CSE 4125: Distributed Database Systems
Chapter – 5

Translation of Global Queries to Fragment Queries.

(Part – D)

Topics to be discussed –

- Some properties/rules for query simplification
- Finding common sub – expressions in operator tree
- Removing common sub-expressions using the rules

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
- $R \text{ NJN } SL_F R \leftrightarrow SL_F R$ → 4
- $R \text{ UN } SL_F R \leftrightarrow R$ → 5
- $R \text{ DF } SL_F R \leftrightarrow SL_{\text{NOT } F} R$ → 6
- $(SL_{F1} R) \text{ NJN } (SL_{F2} R) \leftrightarrow SL_{F1 \text{ AND } F2} R$ → 7
- $(SL_{F1} R) \text{ UN } (SL_{F2} R) \leftrightarrow SL_{F1 \text{ OR } F2} R$ → 8
- $(SL_{F1} R) \text{ DF } (SL_{F2} R) \leftrightarrow SL_{F1 \text{ AND NOT } F2} R$ → 9

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

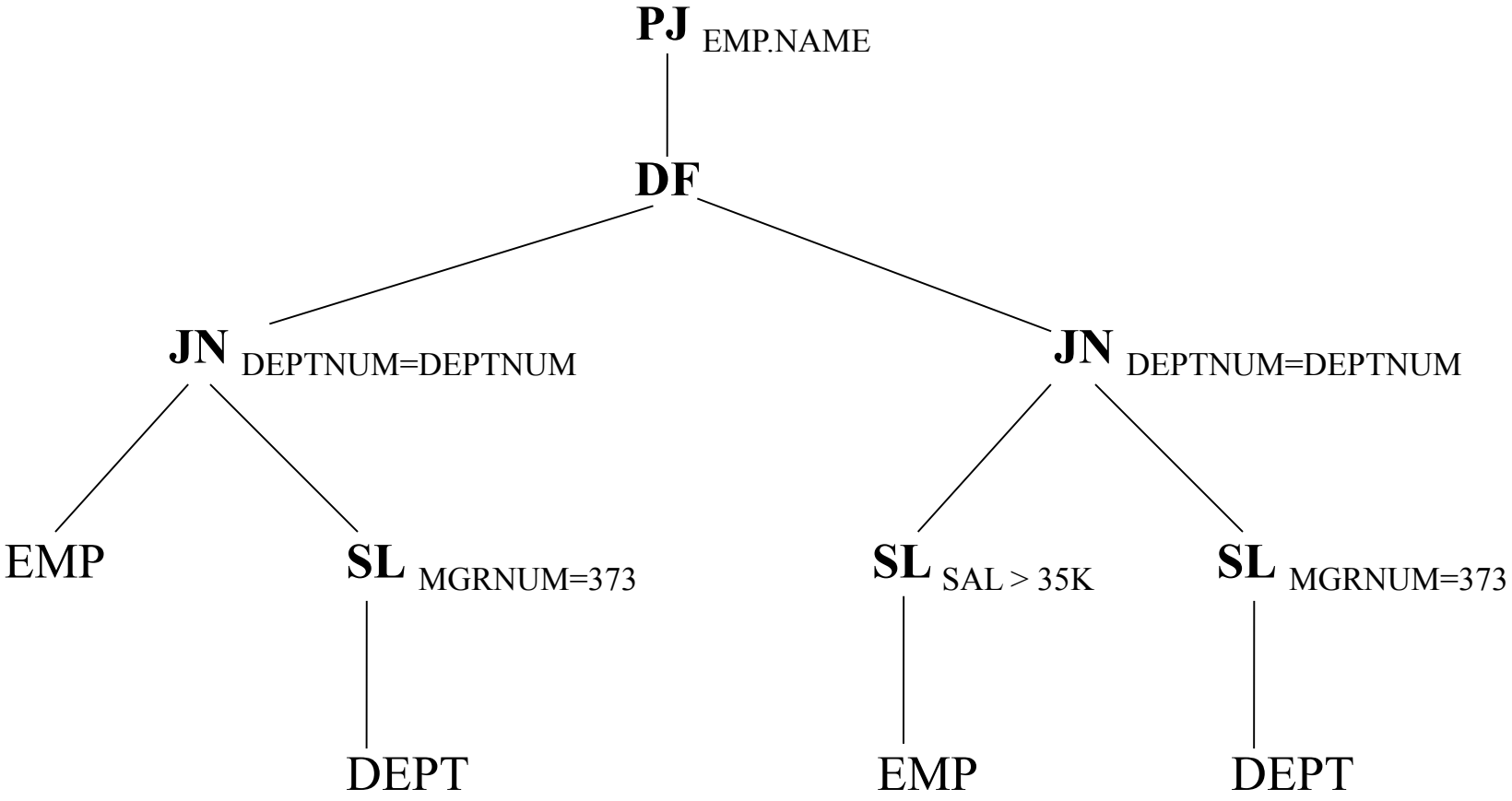
Finding & Removing Common Sub Expression

Example 1

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

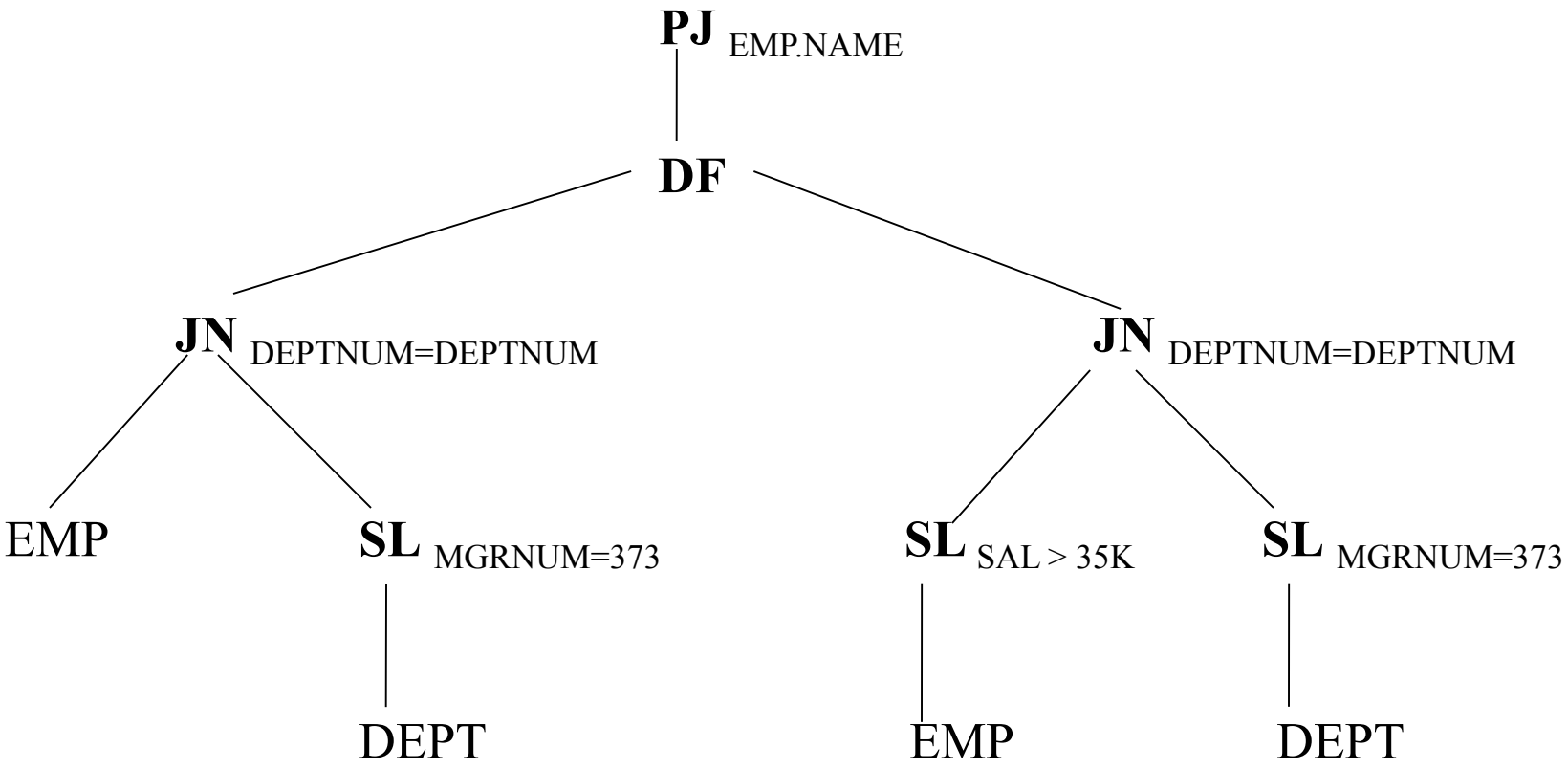
Given query and Operator Tree

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



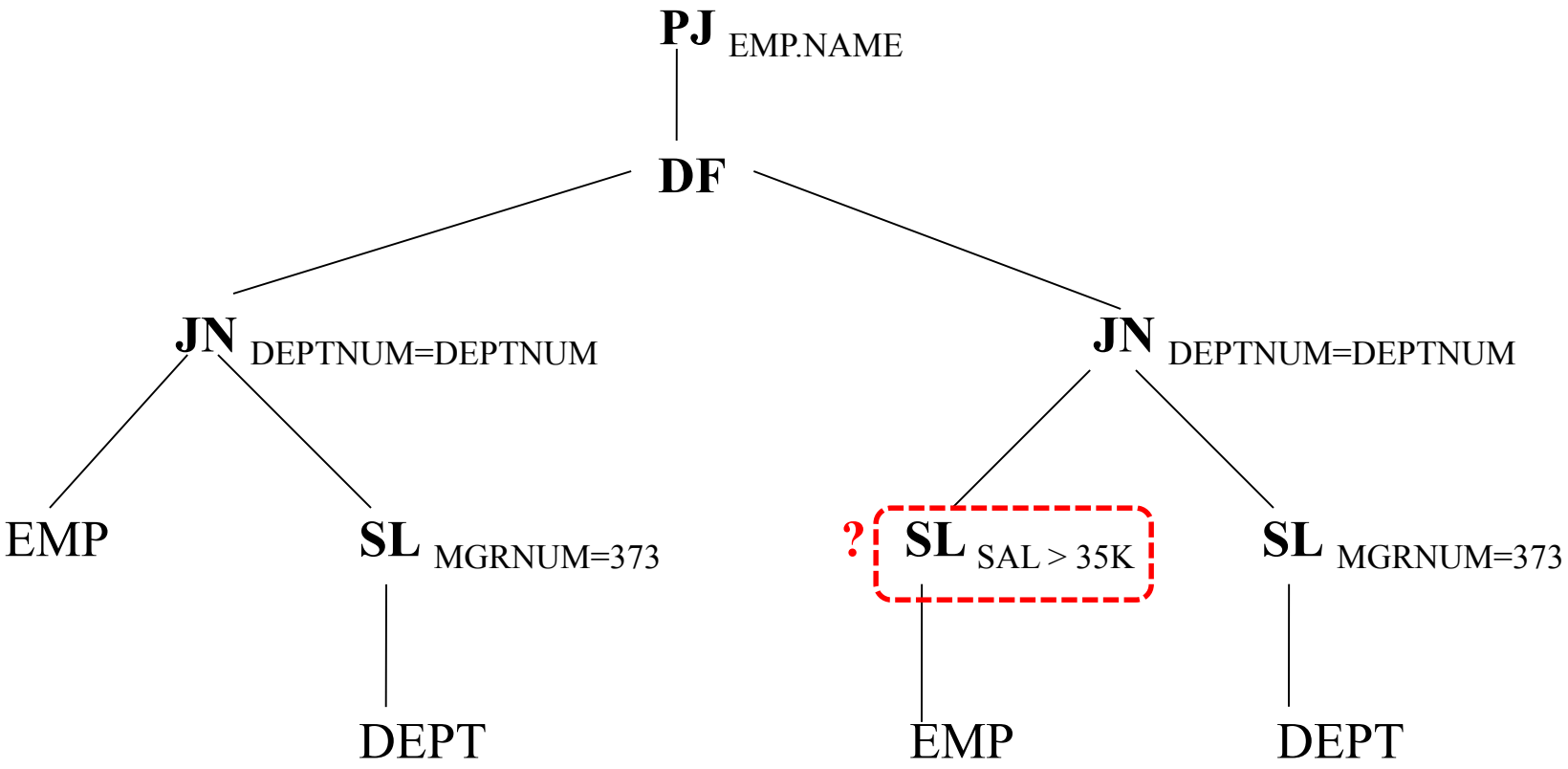
Finding Common Sub-expression

Any common portion?



