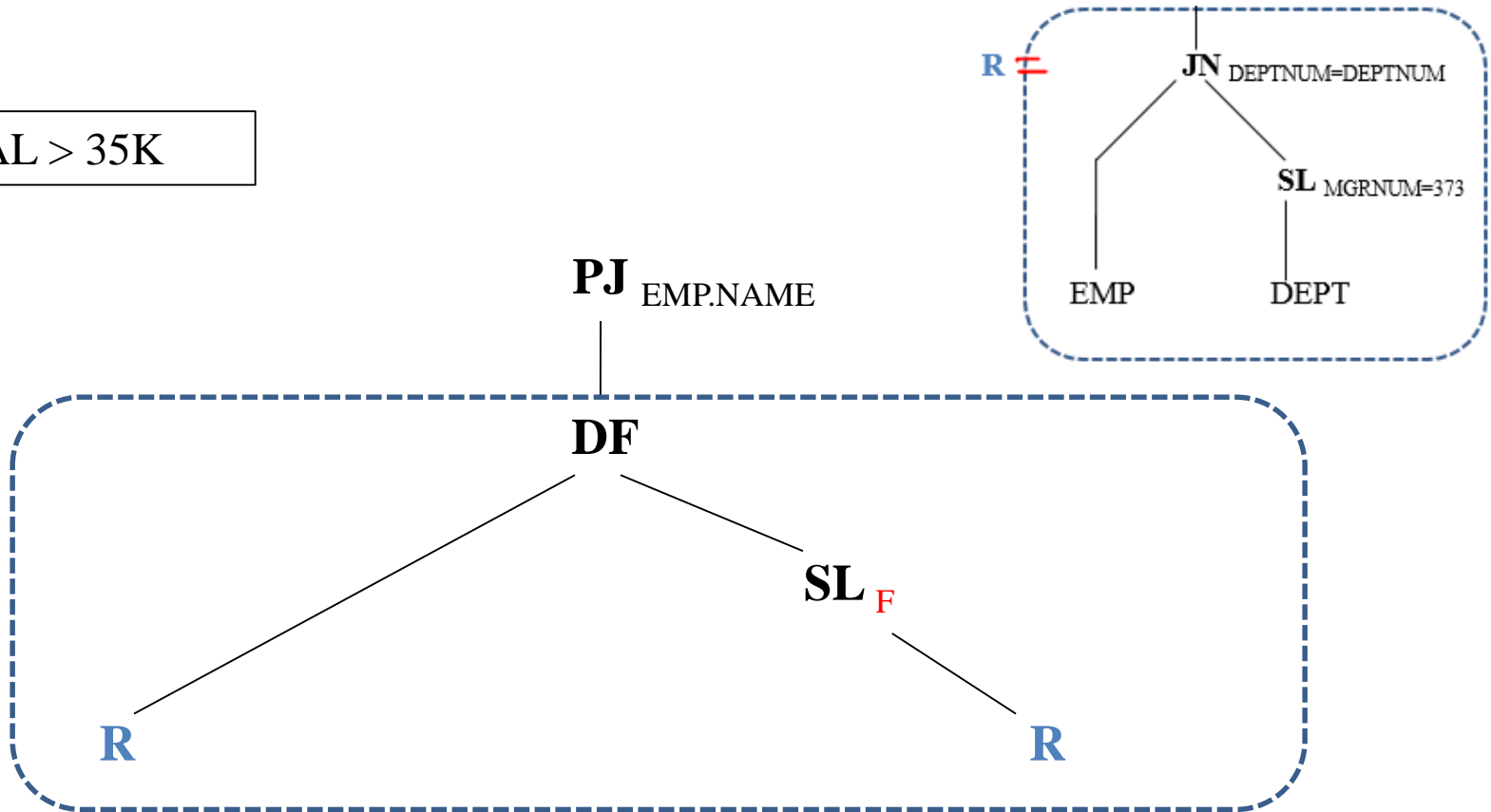


Finding Common Sub-expression



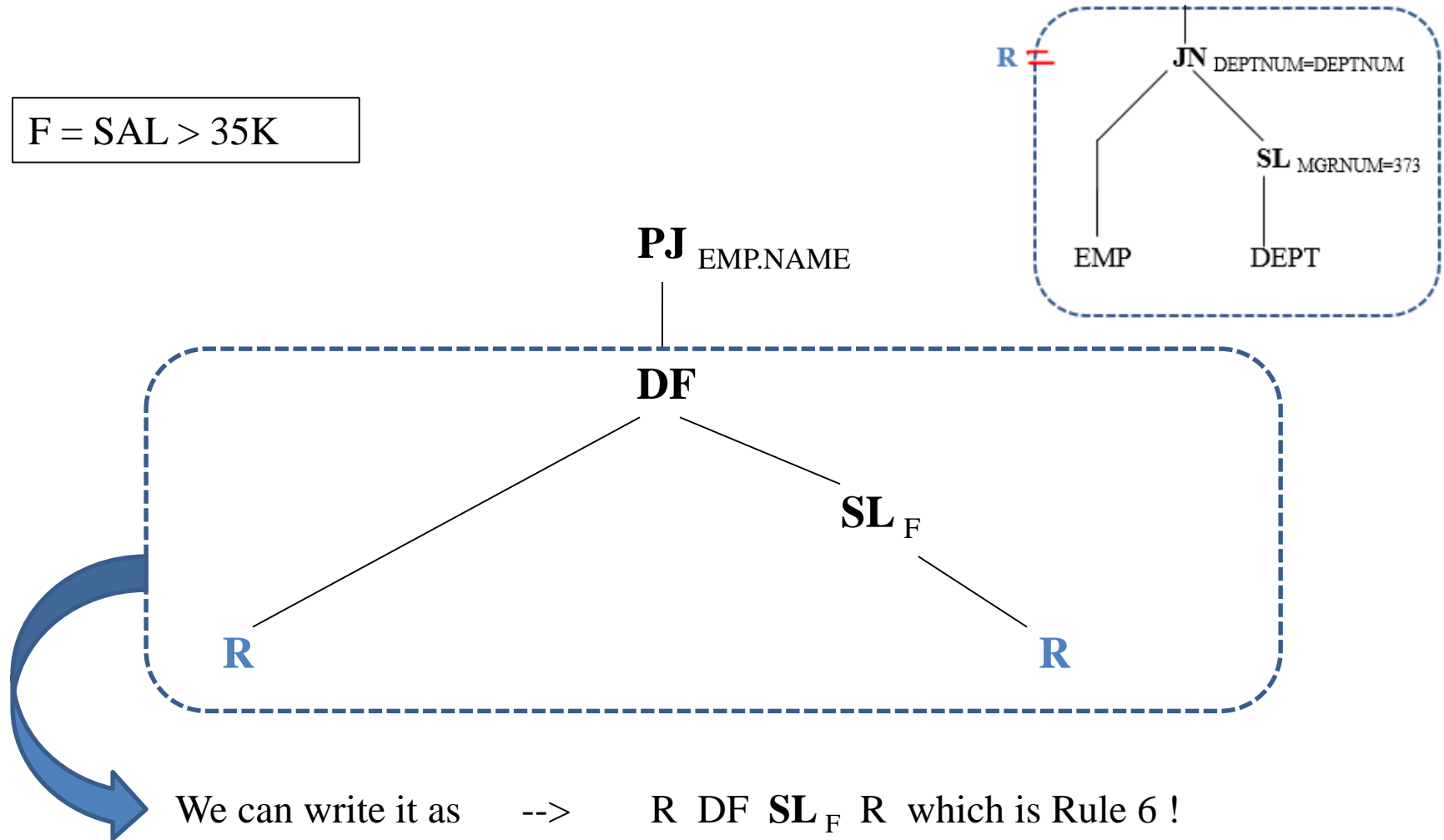
Finding Common Sub-expression

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



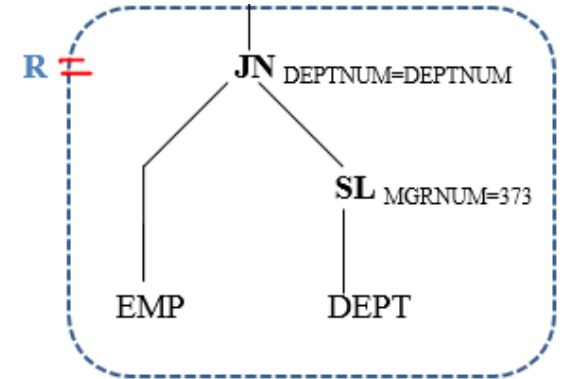
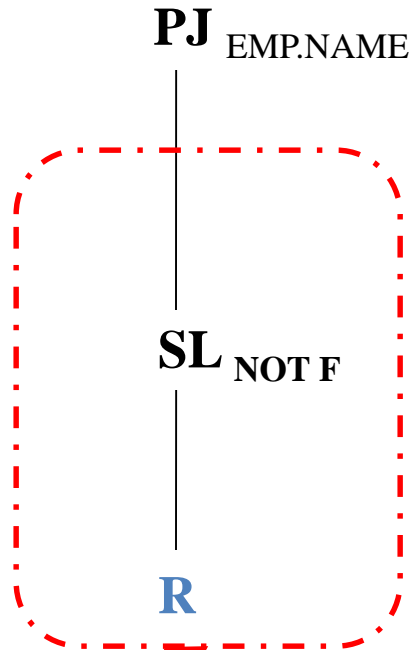
Finding Common Sub-expression

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



Removing Common Sub-expression

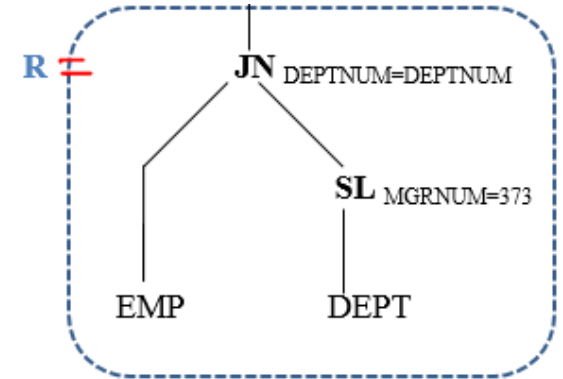
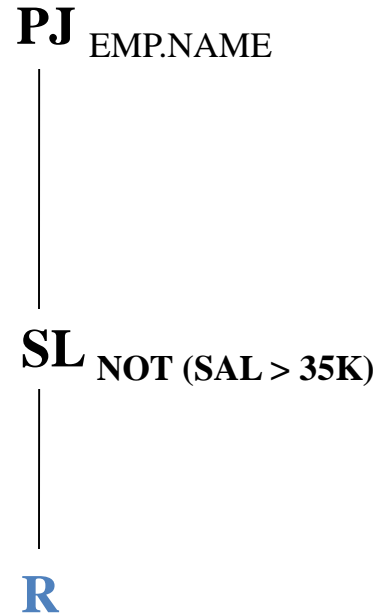
$F = \text{SAL} > 35\text{K}$



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

Removing Common Sub-expression

$F = SAL > 35K$

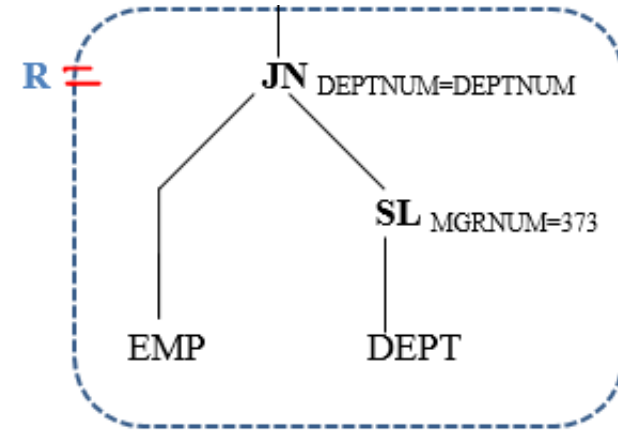


Removing Common Sub-expression

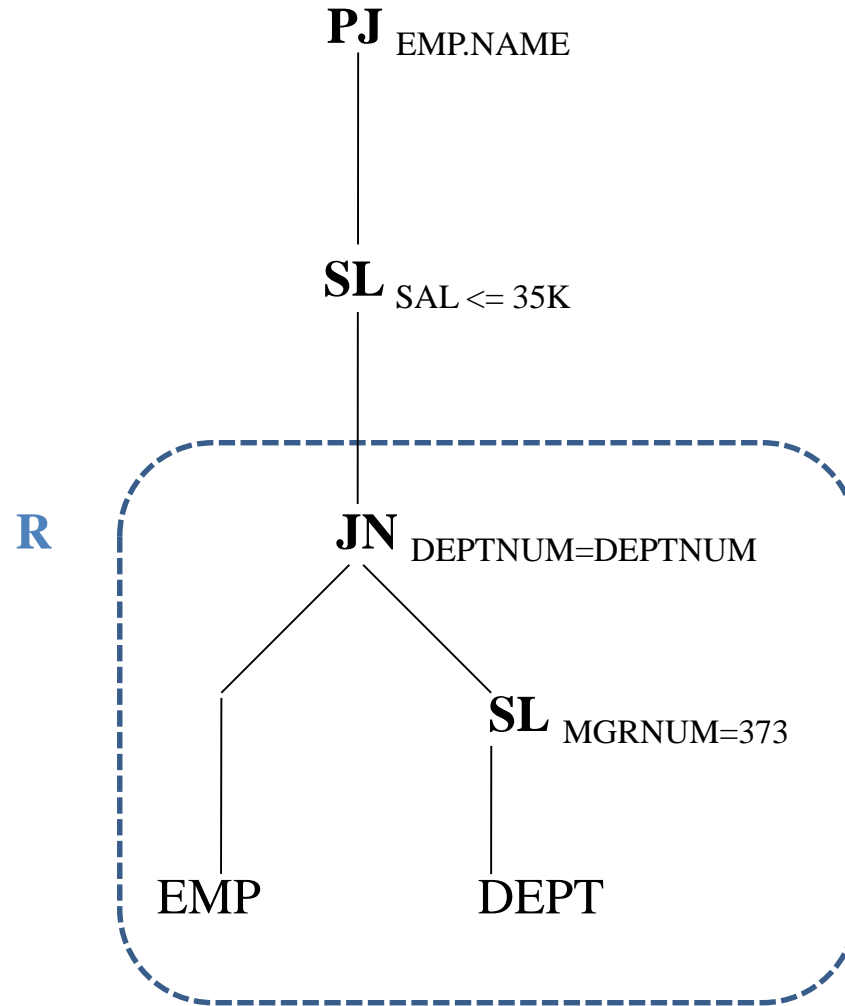
PJ EMP.NAME

SL SAL <= 35K

R



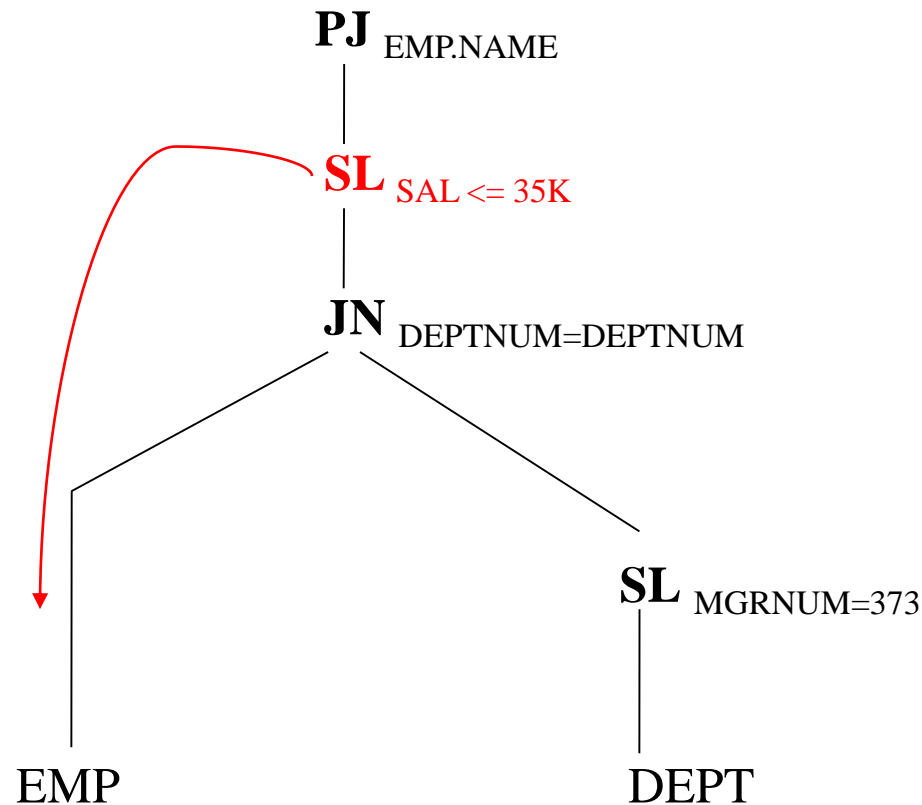
Removing Common Sub-expression



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

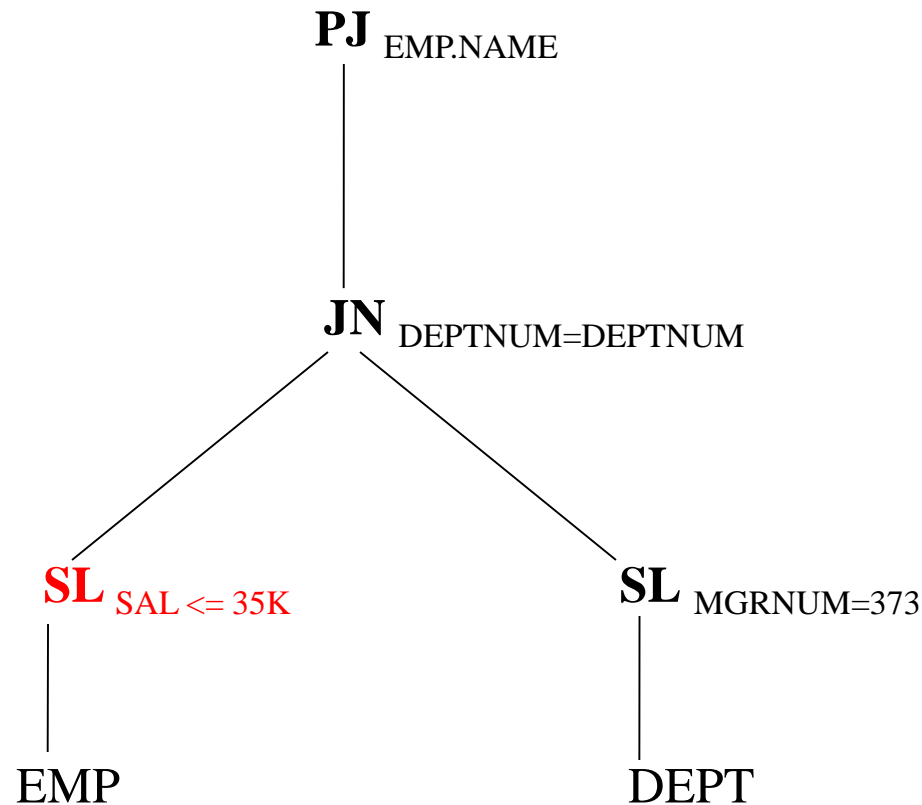
Simplification

Applying criterion – 2



Simplification

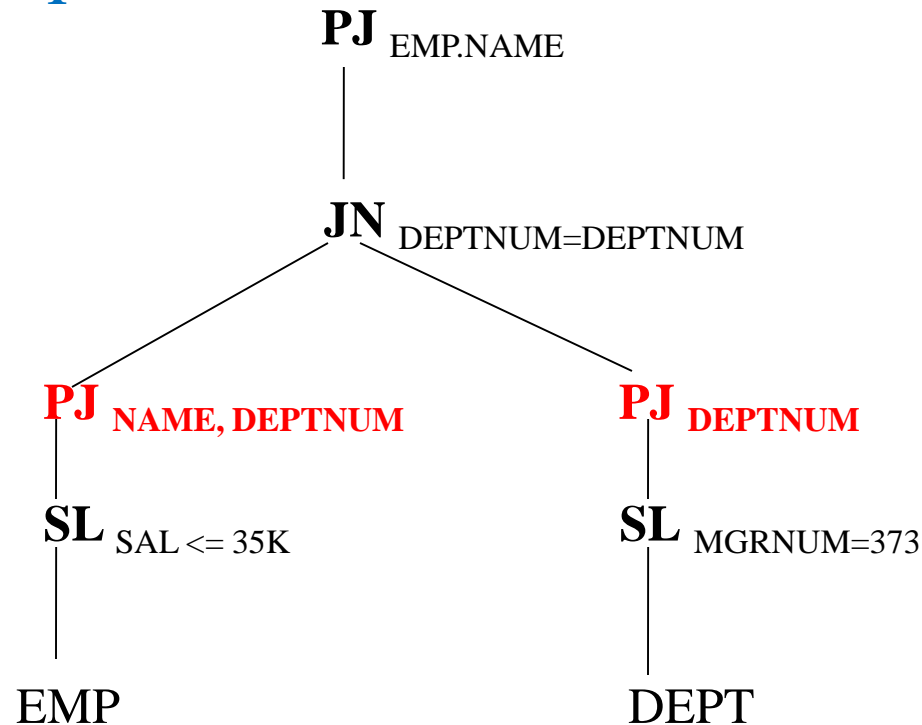
After Applying criterion – 2



Simplification

Q transformed:

After Applying criterion – 1



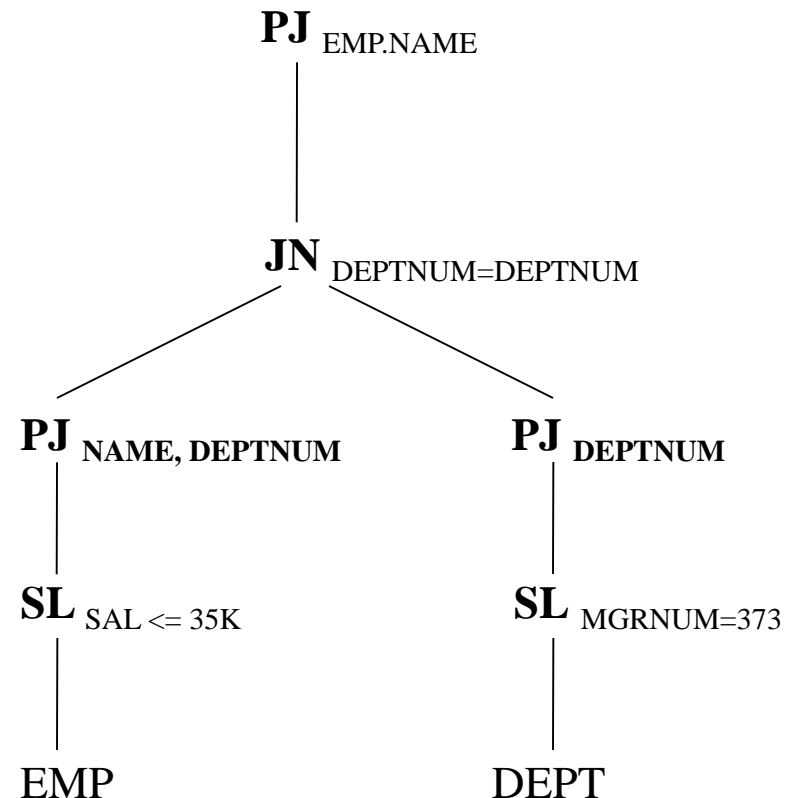
Canonical Expression of a Fragment Query

□ Given, EMP has 2 horizontal fragments: EMP_1 and EMP_2 .
DEPT has 2 horizontal fragments: $DEPT_1$ and $DEPT_2$.

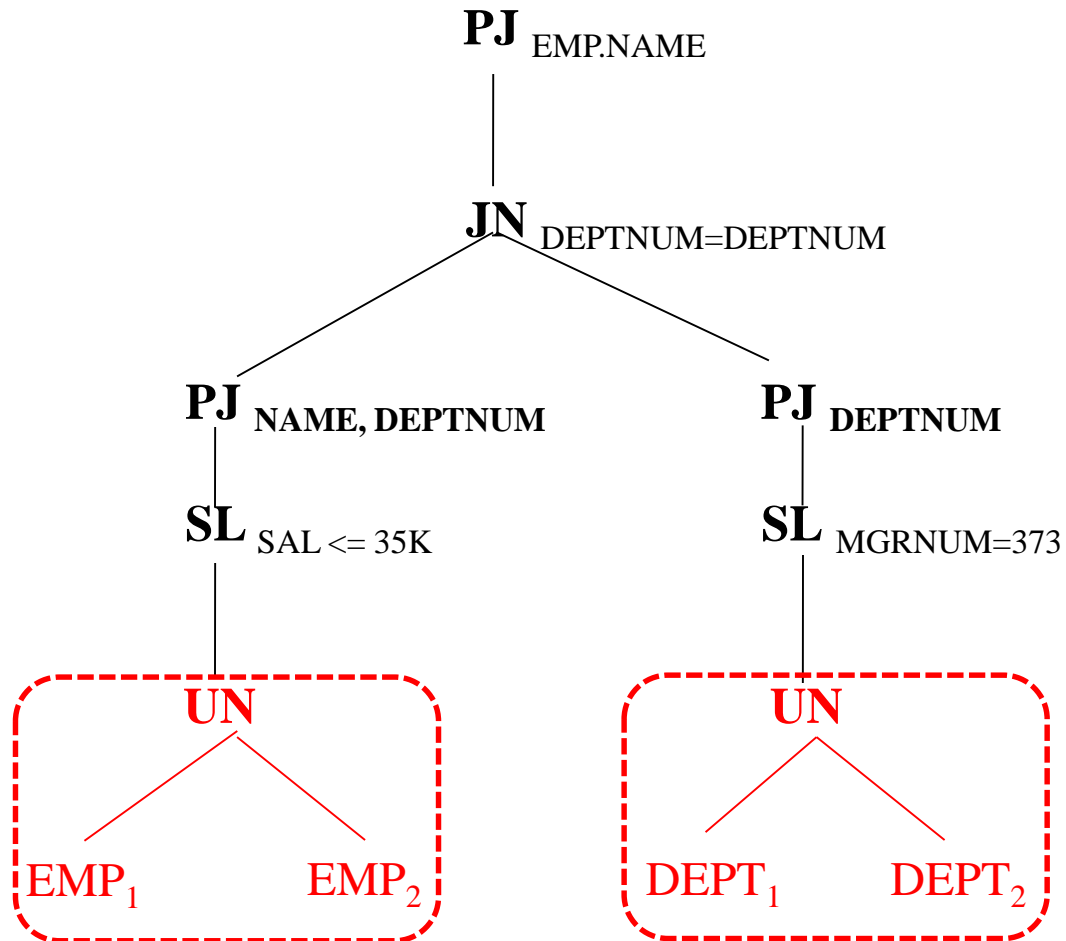
□ How to convert this tree so that the leaves becomes the fragment?