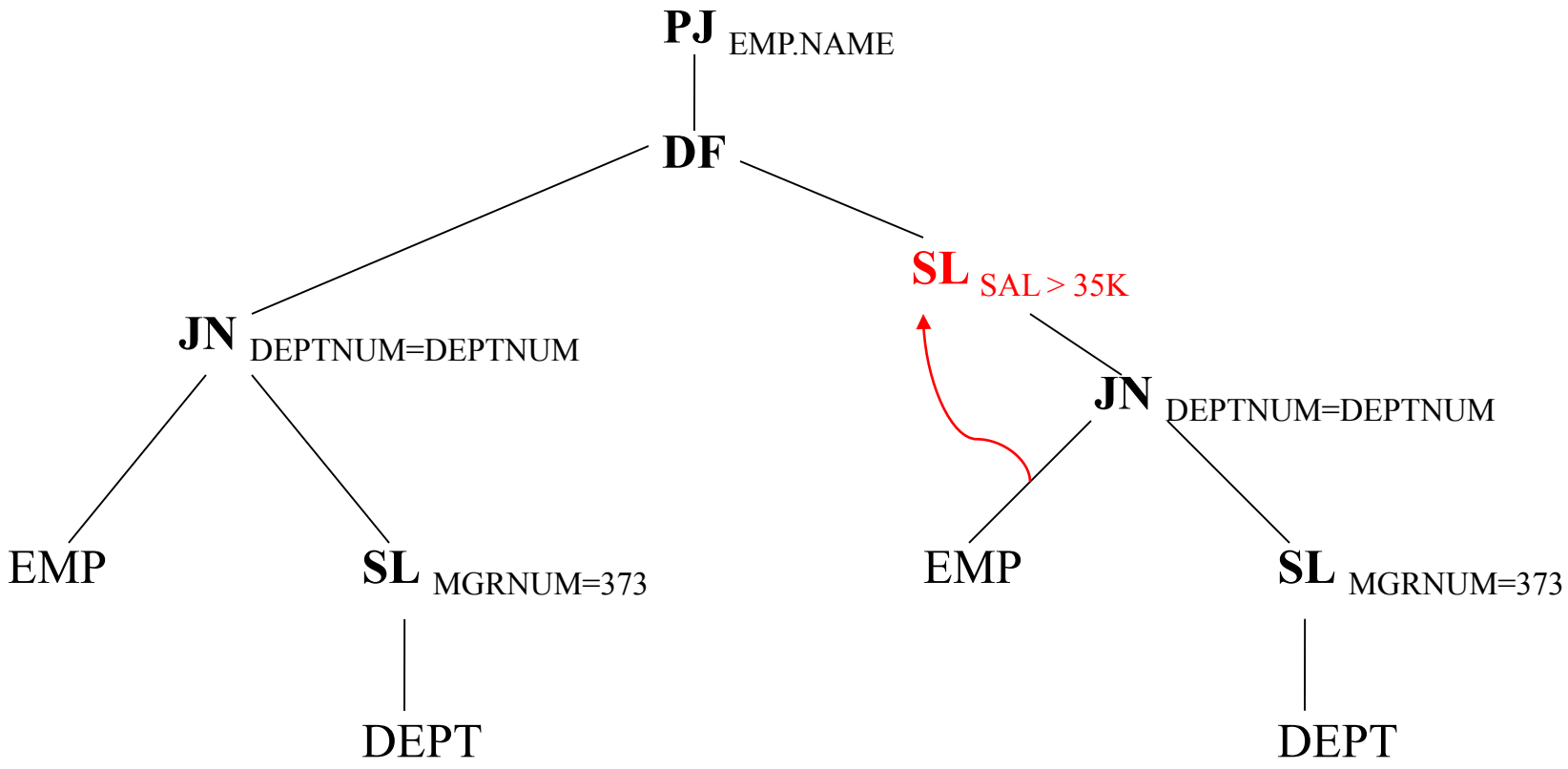
Finding Common Sub-expression

Any common portion?

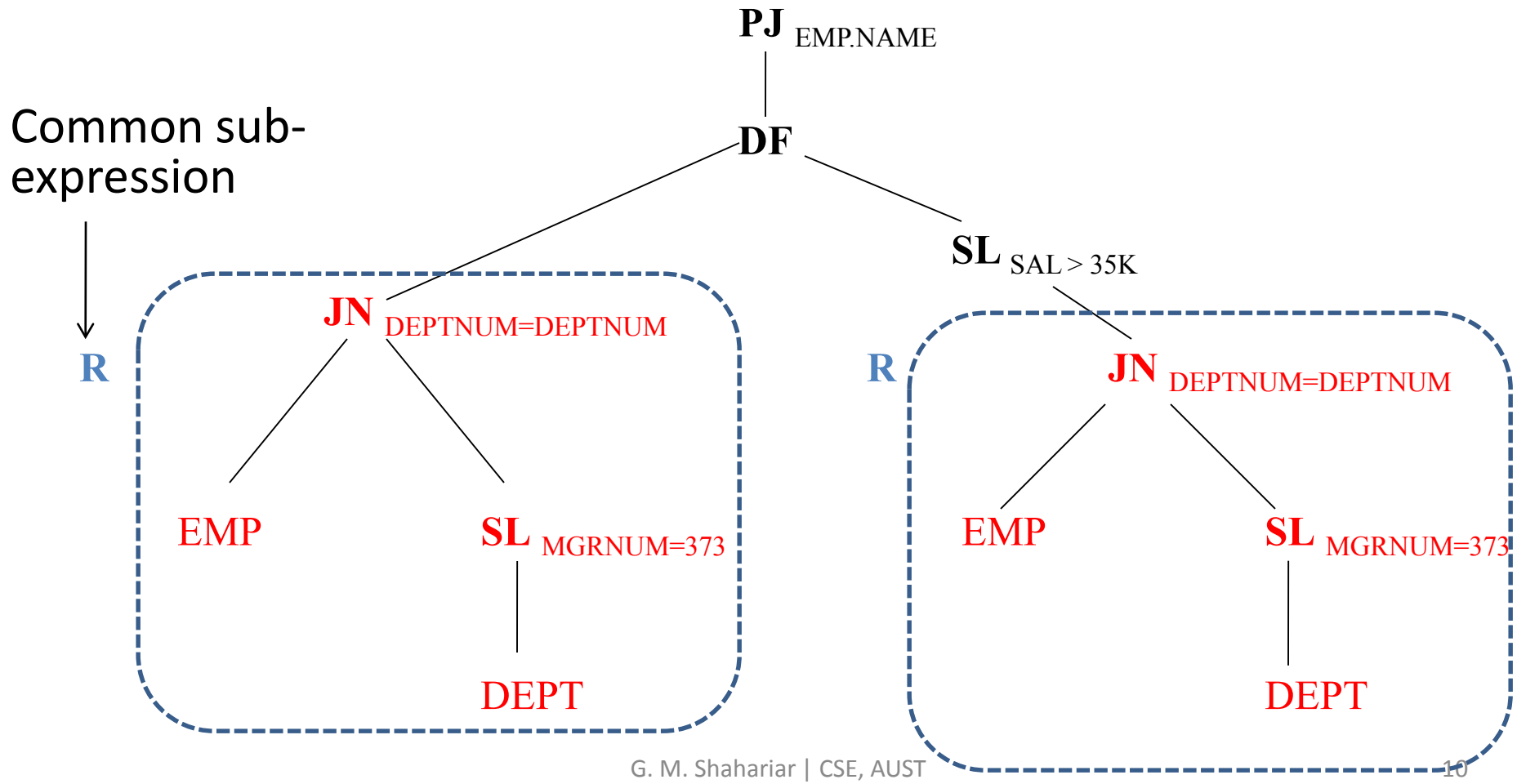


Finding Common Sub-expression

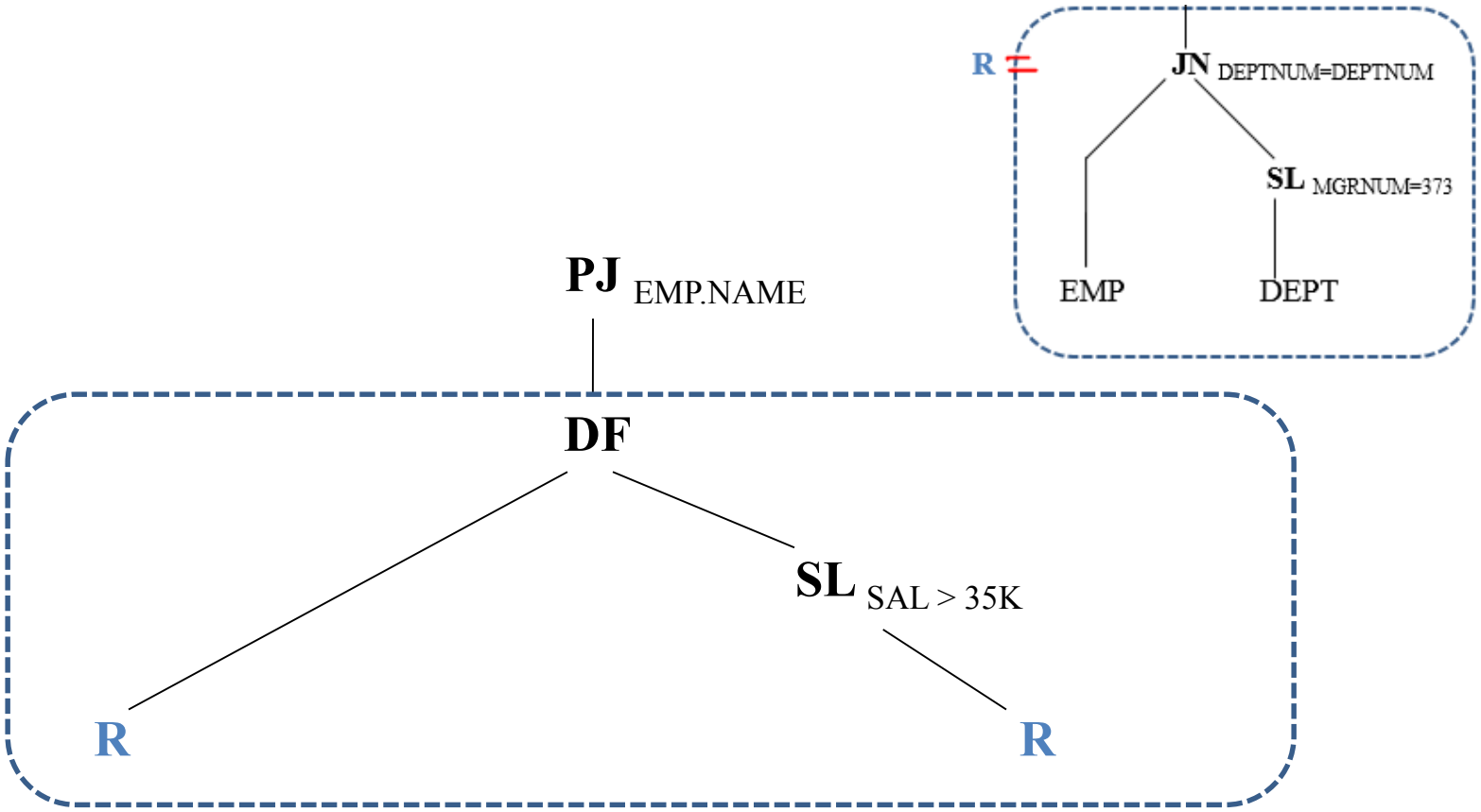
Any common portion? NOW?