–Consider the reconstruction property of Horizontal Fragmentation.

–Replace the leaf with the sub-tree of the reconstruction expression.



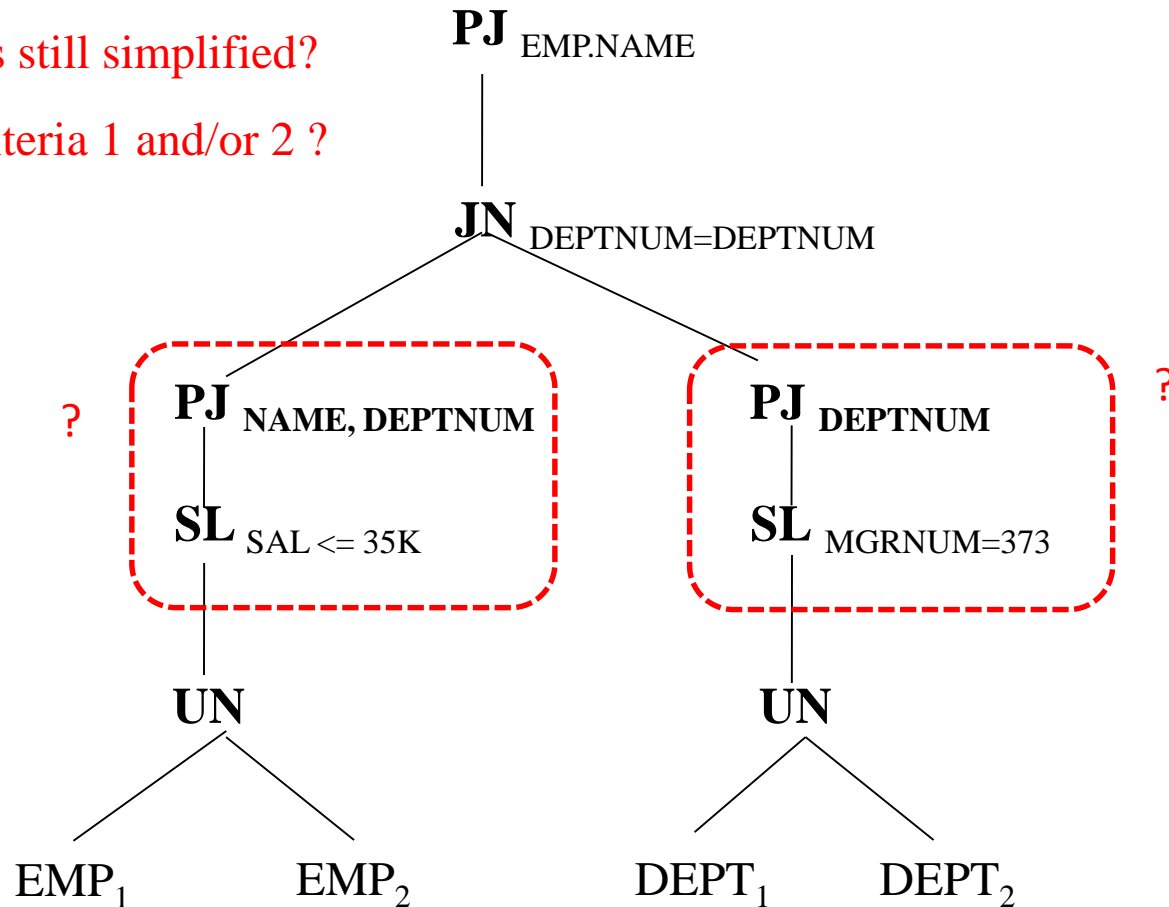
Canonical Expression



Canonical Expression

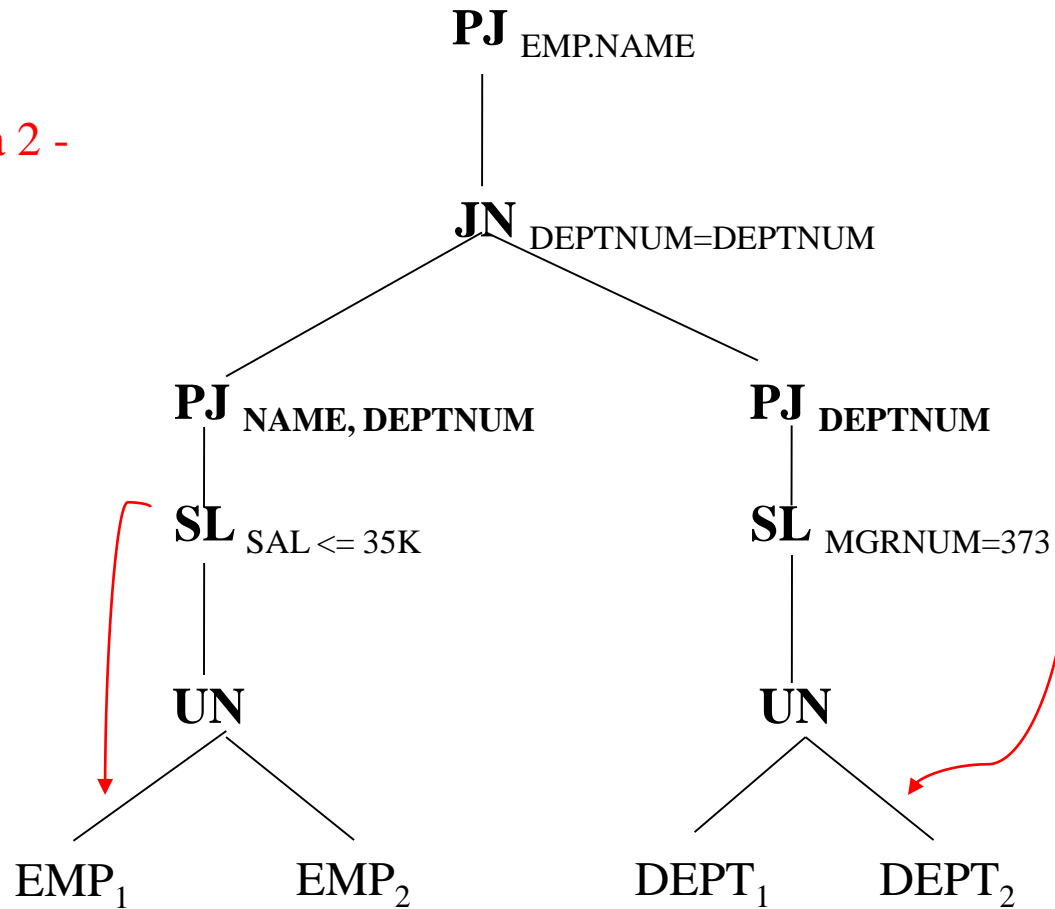
Do you think it is still simplified?

Can we apply Criteria 1 and/or 2 ?



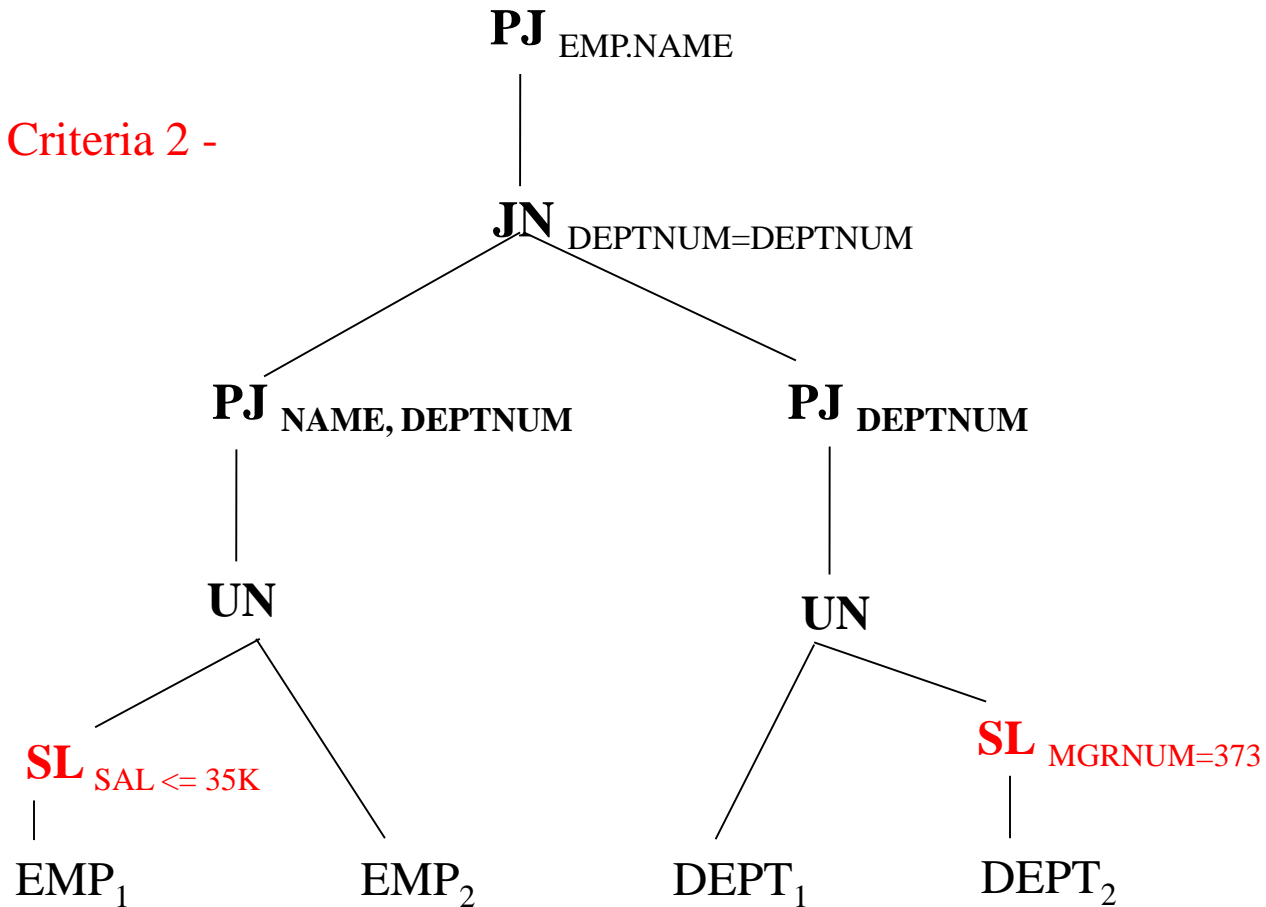
Canonical Expression

Applying Criteria 2 -



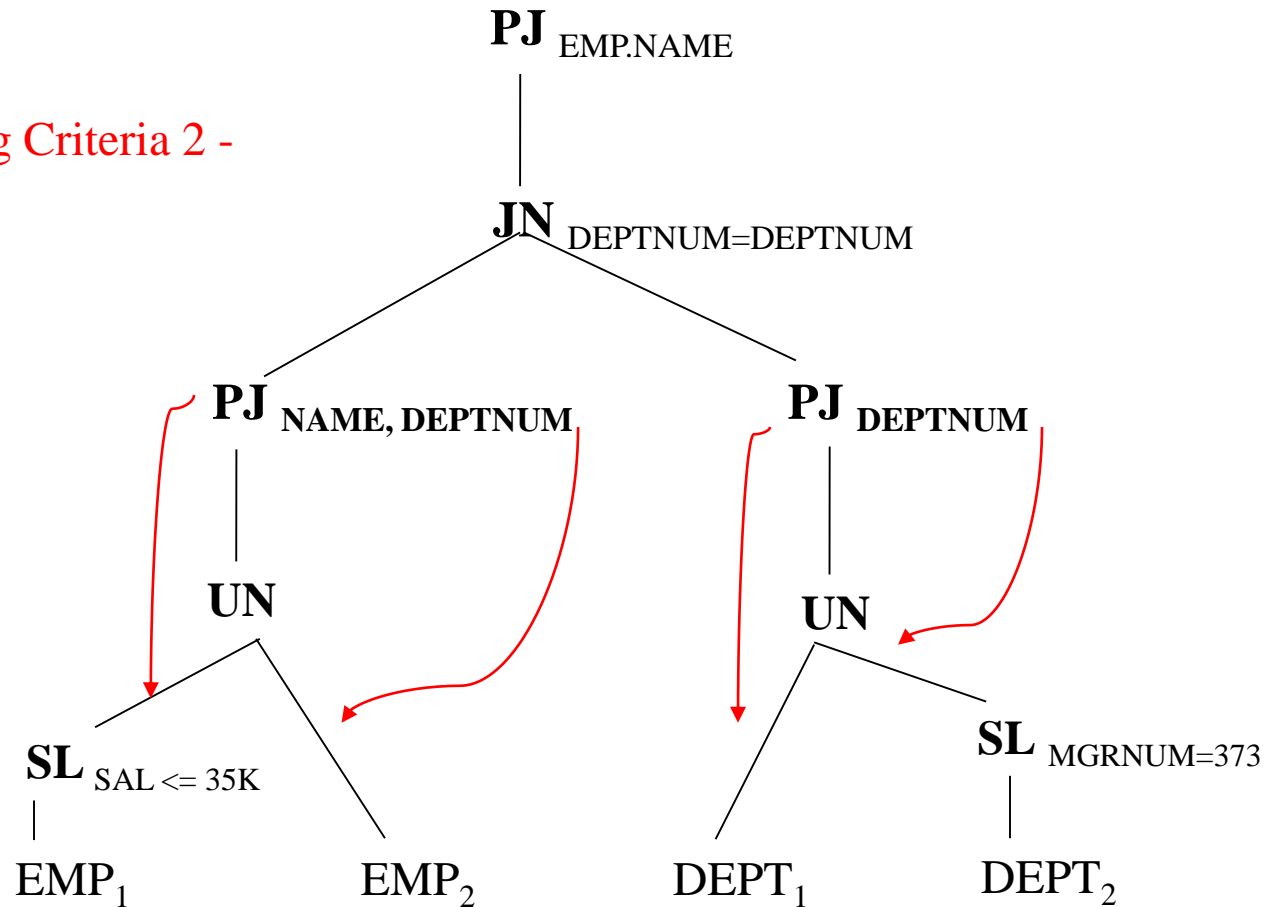
Canonical Expression

After Applying Criteria 2 -



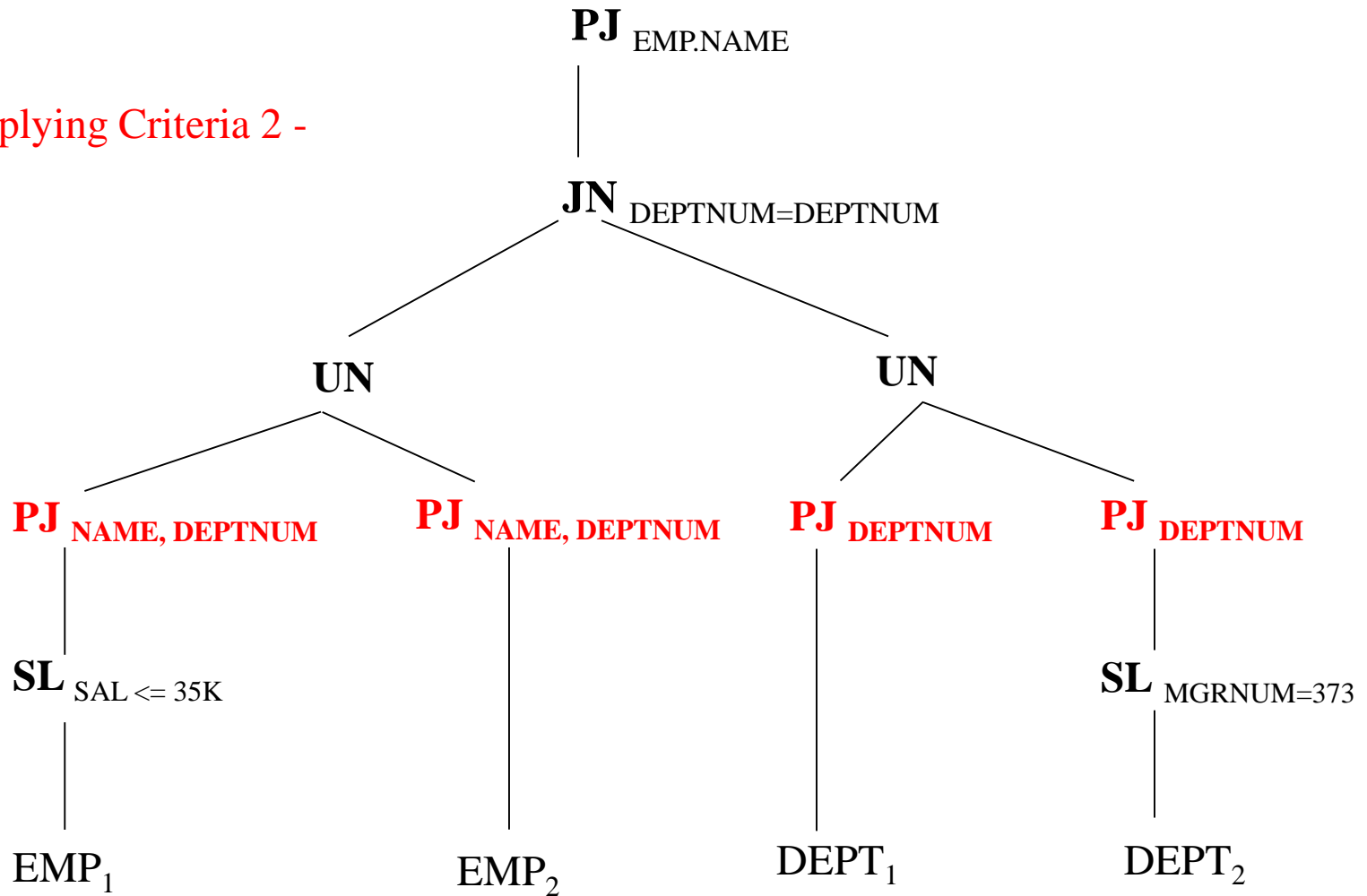
Canonical Expression

Again Applying Criteria 2 -



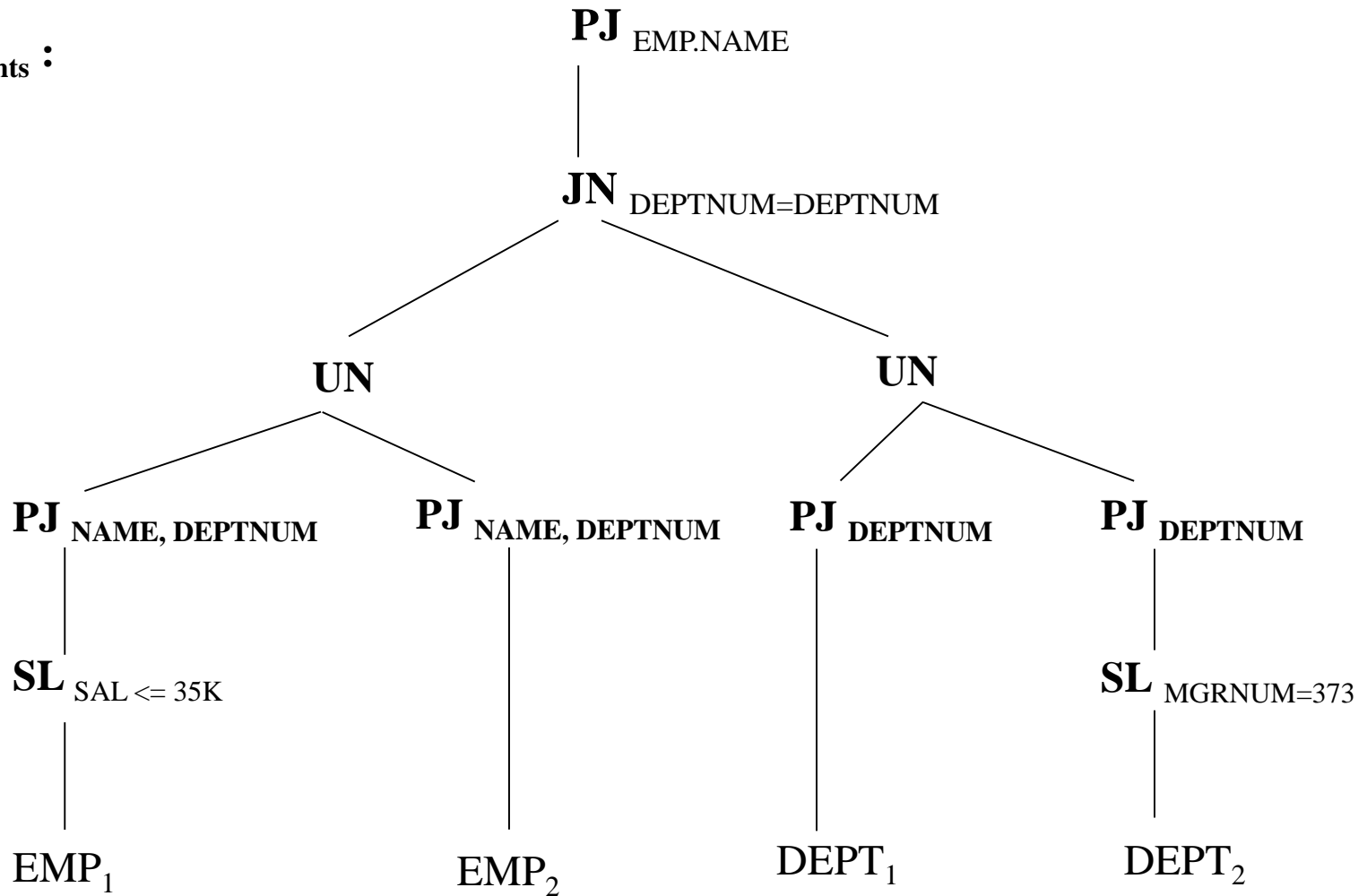
Canonical Expression

After Applying Criteria 2 -



Canonical Expression

Q fragments :



Equivalent Global and Fragment Query

Q_{global} :

PJ_{EMP.NAME} ((**EMP**
JN_{DEPTNUM=DEPTNUM} **SL**_{MGRNUM=373}
DEPT) **DF** (**SL**_{SAL > 35K} **EMP**
JN_{DEPTNUM=DEPTNUM} **SL**_{MGRNUM=373}
DEPT))

Q_{fragments} :

PJ_{EMP.NAME} ((**PJ**_{NAME,DEPTNUM}
SL_{SAL ≤ 35K} **EMP**₁ **UN** **PJ**_{NAME,DEPTNUM}
EMP₂)) **JN**_{DEPTNUM=DEPTNUM}
(**PJ**_{DEPTNUM} **DEPT**₁ **UN** **PJ**_{DEPTNUM}
SL_{MGRNUM=373} **DEPT**₂))

(non-distributed)

Q_{transformed} :

PJ_{EMP.NAME} ((**PJ**_{NAME,DEPTNUM}
SL_{SAL ≤ 35K} **EMP**) **JN**_{DEPTNUM=DEPTNUM}
(**PJ**_{DEPTNUM} **SL**_{MGRNUM=373} **DEPT**))

(distributed)

Steps for Global query to Fragment Query (summary)

-
- The diagram illustrates the steps for converting a global query into a fragment query, organized into two main phases: ND (Non-Distributed) and D (Distributed). The ND phase includes steps 1 through 4, which focus on generating and simplifying the operator tree. The D phase includes steps 5 through 7, which focus on applying canonical expressions and criteria to obtain the final fragments. Red text highlights key terms and actions in each step.
- ND
 - 1. Generate the equivalent operator tree (T_{global}) for the given query (Q_{global}).
 - 2. Find the common sub-expression (R) from T_{global} .
 - 3. Apply rules to remove R and obtain simplified tree T_{removed} .
 - 4. Apply criteria – 1 and 2 on T_{removed} to obtain final simplified operator tree $T_{\text{transformed}}$.
 - D
 - 5. Apply canonical expression on $T_{\text{transformed}}$ to obtain the canonical form $T_{\text{canonical}}$.
 - 6. Apply criterion- 1 and 2 on $T_{\text{canonical}}$ to obtain $T_{\text{fragments}}$.
 - 7. Write the query $Q_{\text{fragments}}$ from $T_{\text{fragments}}$.