Finding Common Sub-expression

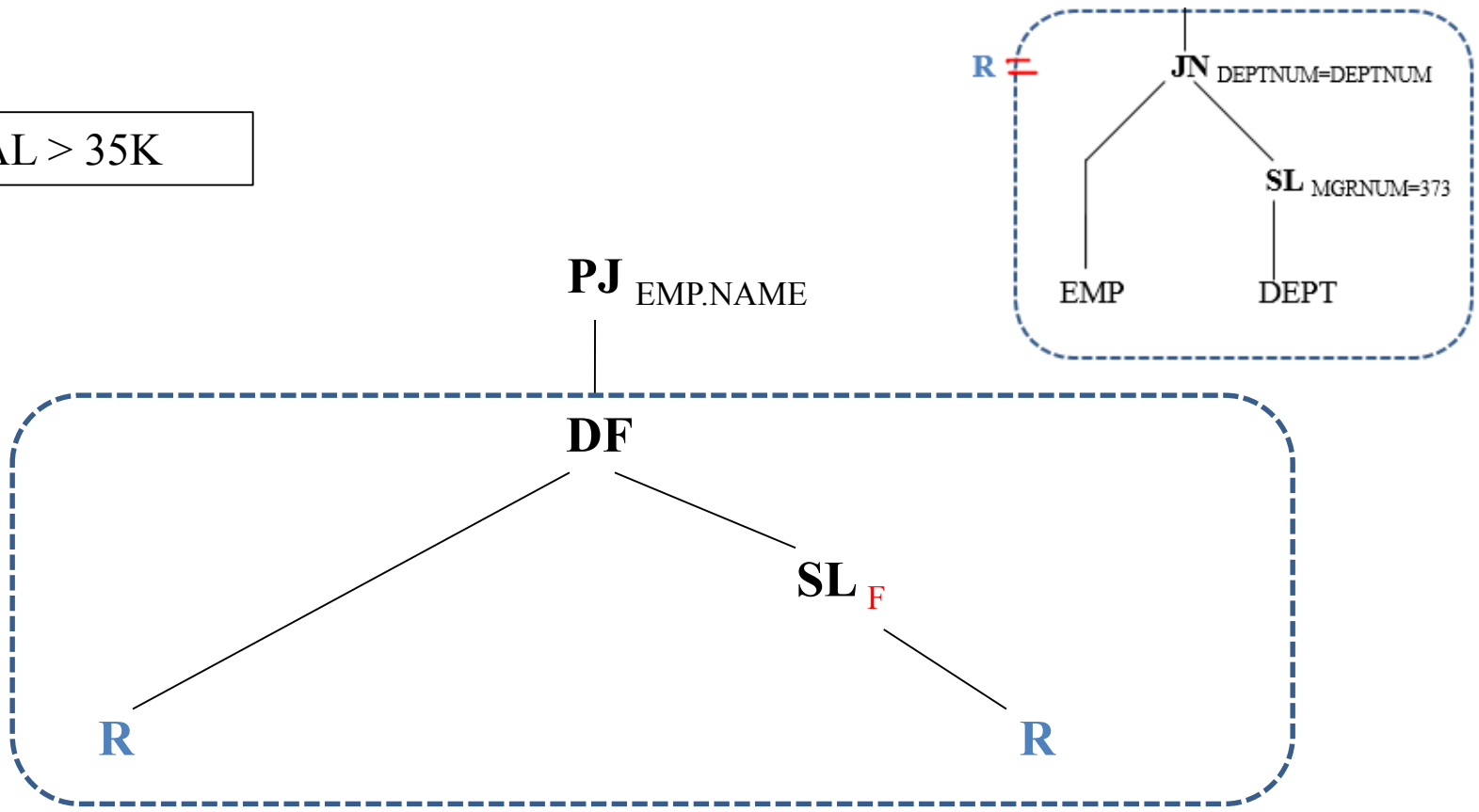


Finding Common Sub-expression



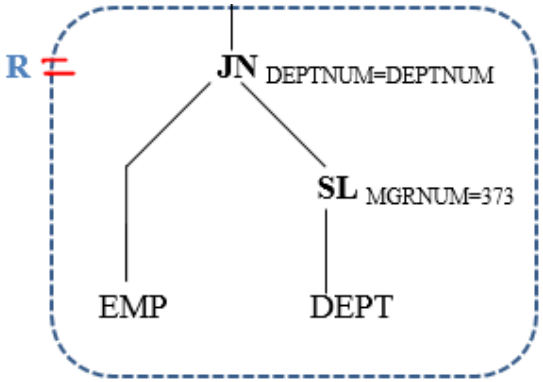
Finding Common Sub-expression

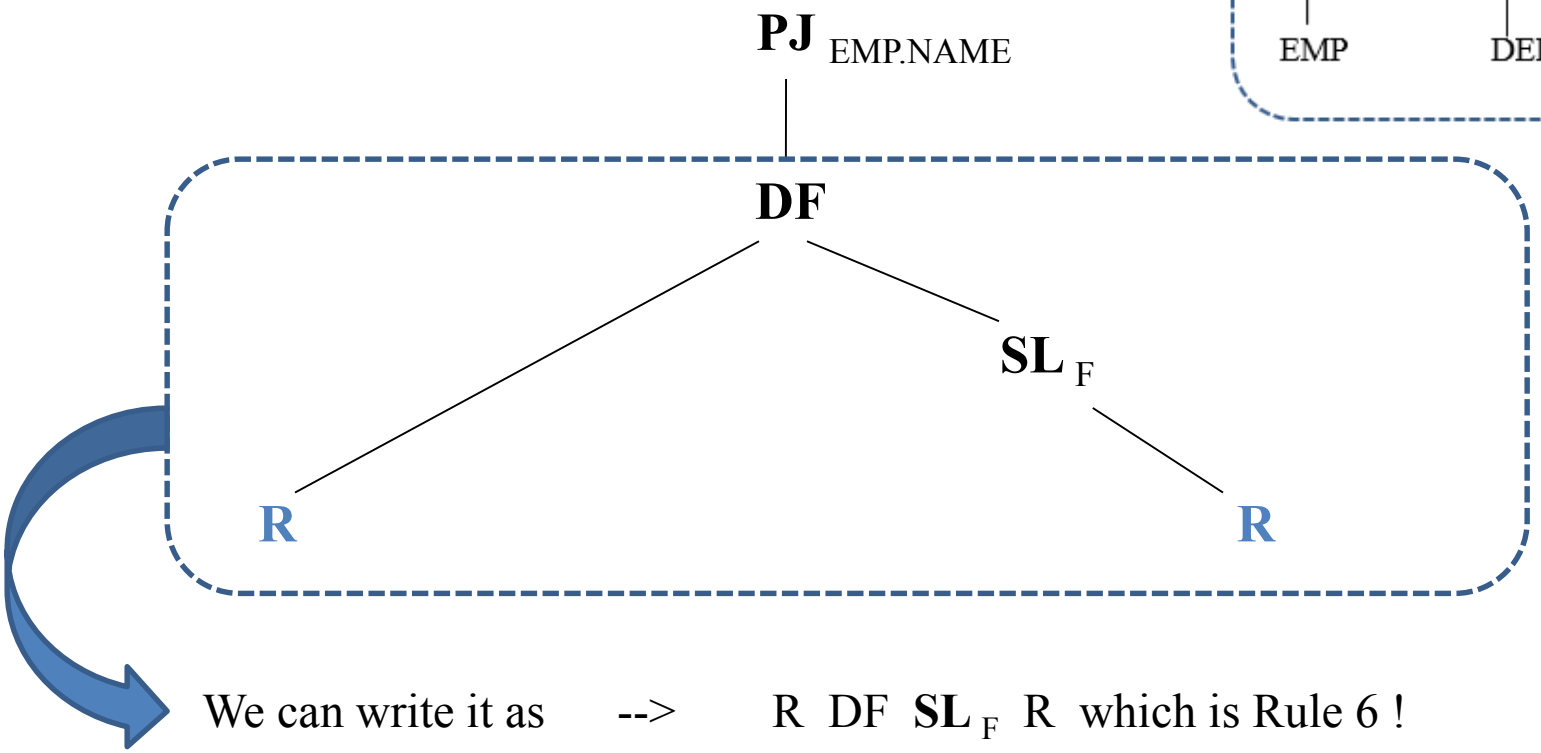
F = SAL > 35K



Finding Common Sub-expression

F = SAL > 35K

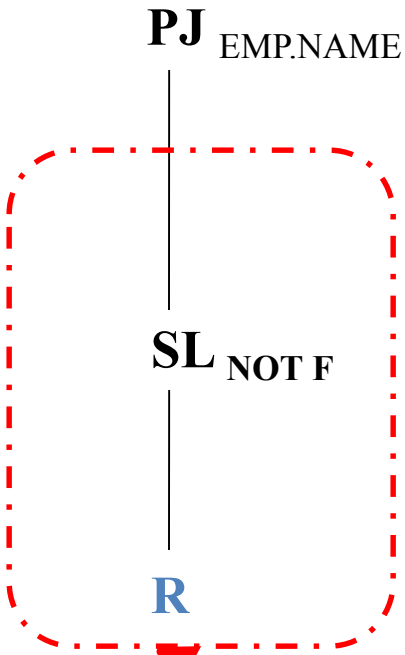


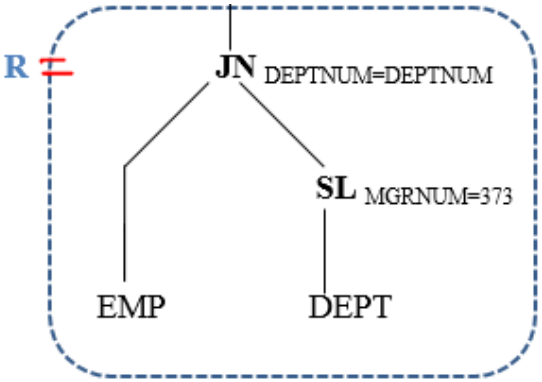


We can write it as --> R DF SL_F R which is Rule 6 !

Removing Common Sub-expression

F = SAL > 35K

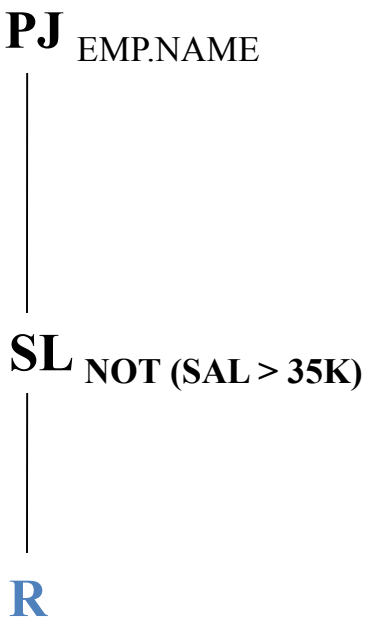


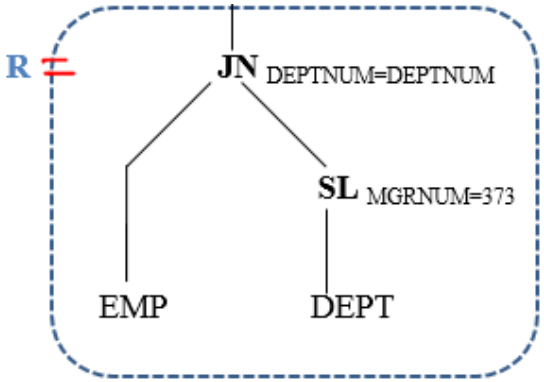


- R DF SL_F R ↔ $SL_{NOT\,F}$ R

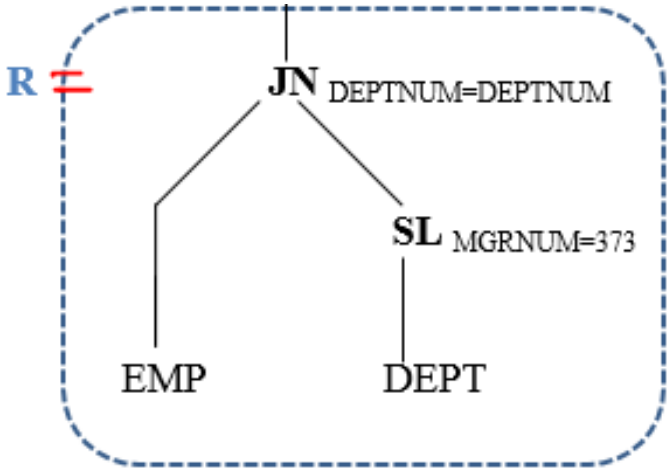
Removing Common Sub-expression

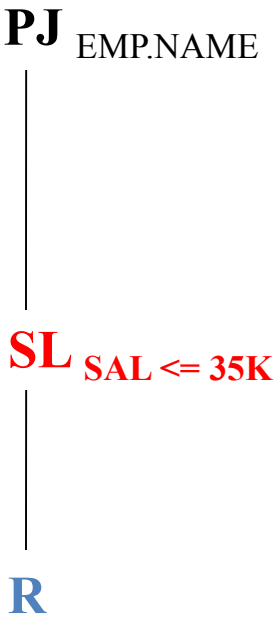
F = SAL > 35K



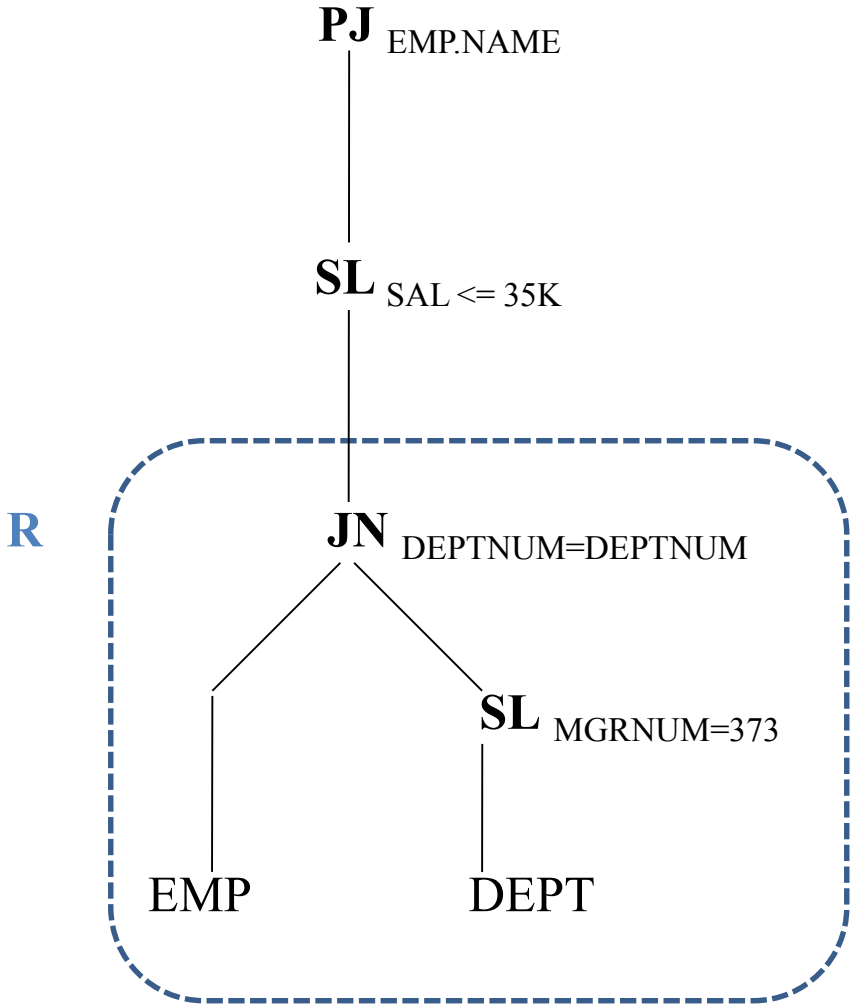


Removing Common Sub-expression





Removing Common Sub-expression



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

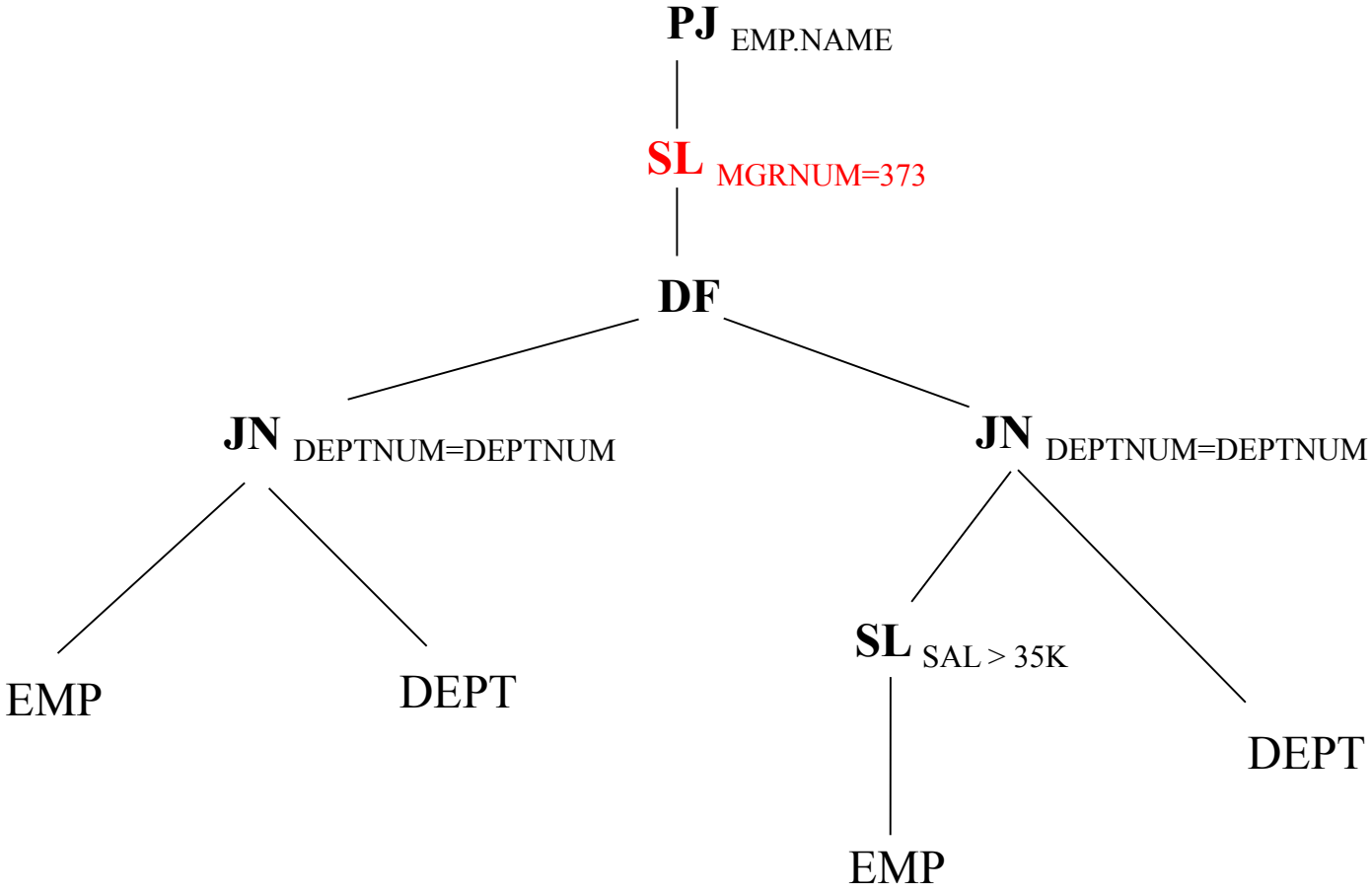
Yes. Both Criteria 1 & 2

Example 2.1

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

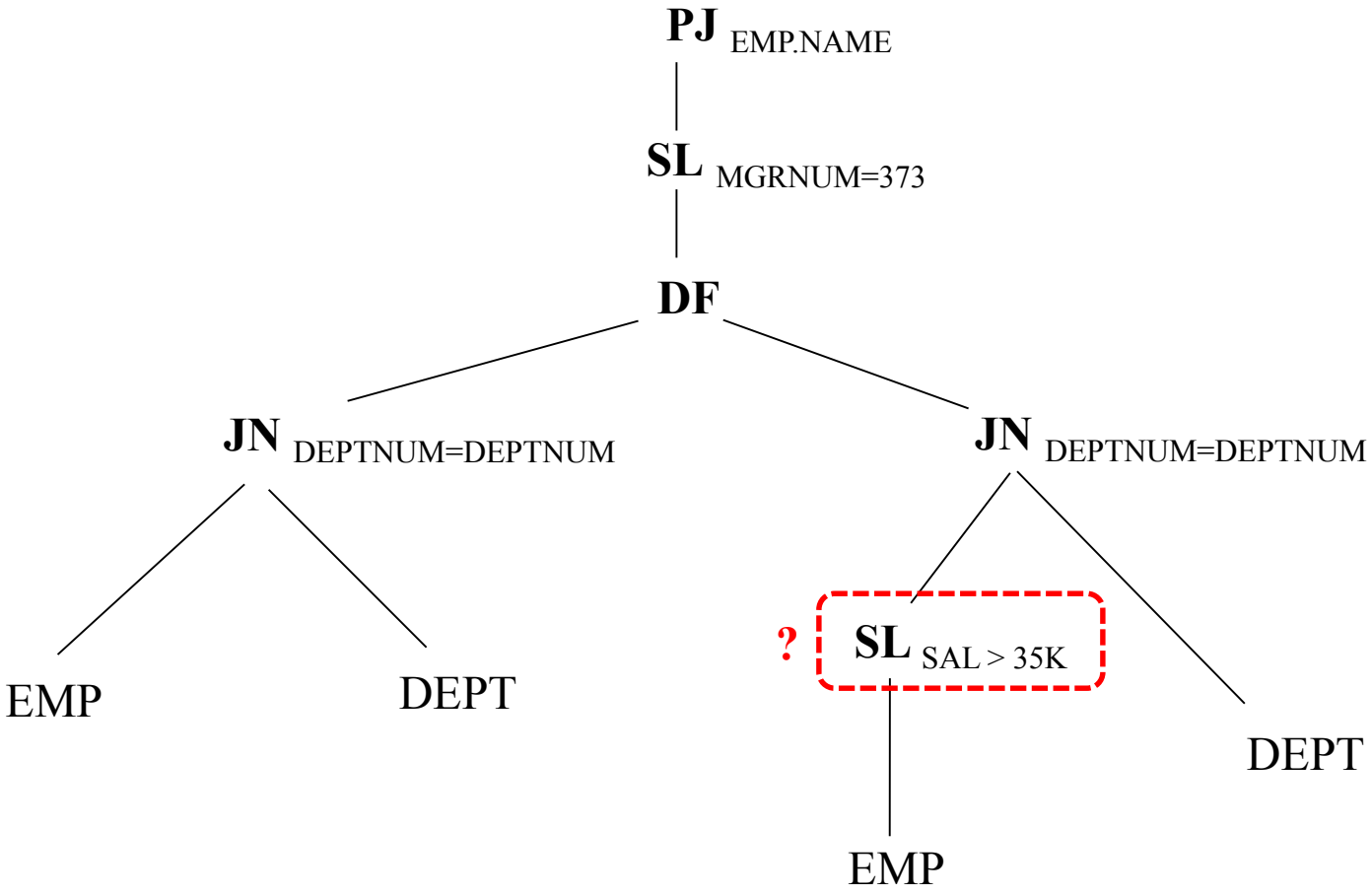
Given query and Operator Tree

Q: PJ $_{\text{EMP.NAME}}$ **SL** $_{\text{MGRNUM=373}}$ ((*EMP* **JN** $_{\text{DEPTNUM=DEPTNUM}}$ *DEPT*) **DF** (**SL** $_{\text{SAL > 35K}}$ *EMP* **JN** $_{\text{DEPTNUM=DEPTNUM}}$ *DEPT*))



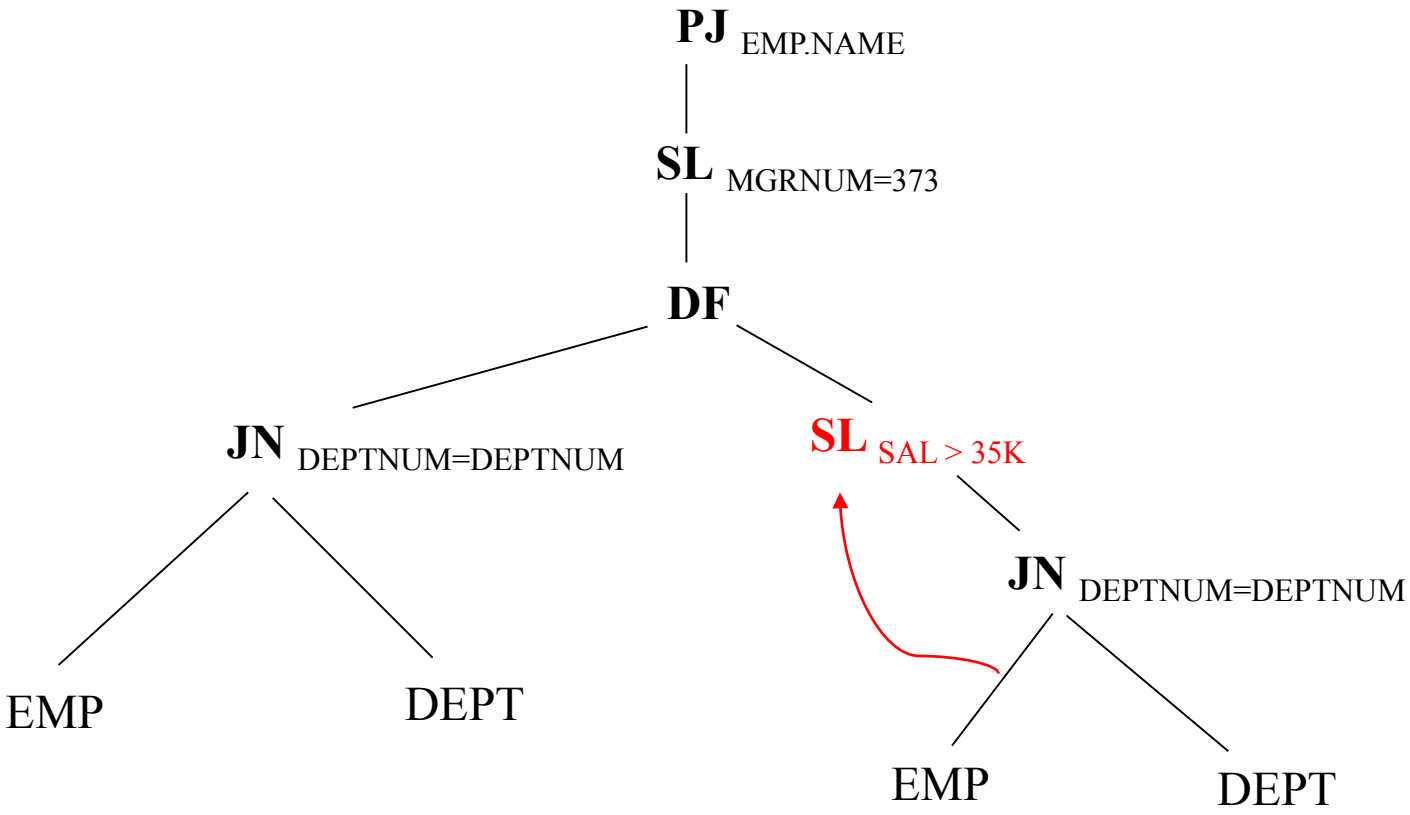
Finding Common Sub-expression

Any common portion?



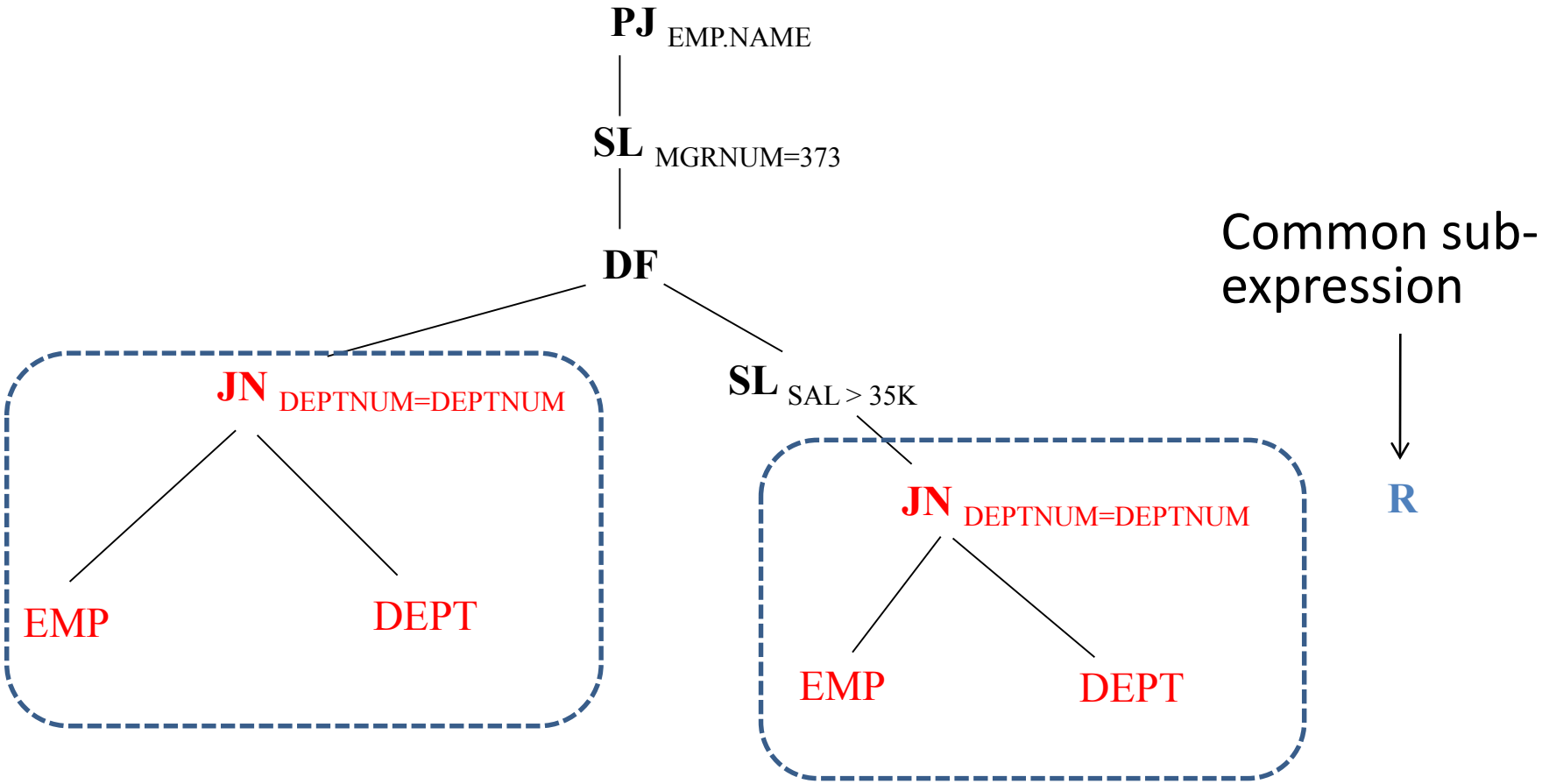
Finding Common Sub-expression

Any common portion? NOW?