Example 2

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

Consider the following global query:

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

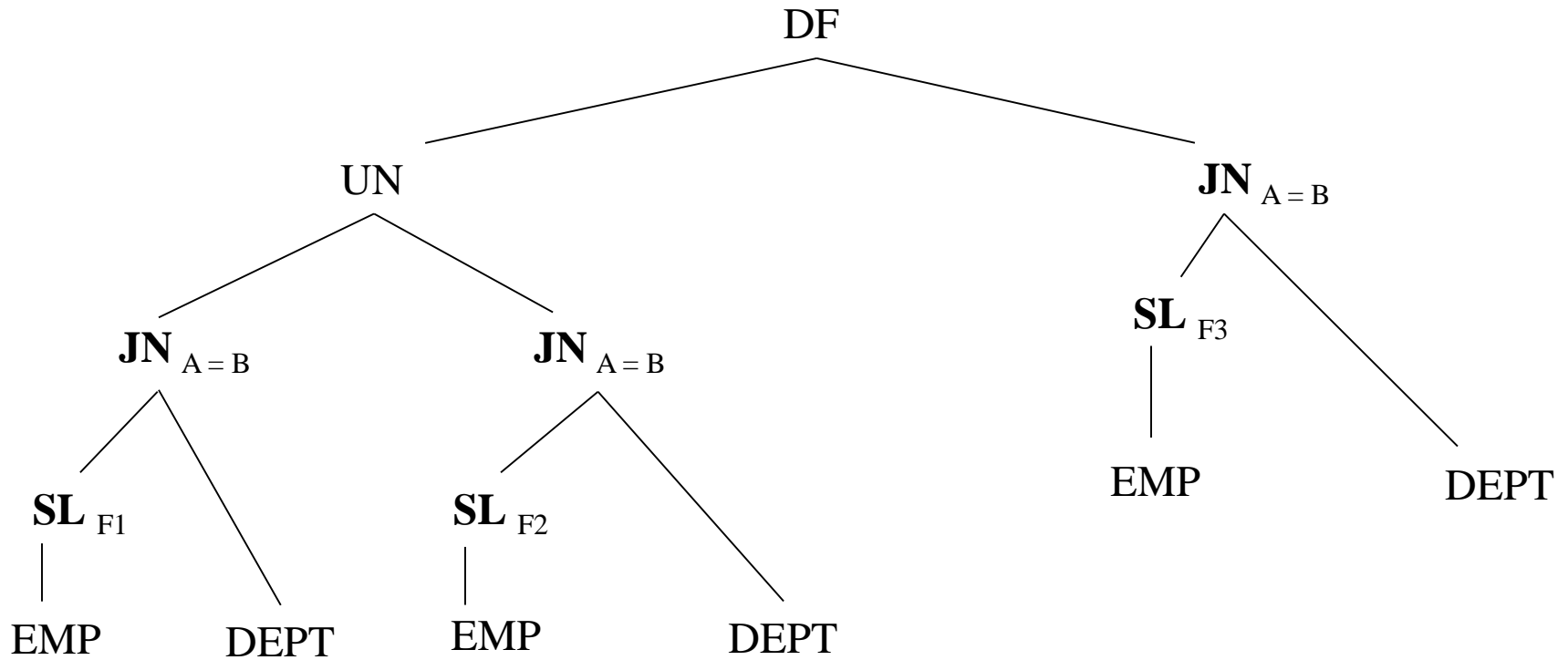
Here,

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

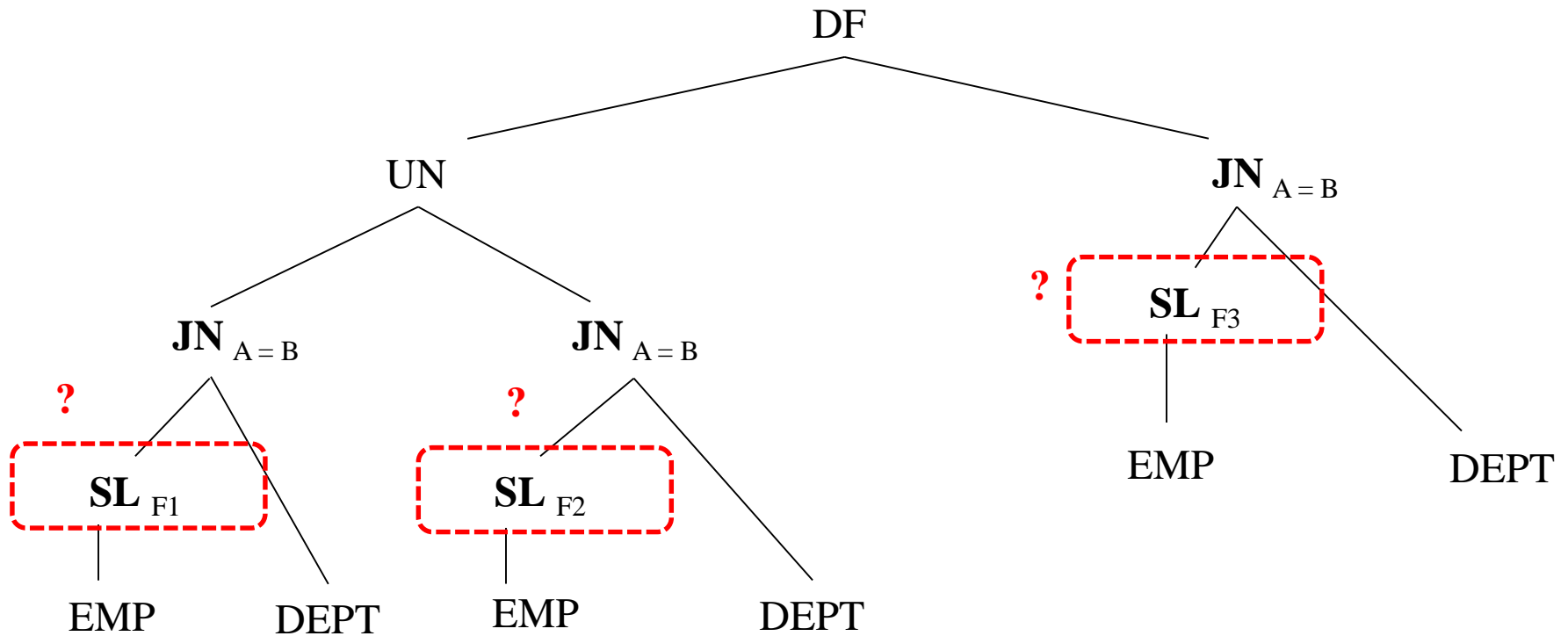
Imagine, $A = B = \text{DEPTNUM}$

- Draw the operator tree.
- Perform step-by-step transformations to simplify the tree, indicating which rule and criterion is applied at each step.
- Write down the query from the simplified tree.
- Transform the simplified query into fragment query by applying canonical expression on the simplified tree. Consider, *EMP* has three fragments - EMP_1 , EMP_2 and EMP_3

Operator Tree

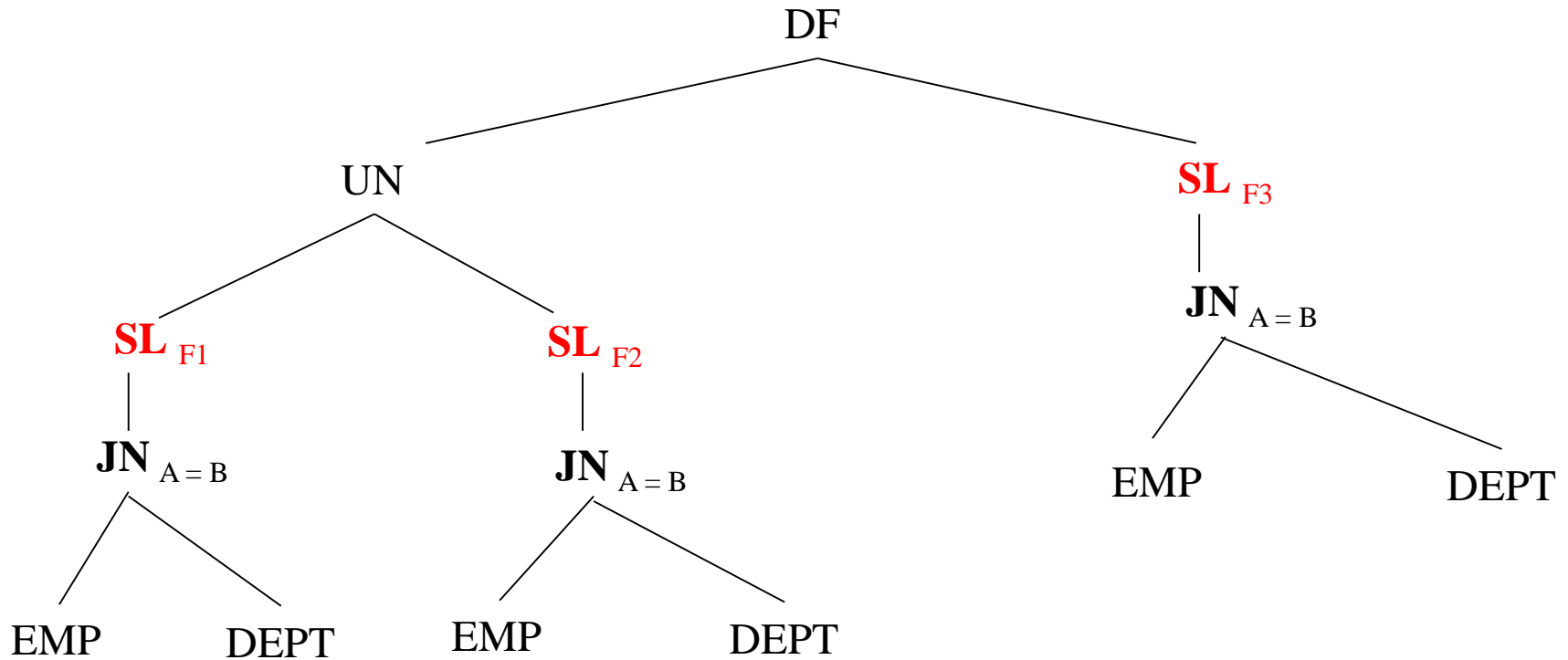


Finding Common Sub-Expression



Finding Common Sub-expression

Any common portion?

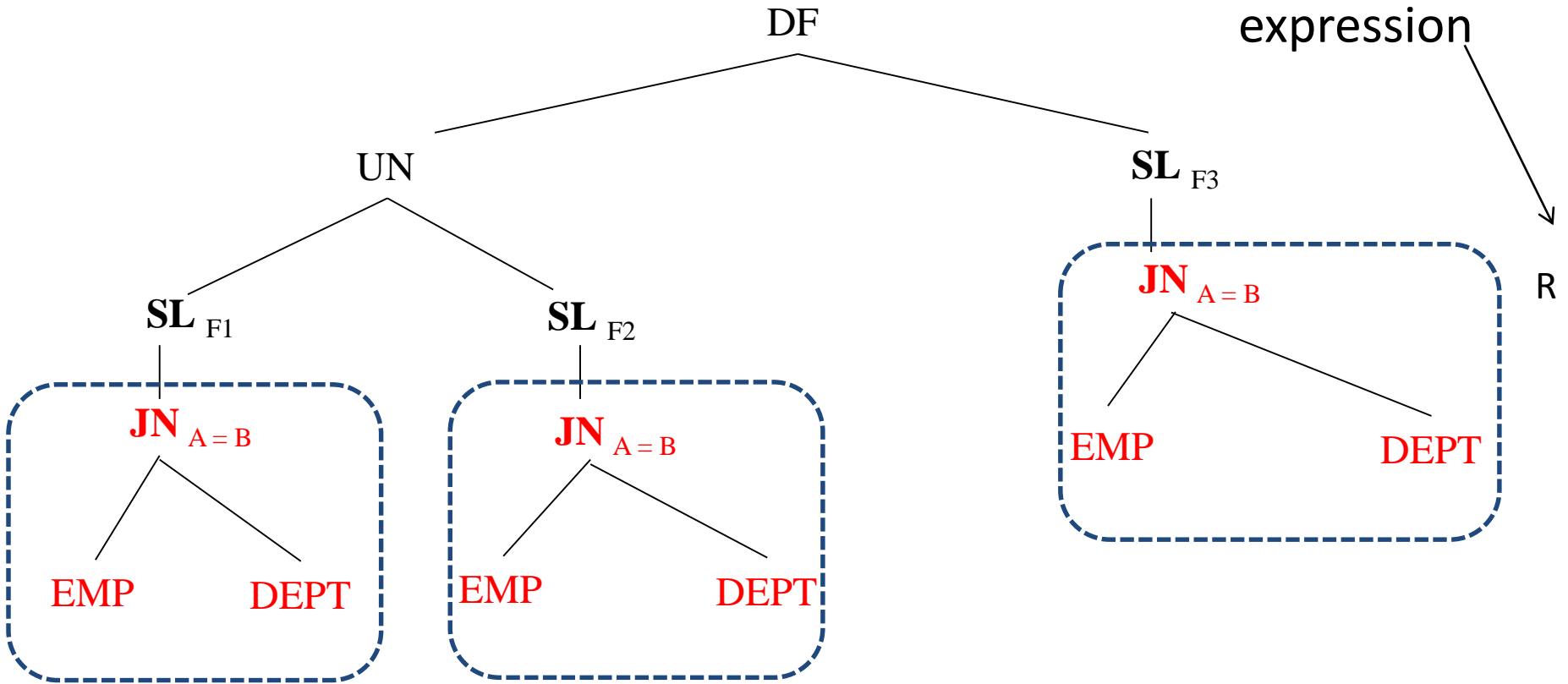


Finding Common Sub-expression

Any common portion?