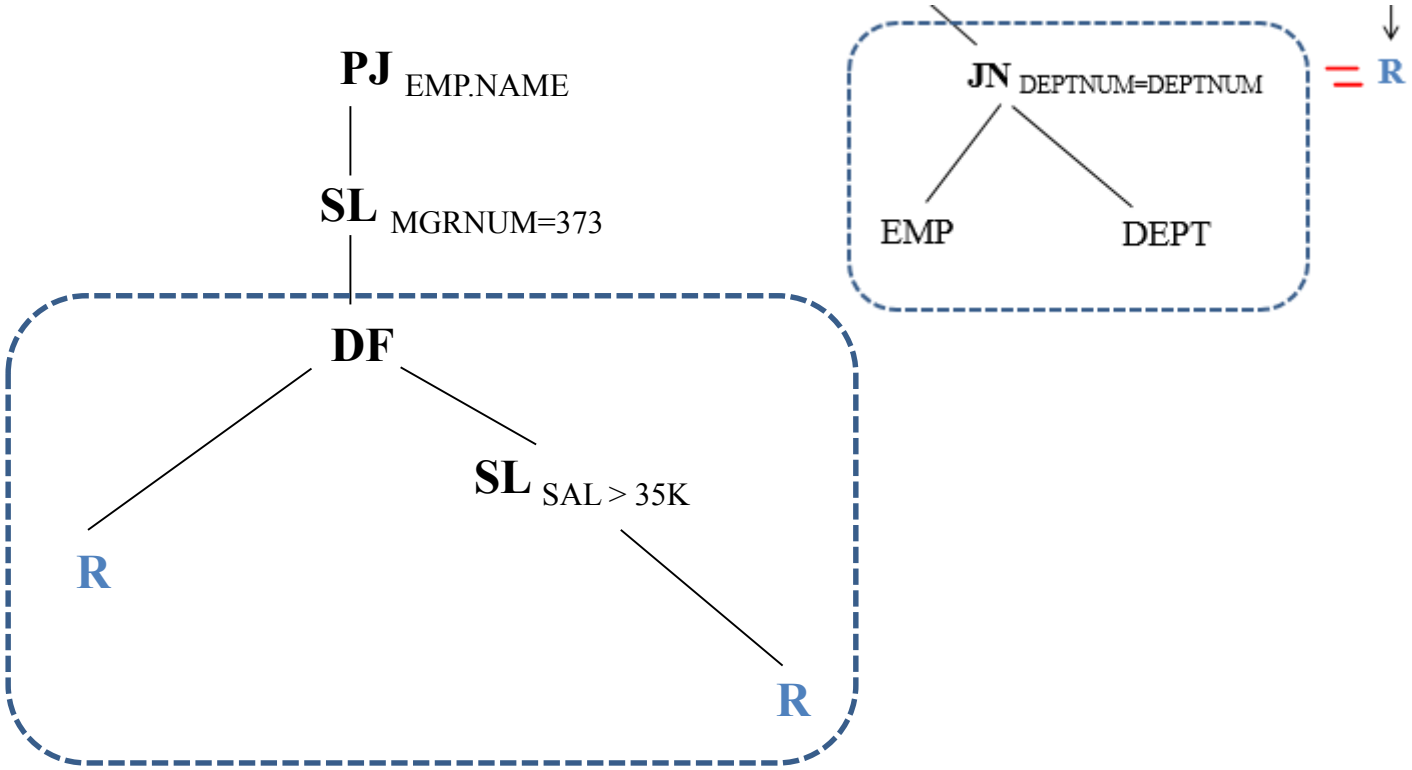


Finding Common Sub-expression

Any common portion? NOW?

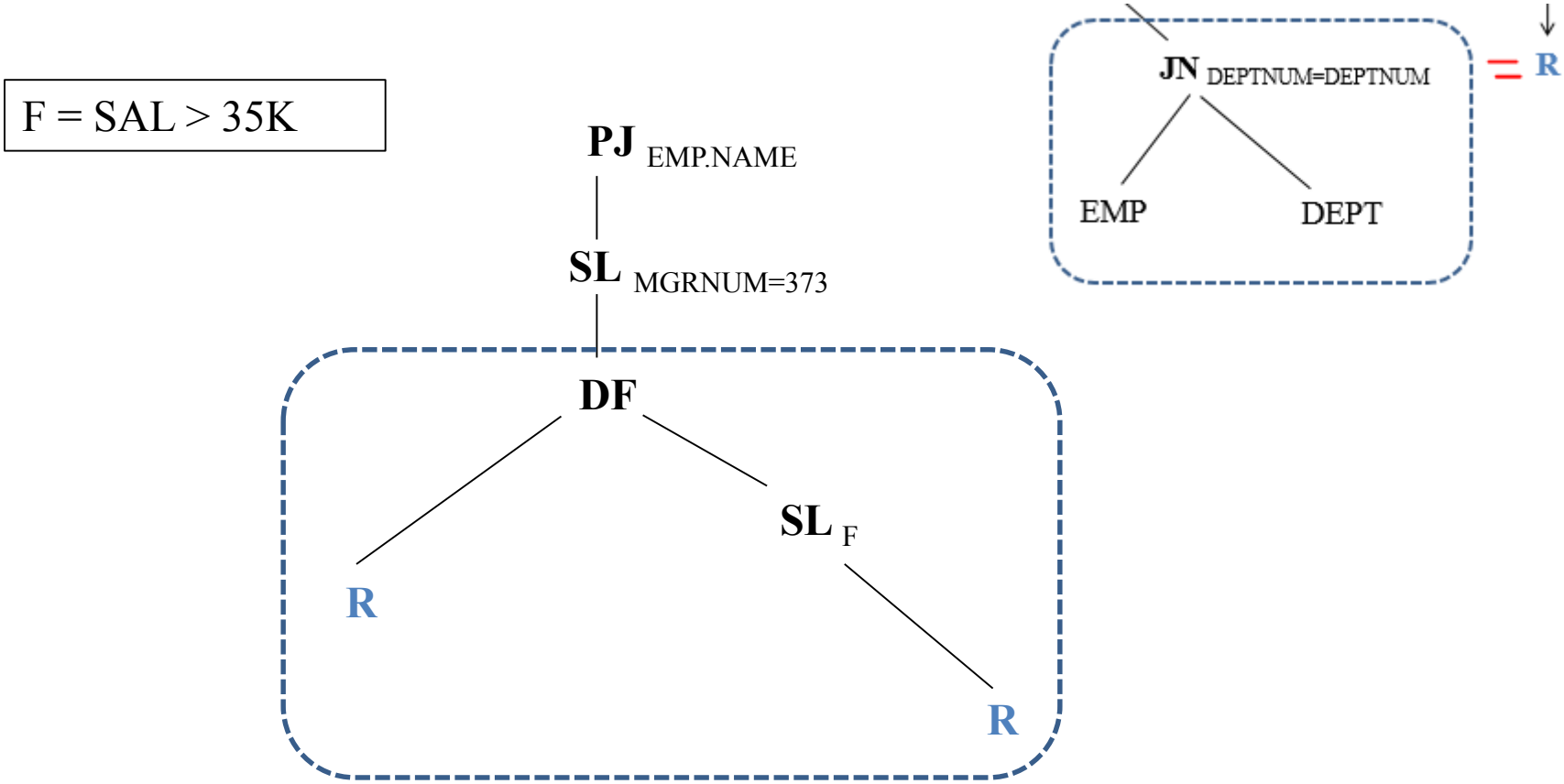


Finding Common Sub-expression



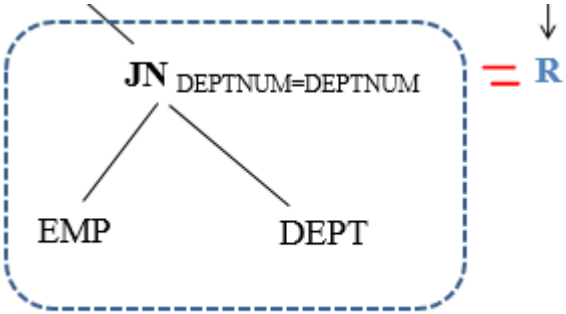
Finding Common Sub-expression

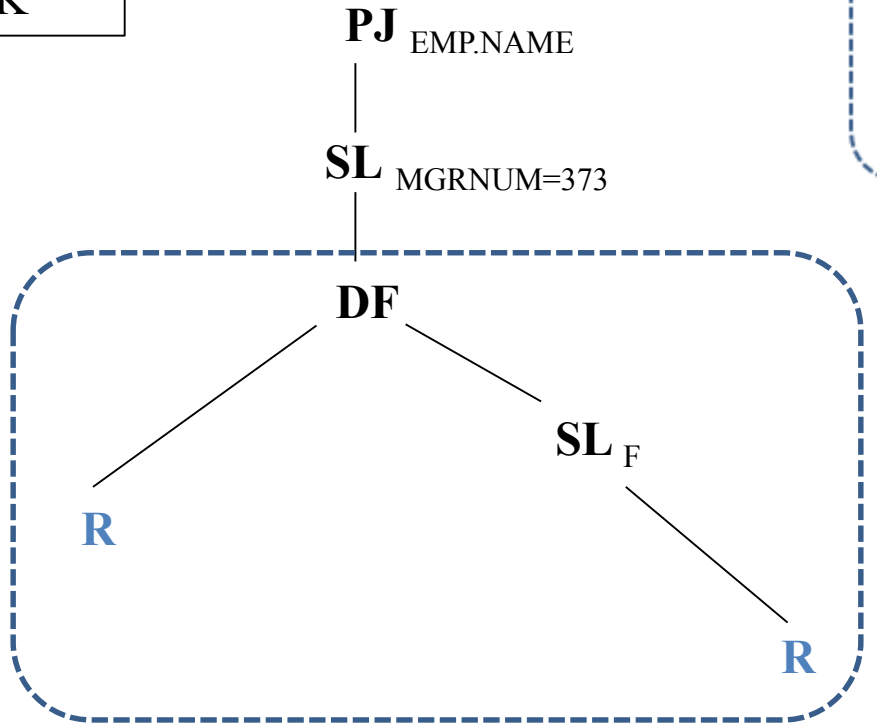
F = SAL > 35K



Finding Common Sub-expression

F = SAL > 35K

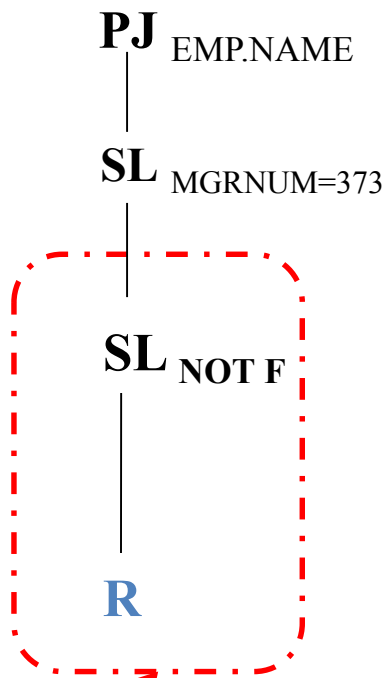


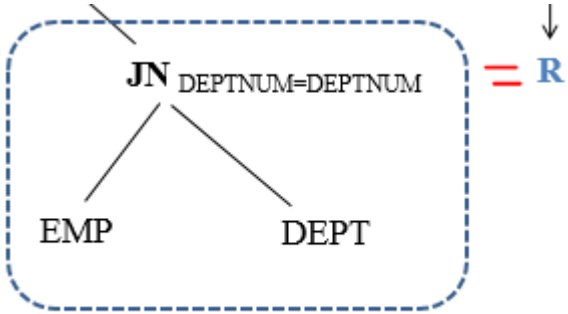


We can write it as --> R DF SL F R which is Rule 6 !

Removing Common Sub-expression

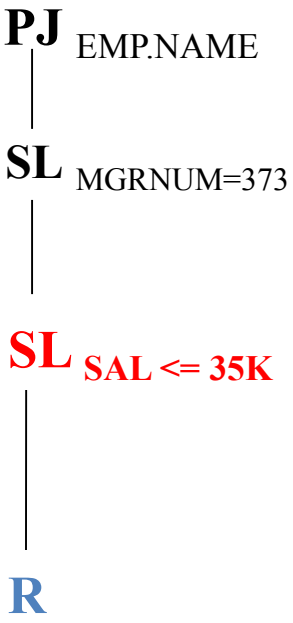
F = SAL > 35K

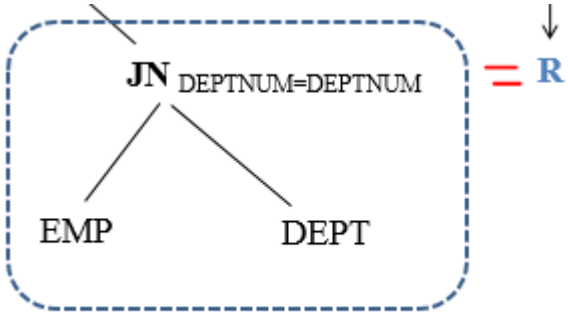




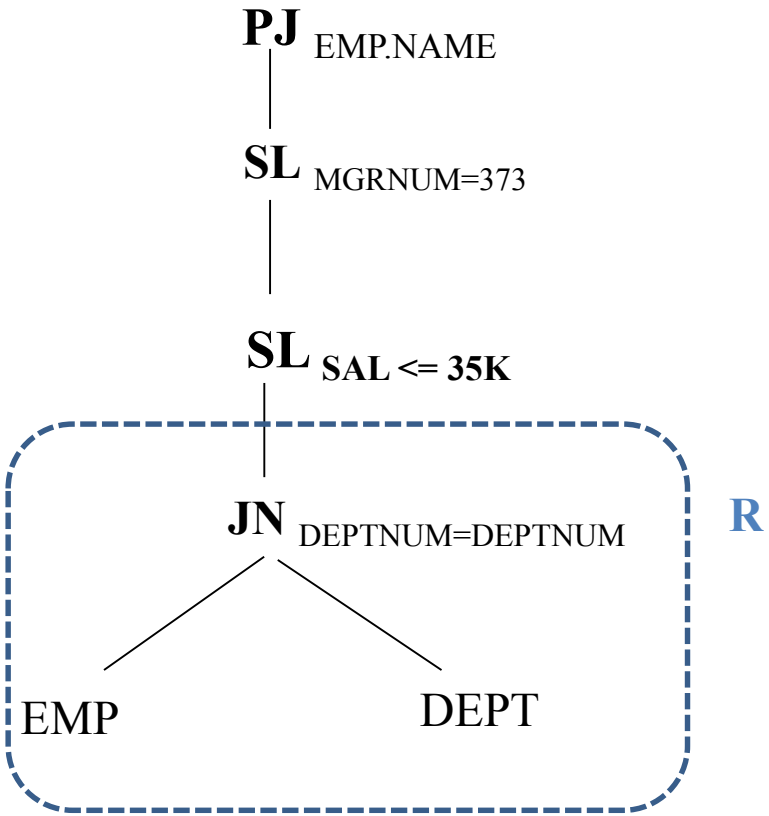
- R DF SL_F R ↔ $SL_{NOT\,F}$ R

Removing Common Sub-expression





Removing Common Sub-expression



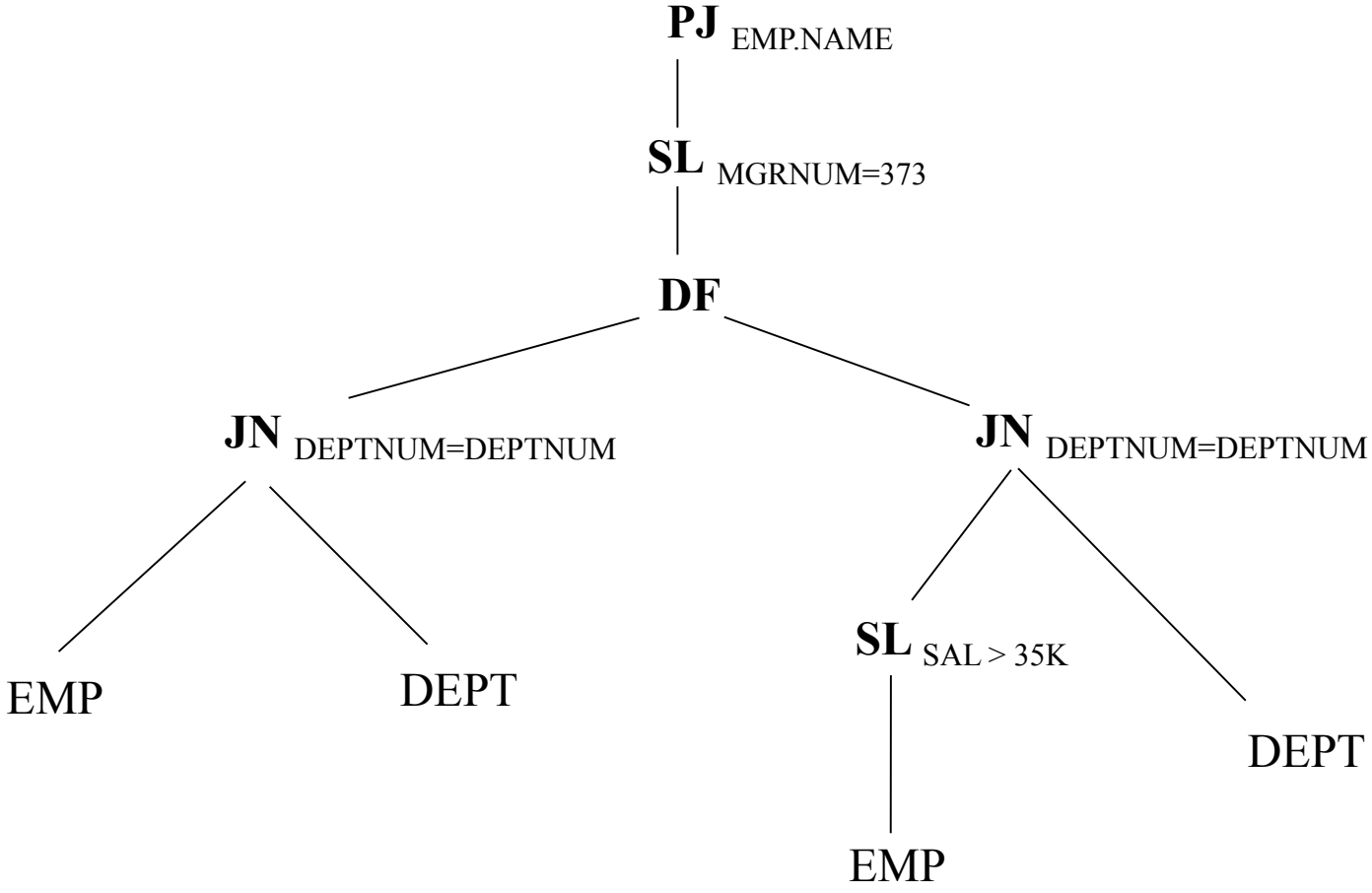
Can you apply Criterion 1 and/or 2 on this tree? Yes. Both Criteria 1 & 2

Example 2.2

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

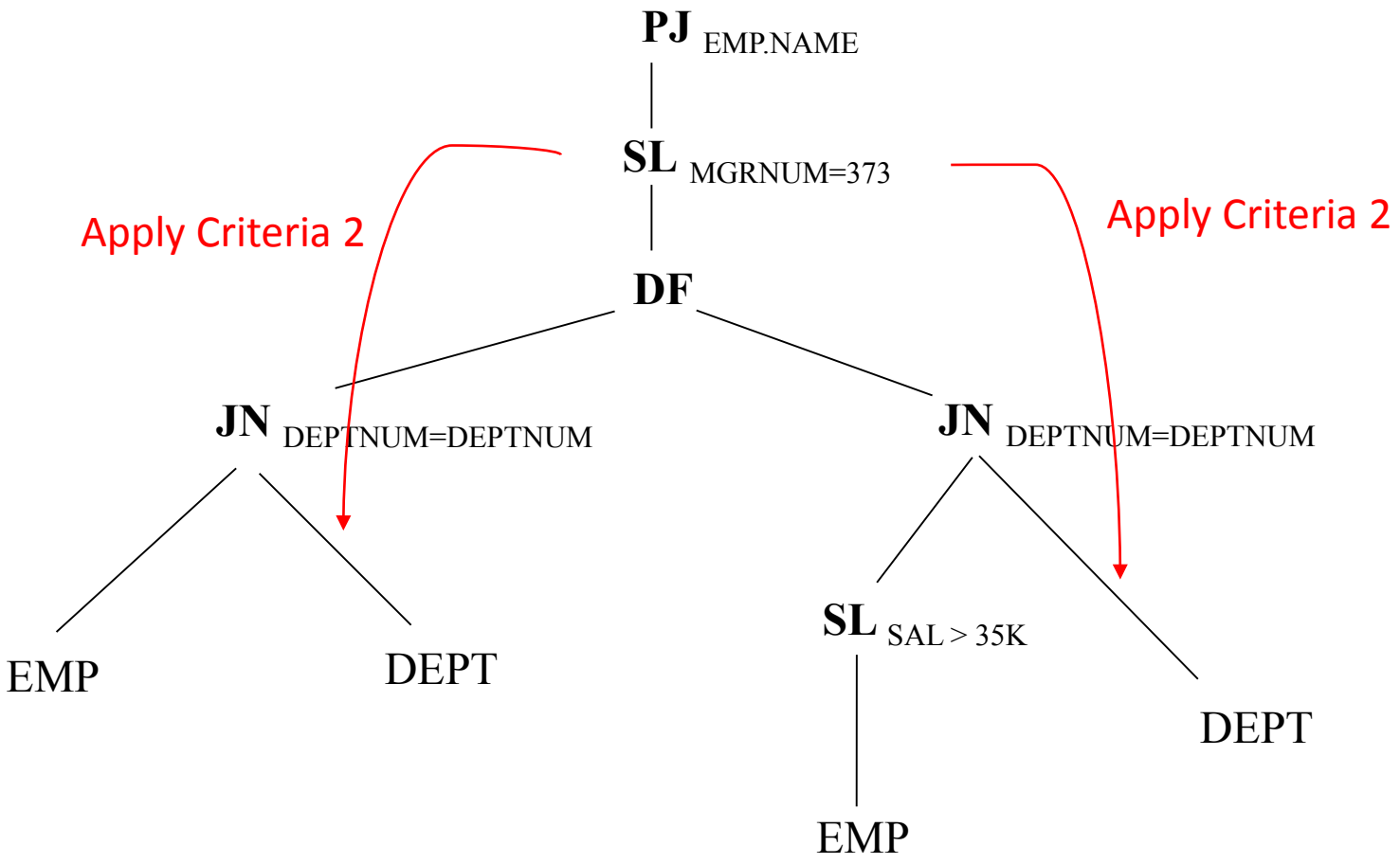
Given query and Operator Tree

Q: PJ $_{\text{EMP.NAME}}$ **SL** $_{\text{MGRNUM=373}}$ ((*EMP* **JN** $_{\text{DEPTNUM=DEPTNUM}}$ *DEPT*) **DF** (**SL** $_{\text{SAL > 35K}}$ *EMP* **JN** $_{\text{DEPTNUM=DEPTNUM}}$ *DEPT*))



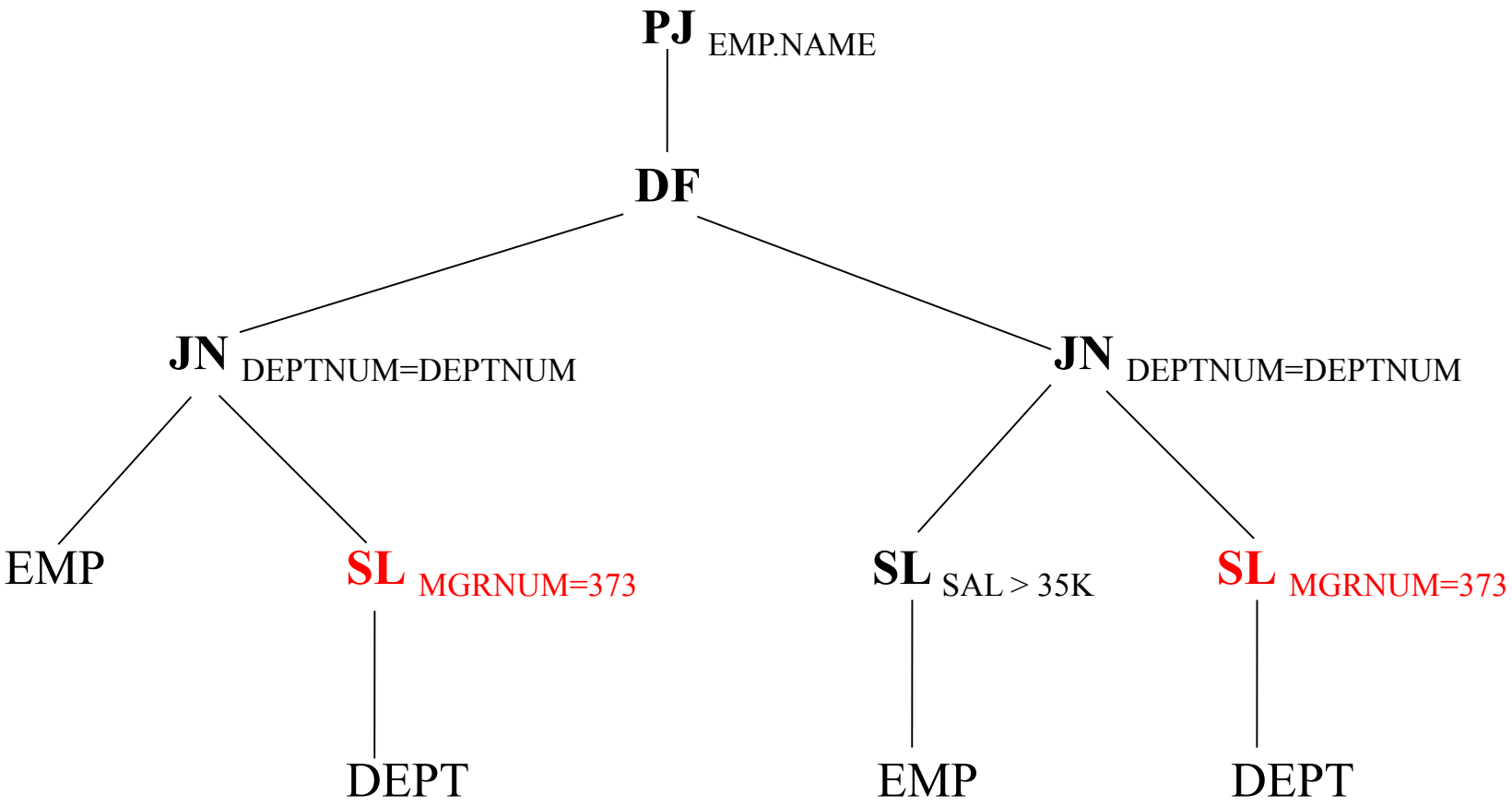
Finding Common Sub-expression

Any common portion?