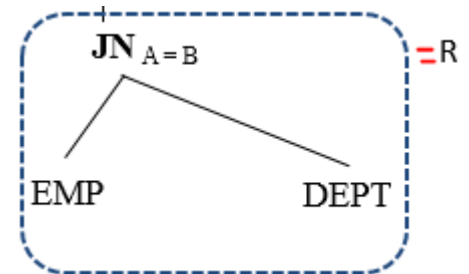
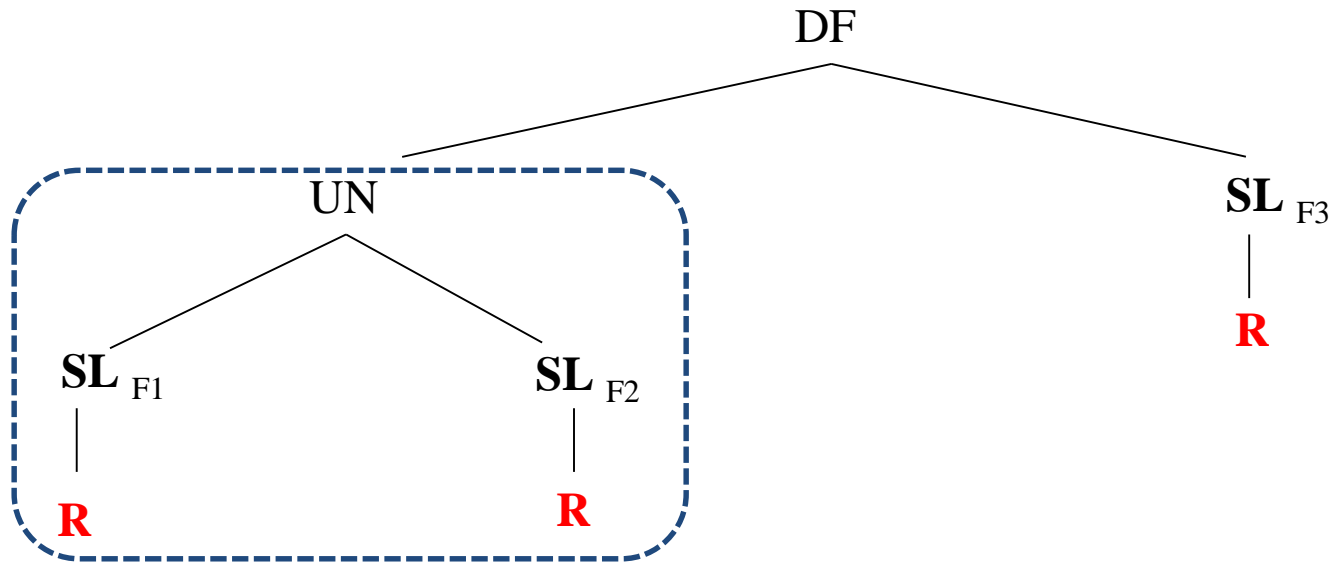
Common sub-expression

R



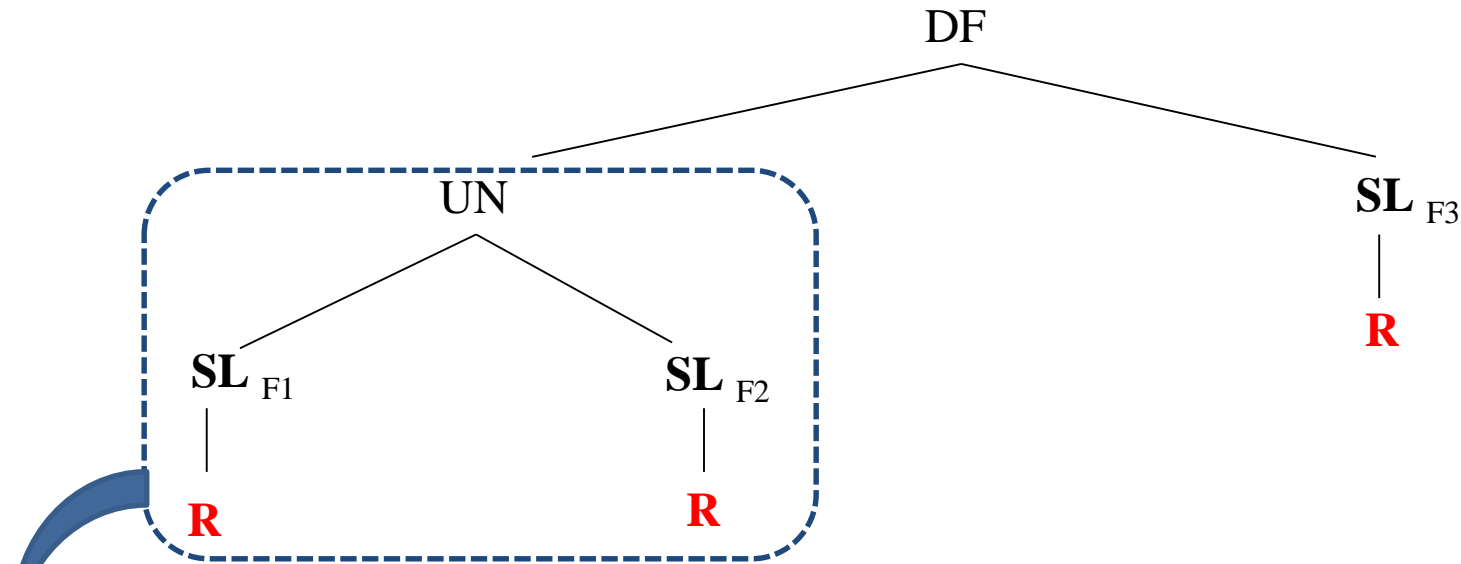
Finding Common Sub-expression

Any common portion?

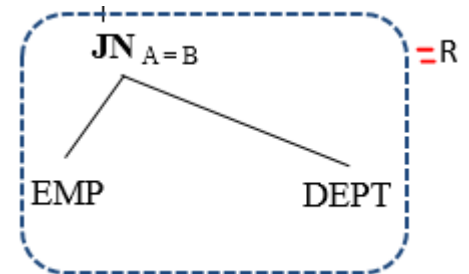


Finding Common Sub-expression

Any common portion?

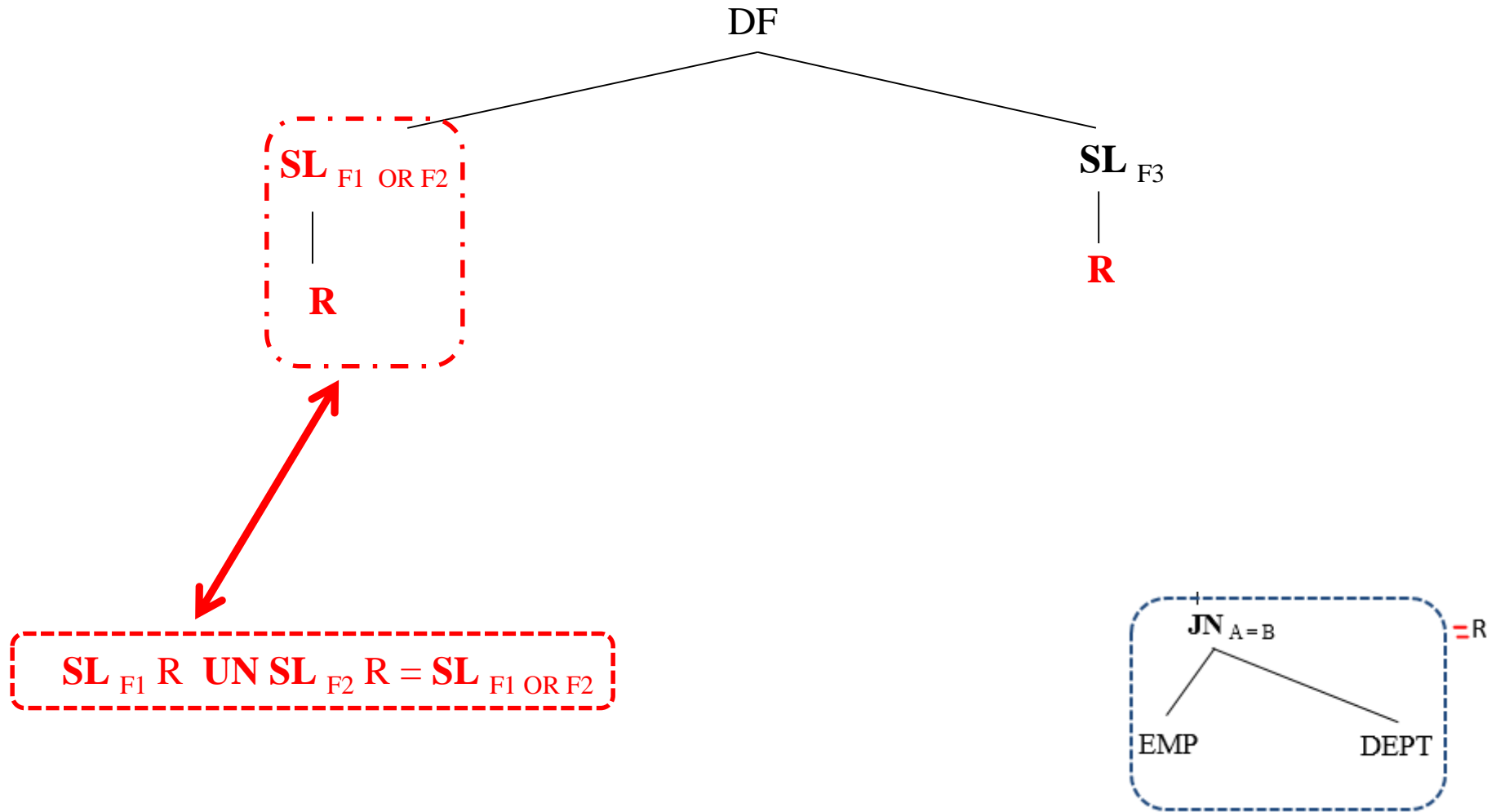


We can write it as $SL_{F1} R \text{ UN } SL_{F2} R$ which is Rule 8 !



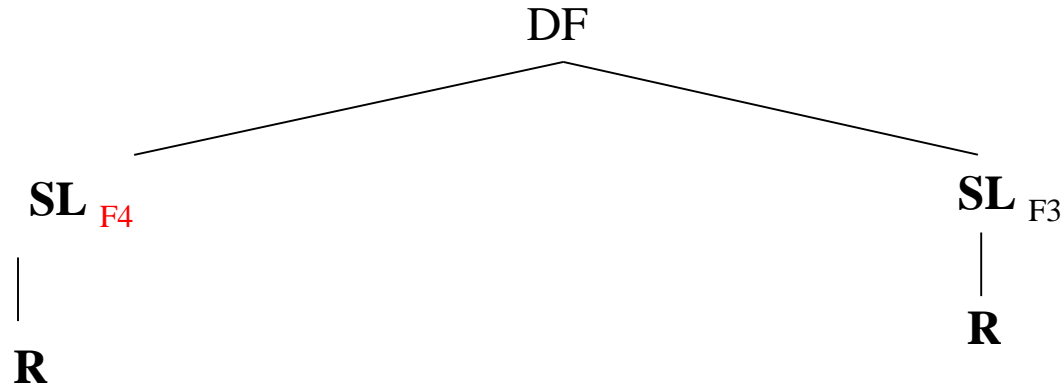
Removing Common Sub-expression

Any common portion?

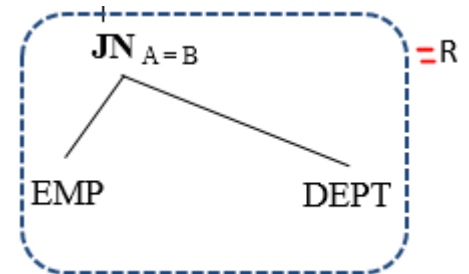


Finding Common Sub-expression

Any common portion?