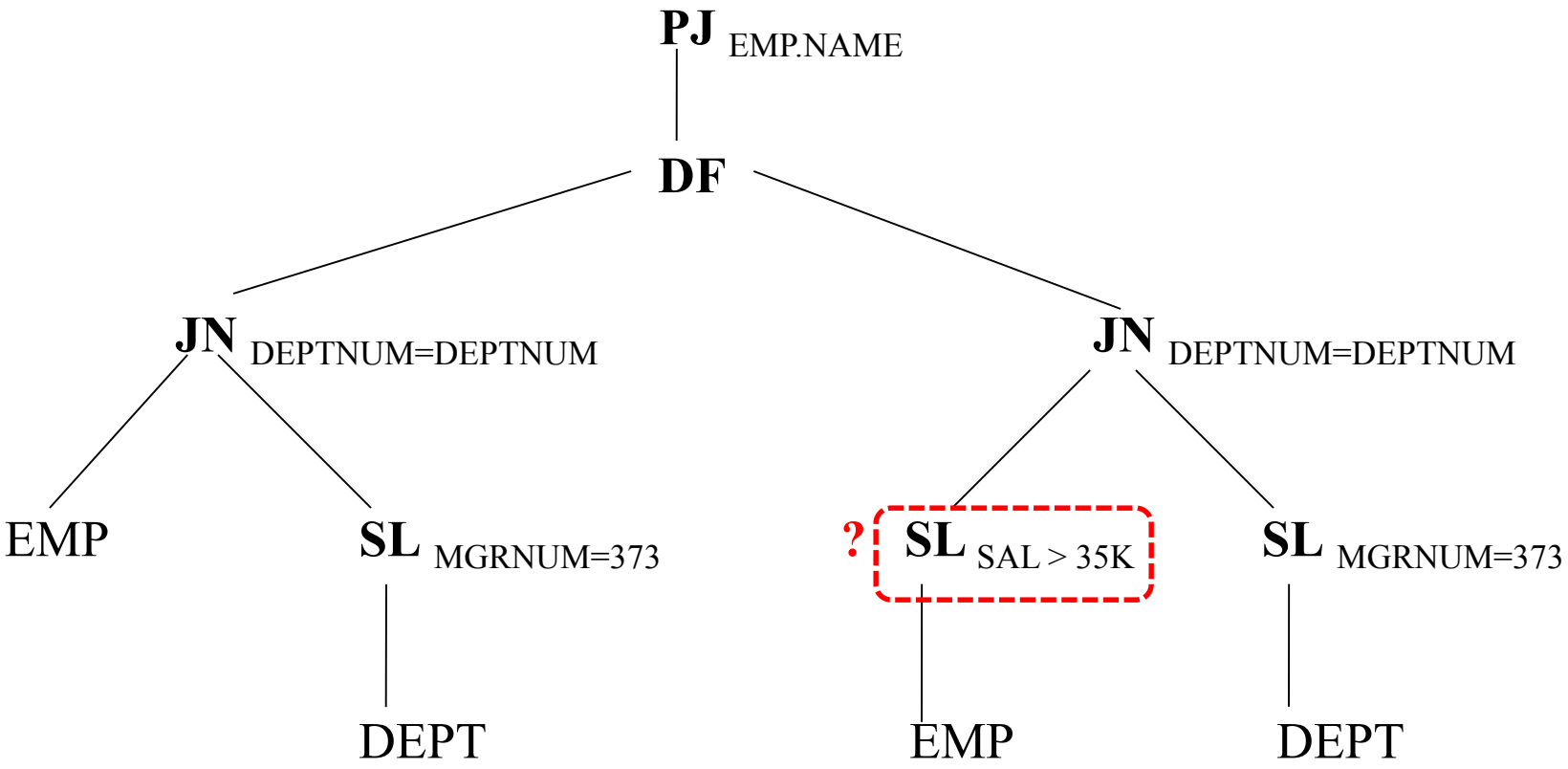
Finding Common Sub-expression

Any common portion?



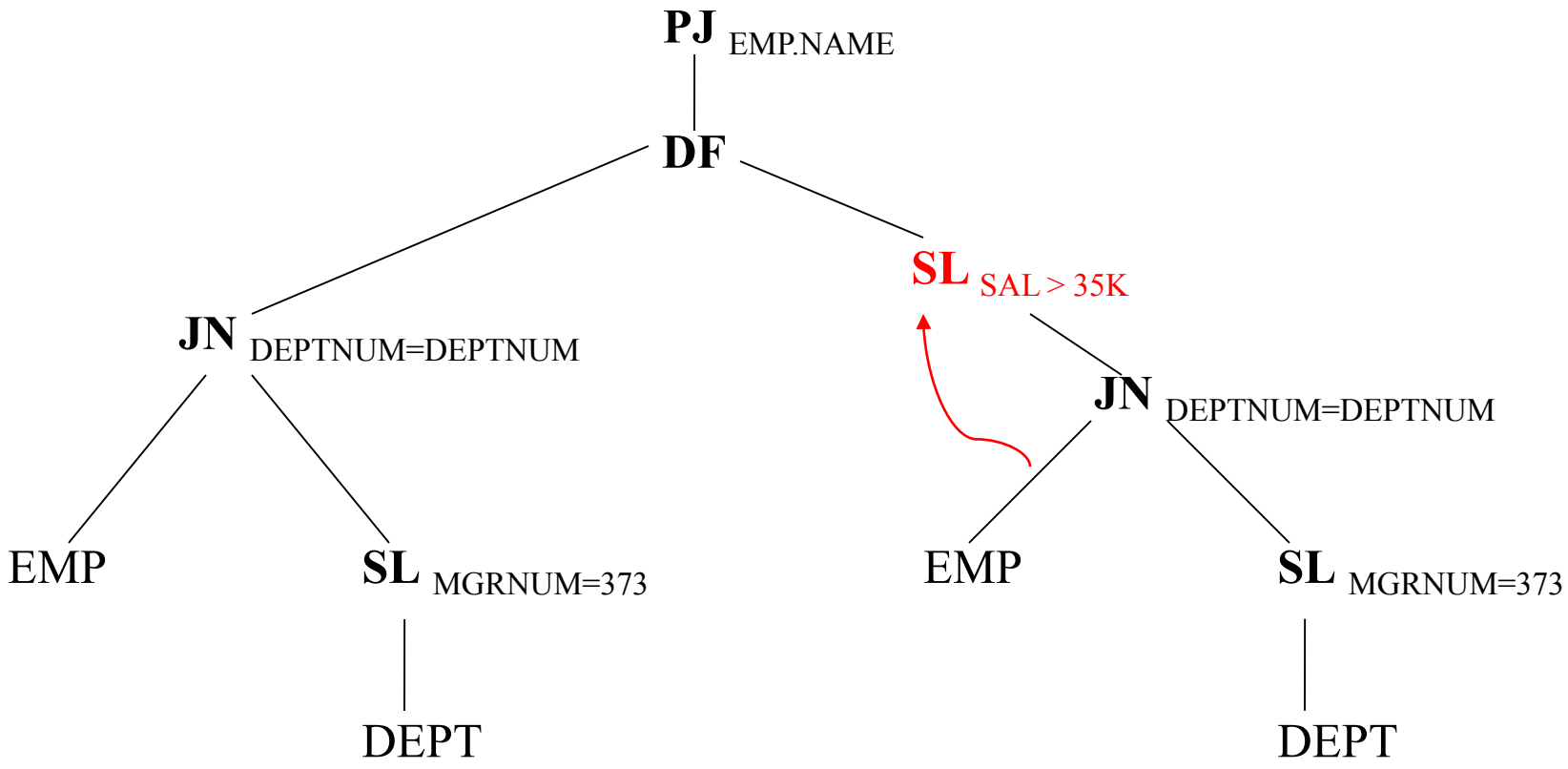
Finding Common Sub-expression

Any common portion?