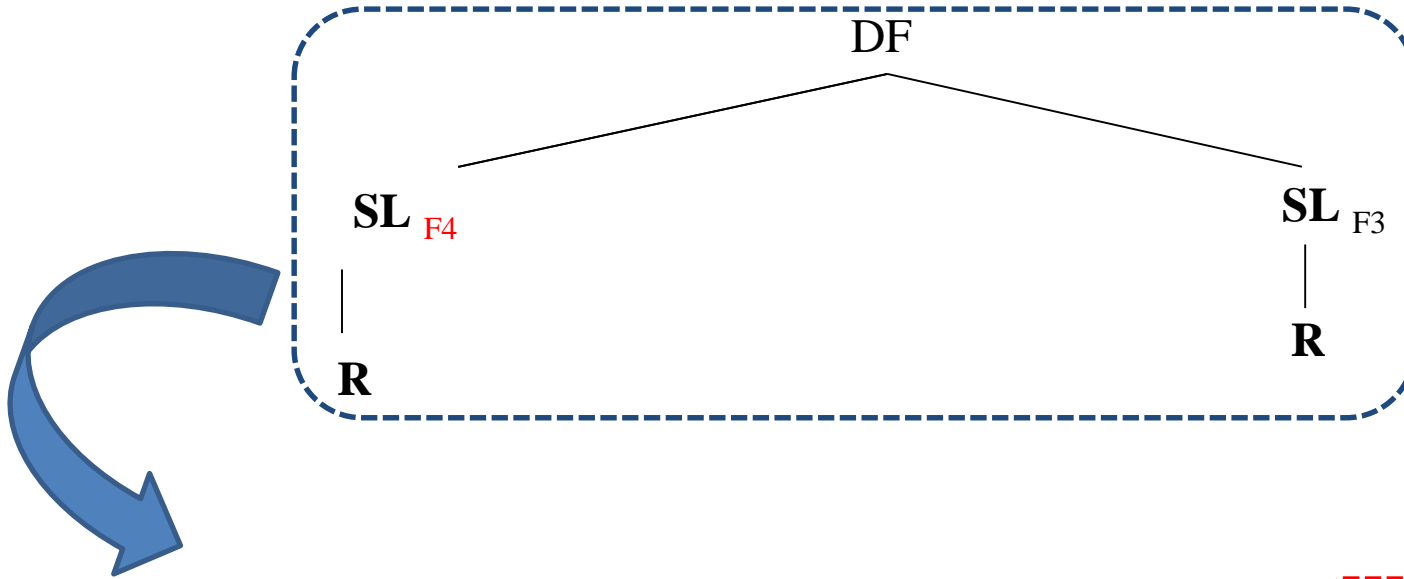


Let, $F4 = F1 \text{ OR } F2$



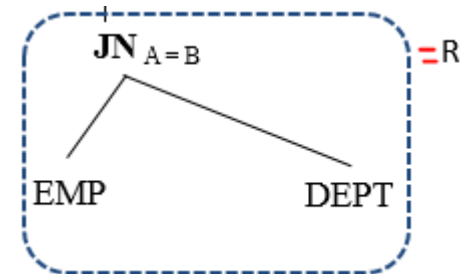
Finding Common Sub-expression

Any common portion?



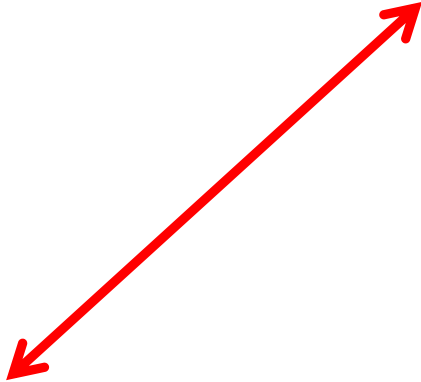
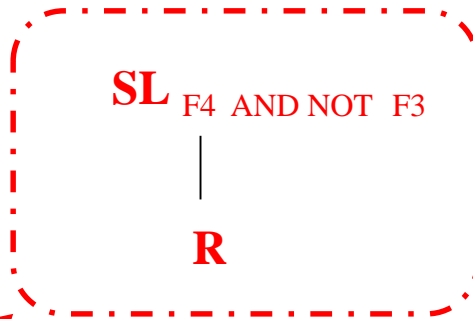
We can write it as **SL_{F4} R DF SL_{F3} R** which is Rule 9 !

Let, F4 = F1 OR F2

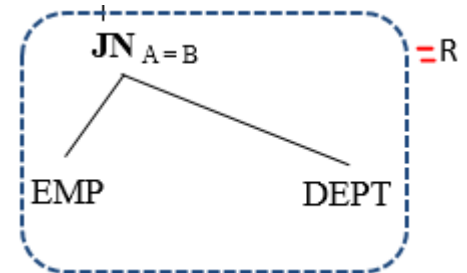


Removing Common Sub-expression

Any common portion?



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

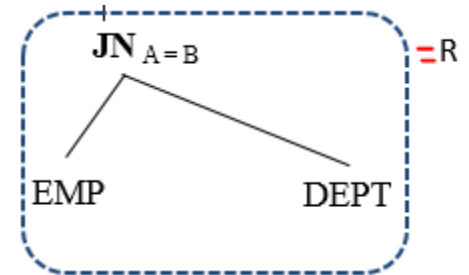


Removing Common Sub-expression

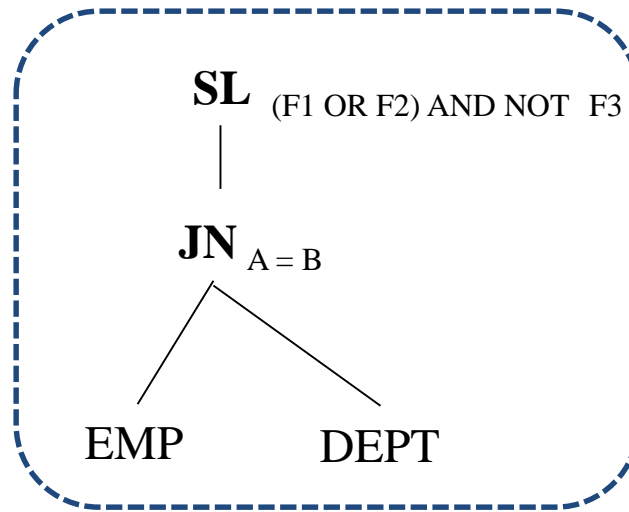
SL (F1 OR F2) AND NOT F3

|

R



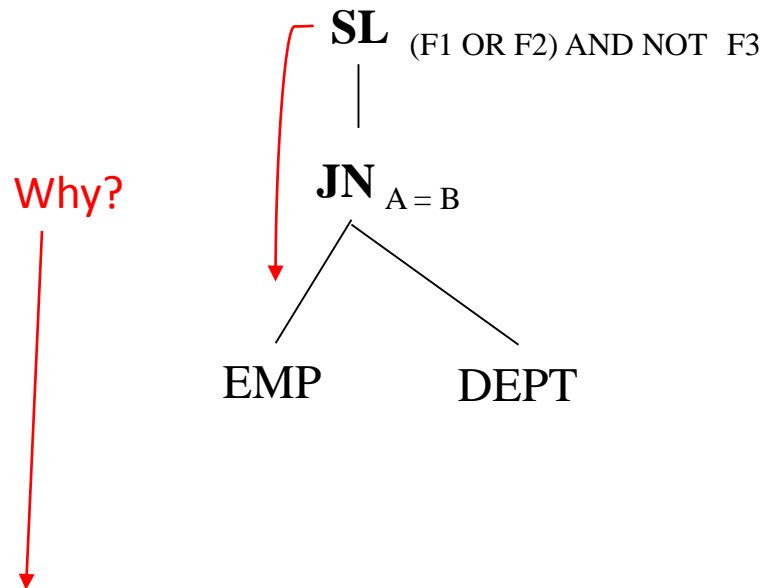
Removing Common Sub-expression



Can we apply Criterion 1 and/or 2?

Simplification

Applying Criterion 2 -

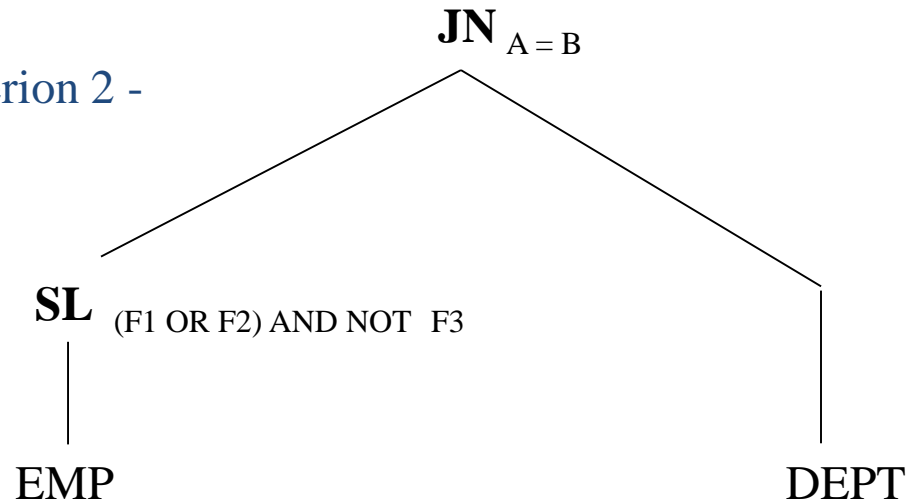


Because in the original query, all **SL**_{F1}, **SL**_{F2}, **SL**_{F3} were applied on EMP relation.

Simplification

Q transformed:

After Applying Criterion 2 -



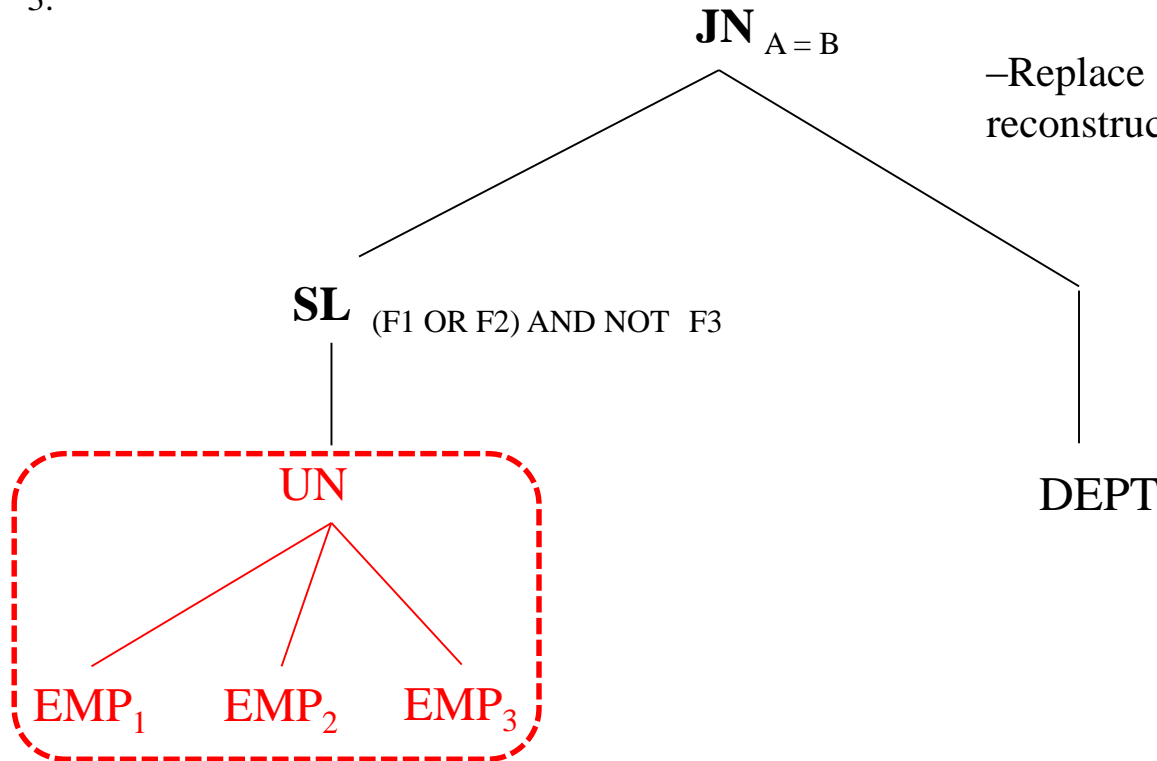
Canonical Expression

- Given, EMP has 2 horizontal fragments: EMP₁, EMP₂, EMP₃.

How to convert this tree so that the leaves becomes the fragment?

–Consider the reconstruction property of Horizontal Fragmentation.

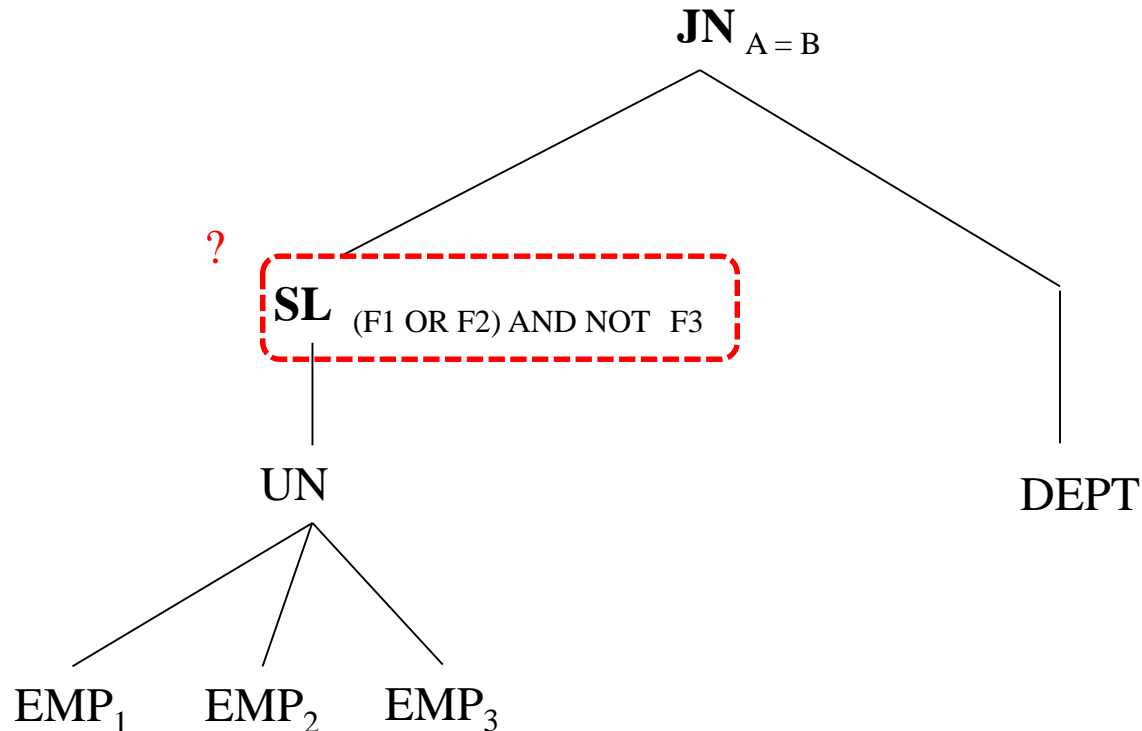
–Replace the leaf with the sub-tree of the reconstruction expression.



Canonical Expression

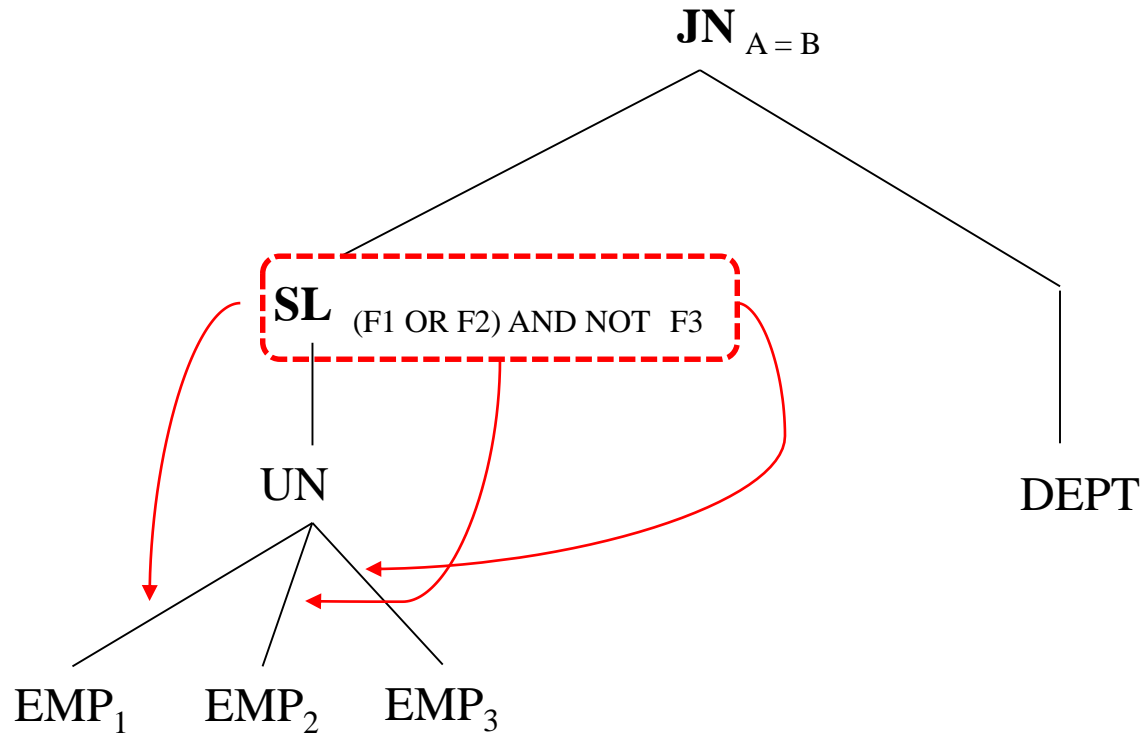
Do you think it is still simplified?

Can we apply Criteria 1 and/or 2 ?



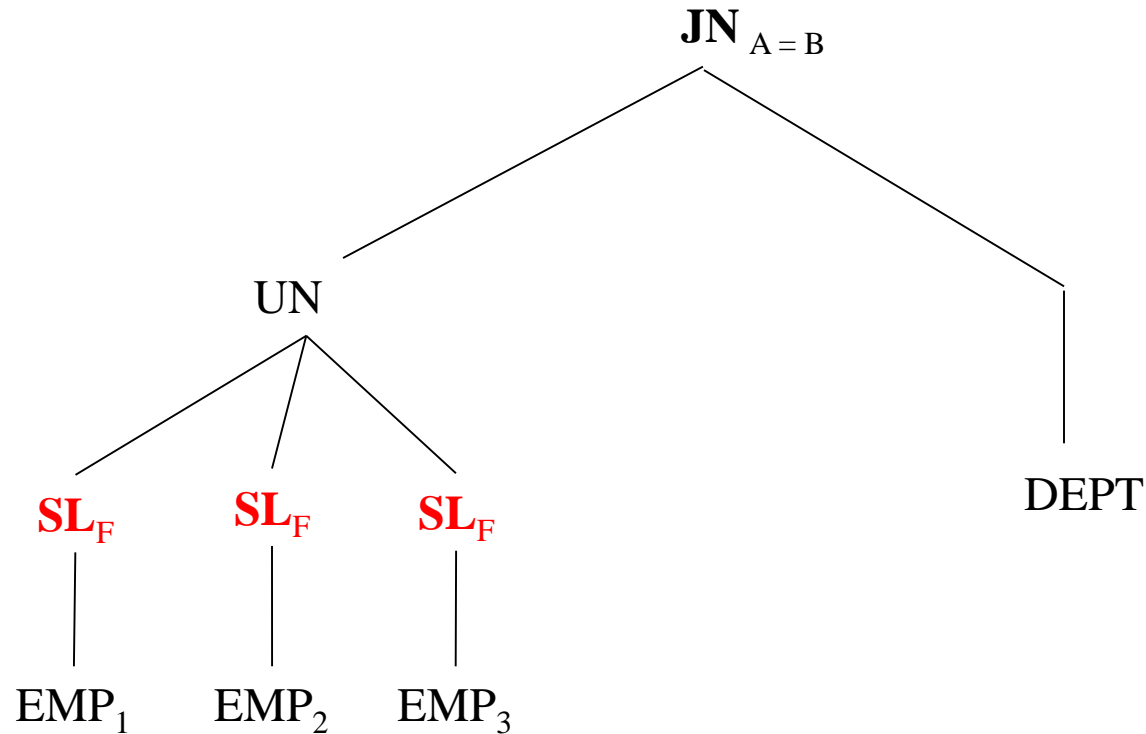
Canonical Expression

Applying Criteria 2 -



Canonical Expression

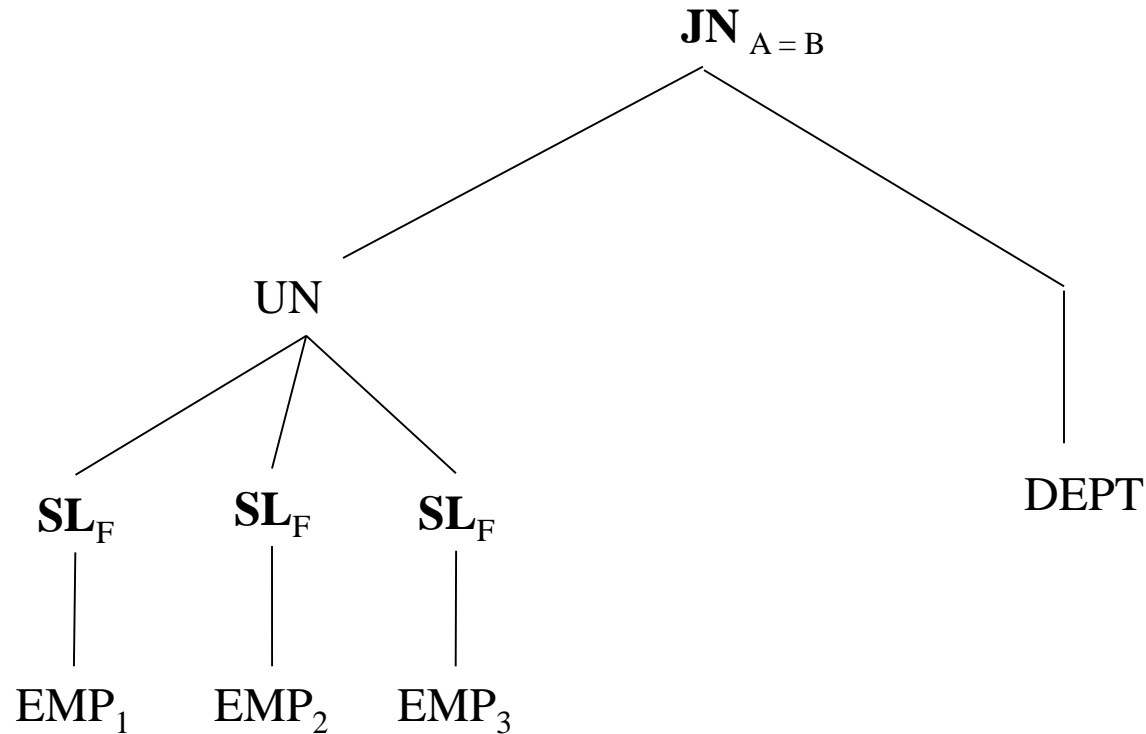
After Applying Criteria 2 -



Let $F = (F1 \text{ OR } F2) \text{ AND NOT } F3$

Canonical Expression

Q fragments :



Let $F = (F1 \text{ OR } F2) \text{ AND NOT } F3$

Equivalent Global and Fragment Query

$Q_{\text{global}} : ((SL_{F1} EMP \text{ JN}_{A=B} DEPT) \text{ UN } (SL_{F2} EMP \text{ JN}_{A=B} DEPT)) \text{ DF } (SL_{F3} EMP \text{ JN}_{A=B} DEPT)$



(non-distributed)

$Q_{\text{transformed}} : SL_{(F1 \text{ OR } F2) \text{ AND NOT } F3} EMP \text{ JN}_{A=B} DEPT$



(distributed)

$Q_{\text{fragments}} : (SL_{(F1 \text{ OR } F2) \text{ AND NOT } F3} EMP_1 \text{ UN } SL_{(F1 \text{ OR } F2) \text{ AND NOT } F3} EMP_2 \text{ UN } SL_{(F1 \text{ OR } F2) \text{ AND NOT } F3} EMP_3) \text{ JN}_{A=B} DEPT$

Exercise 1

Practise

* Draw Operator Tree for the following queries:

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

Query:

Consider SUPPLY and DEPT both have two horizontal fragments.

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

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]
- iii. Transform the simplified query into fragment query by applying canonical expression based on the given fragmentation schema. [2]
- iv. Write the equivalent query obtained from the simplified tree. [1]

Exercise 2

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))

Consider the Fragmentation Schemata and answer the questions of **Exercise 3**.

Exercise 3

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.

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

Now, answer the following questions.

- Draw the *operator tree*. [2]
- Perform step-by-step transformations to simplify the operator tree, indicating which rule and criterion is applied at each step. [5]
- Transform the simplified query into fragment query by applying canonical expression based on the given fragmentation schema. [2]
- Write the equivalent query obtained from the simplified tree. [1]

Exercise 4

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

Consider the following global query:

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

Here,

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

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

- Draw the operator tree.
- Perform step-by-step transformations to simplify the tree, indicating which rule and criterion is applied at each step.
- Write down the query from the simplified tree.
- Transform the simplified query into fragment query by applying canonical expression on the simplified tree. Consider, *EMP* has three fragments - EMP_1 , EMP_2 and EMP_3

Exercise 5

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

Consider the following global query:

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

Here,

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

- Draw the operator tree.
- Perform step-by-step transformations to simplify the tree, indicating which rule and criterion is applied at each step.
- Write down the query from the simplified tree.
- Transform the simplified query into fragment query by applying canonical expression on the simplified tree. Consider, *EMP* has three fragments - EMP_1 , EMP_2 and EMP_3

Exercise 6

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

$$(((SL_{F1} R \Join_{A=B} S) \text{ DF } (SL_{F2} R \Join_{A=B} S)) \text{ NJN } (R \Join_{A=B} S)) \\ \text{UN } (SL_{F1 \text{ AND } NOT F2} R \Join_{A=B} S)$$

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]

iv) Transform the simplified query into Fragment query by applying canonical expression. Consider relation R has three and relation S has two horizontal fragments.

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 7

Consider the following global relational schemata.

$EMP (EMPNUM, ID, NAME, SAL, AGE)$

$DEPT (ID, NAME, AREA, MGRNUM)$

Corresponding fragmentation schemata:

$$EMP_1 = SL_{SAL \leq 35K} EMP$$

$$DEPT_1 = SL_{MGRNUM = 375} DEPT$$

$$EMP_2 = SL_{SAL > 35K} EMP$$

$$DEPT_2 = SL_{MGRNUM = 373} DEPT$$

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

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]
- iii. Transform the simplified query into fragment query by applying canonical expression based on the given fragmentation schema. [2]
- iv. Write the equivalent query obtained from the simplified tree. [1]