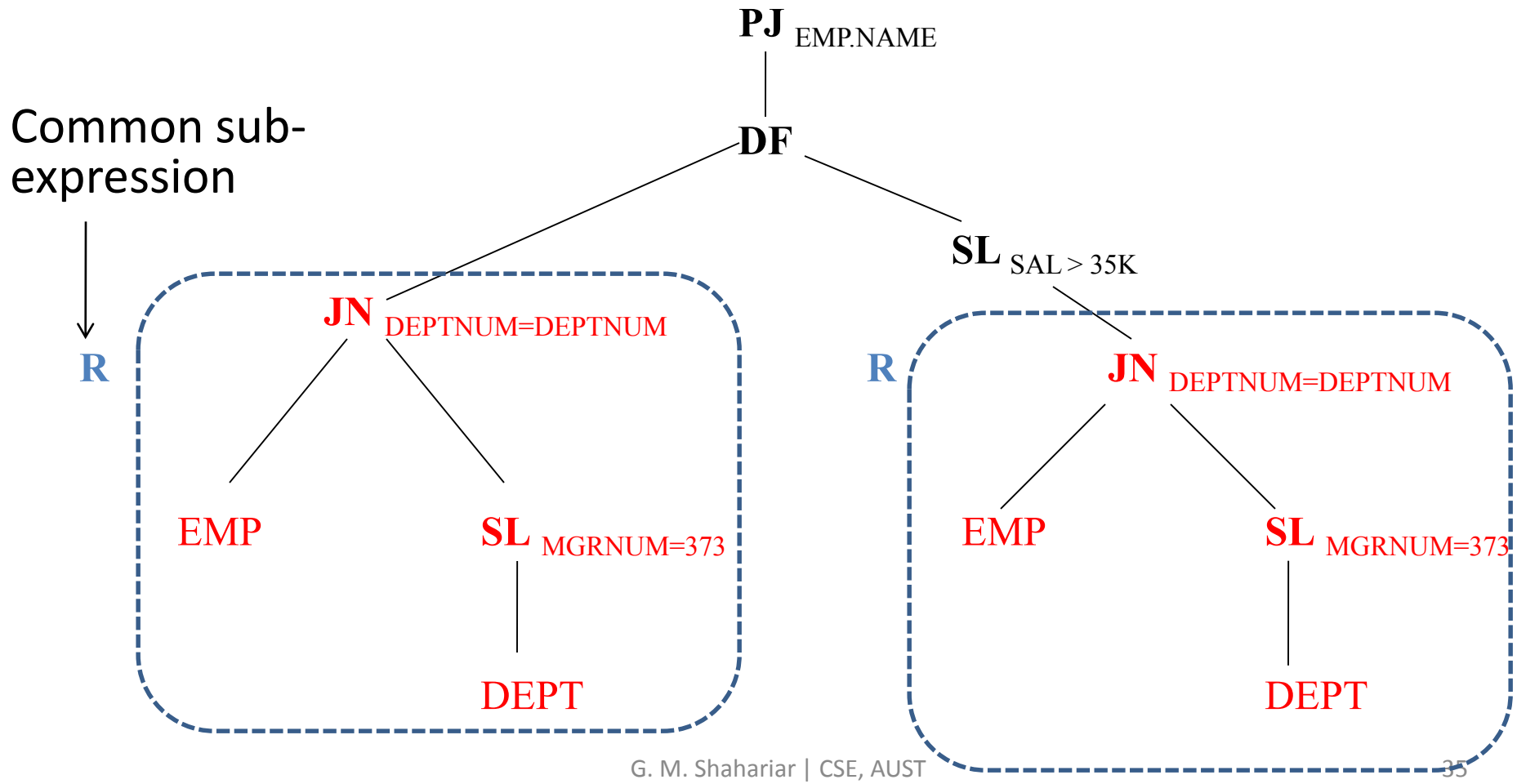


Finding Common Sub-expression

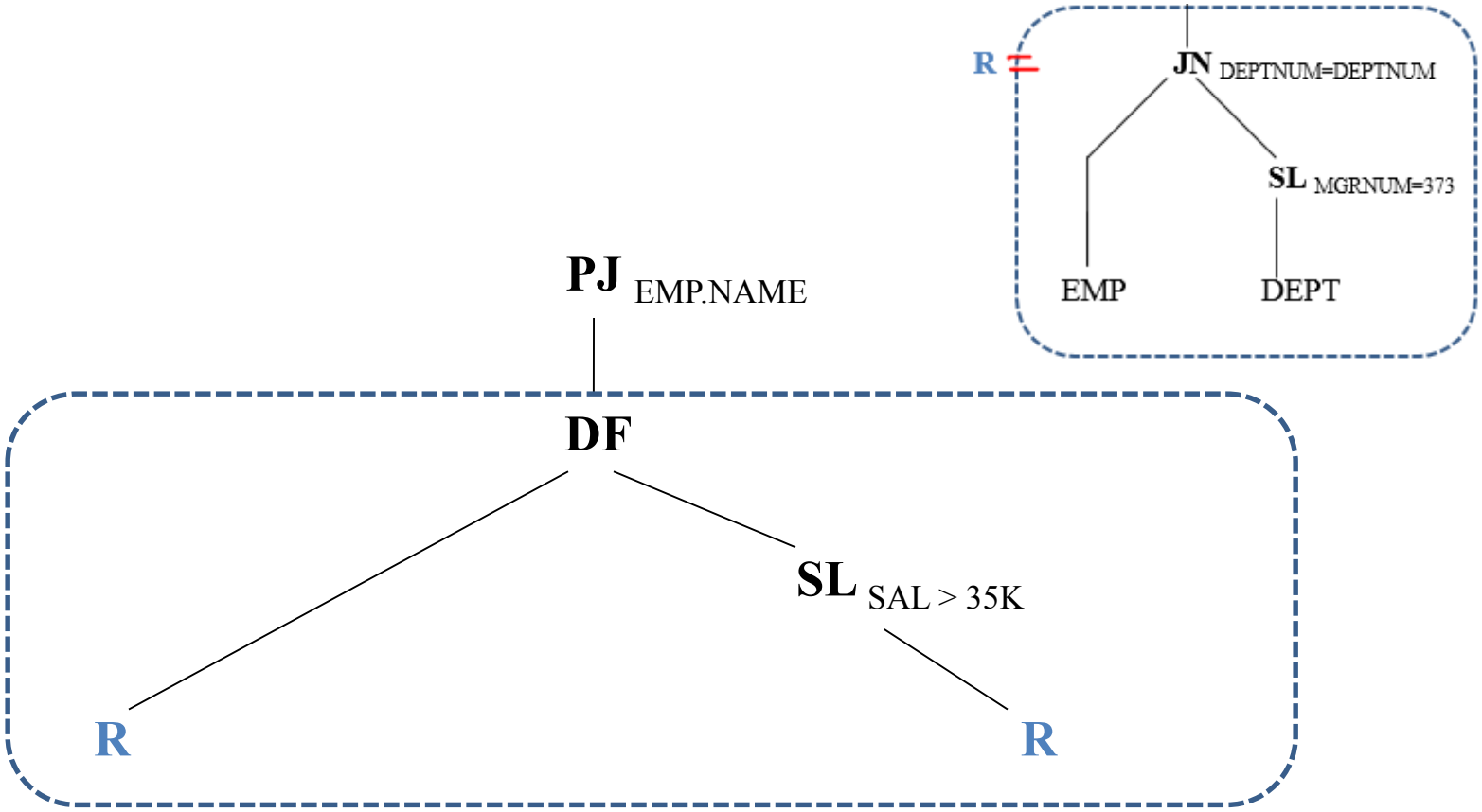
Any common portion? NOW?



Finding Common Sub-expression

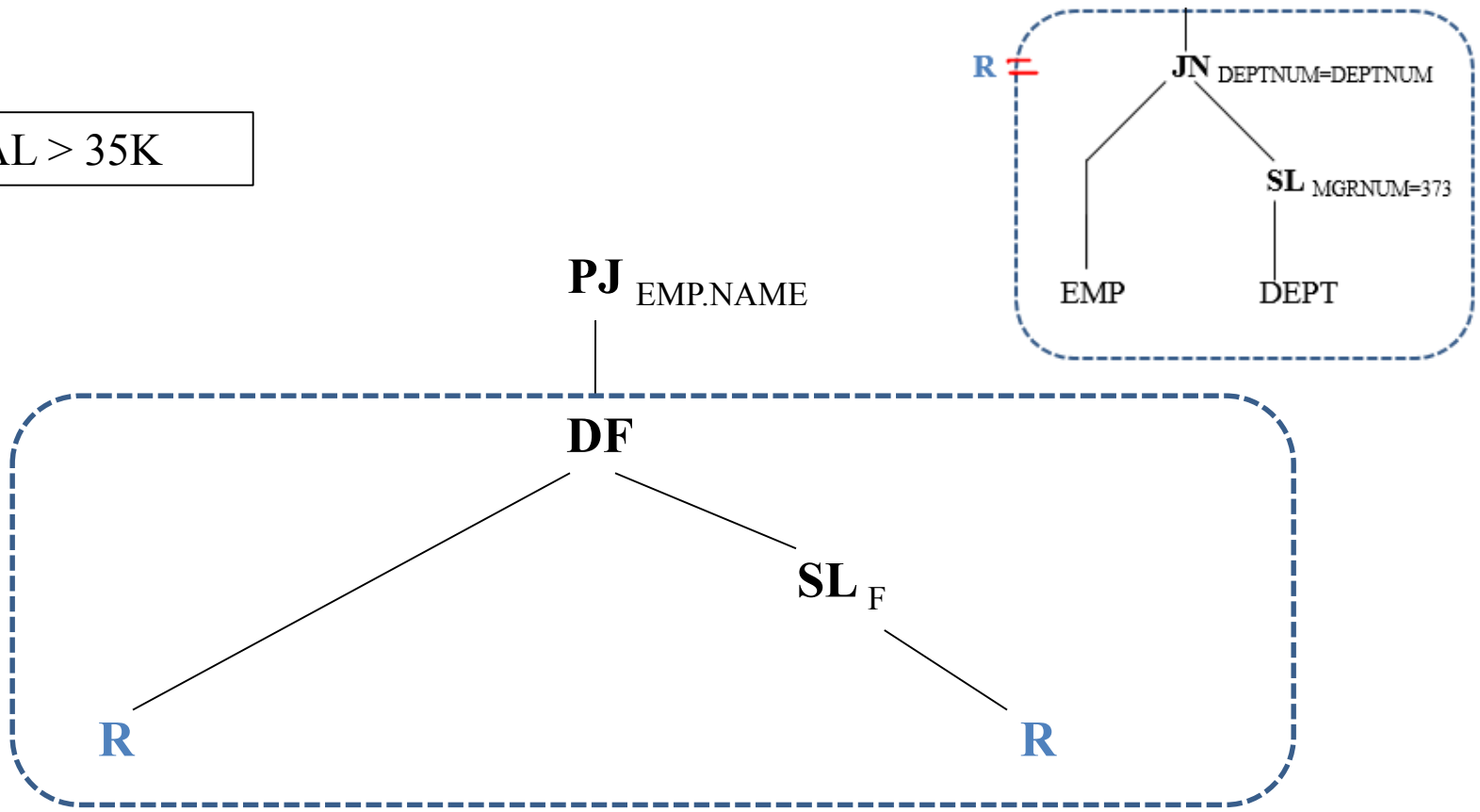


Finding Common Sub-expression



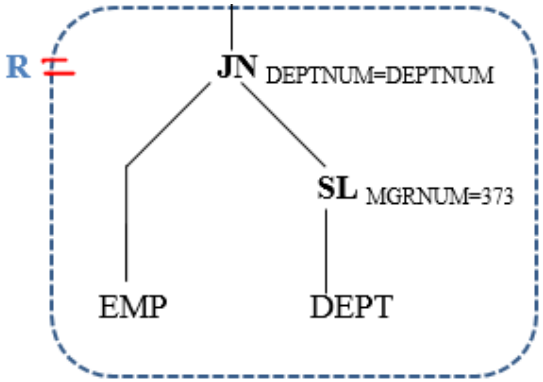
Finding Common Sub-expression

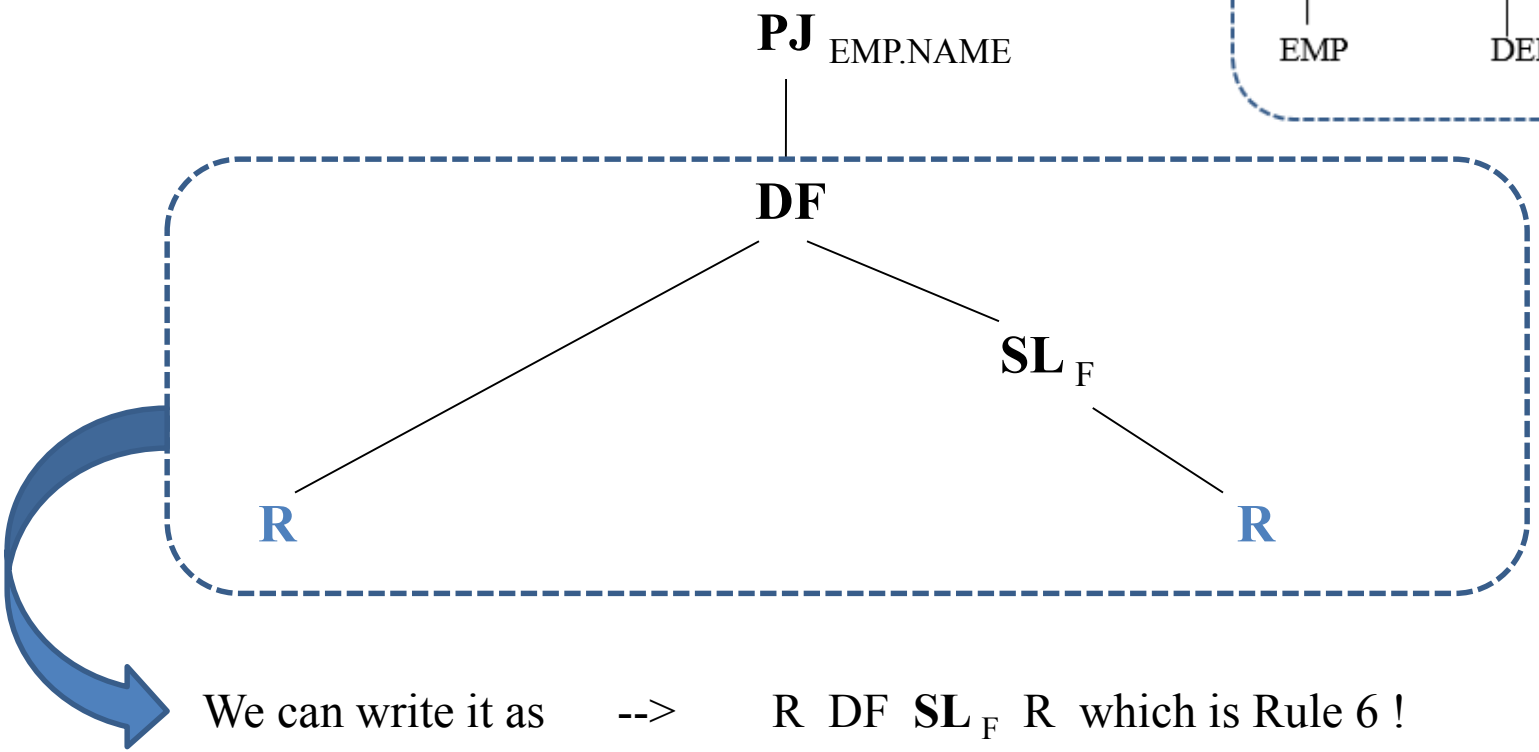
F = SAL > 35K



Finding Common Sub-expression

F = SAL > 35K

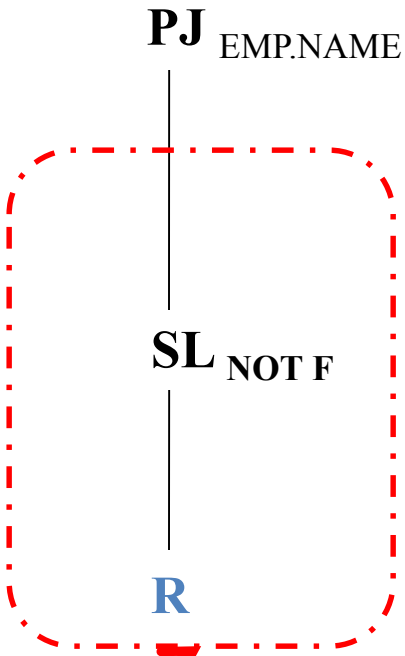


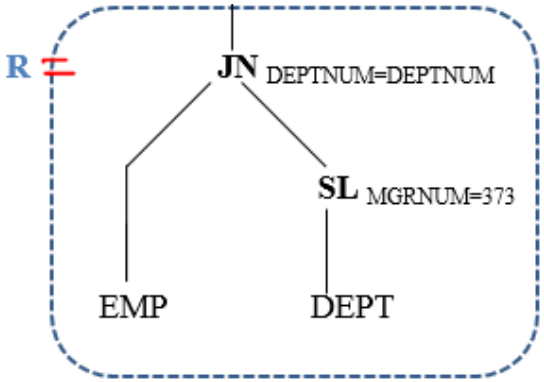


We can write it as --> R DF SL_F R which is Rule 6 !

Removing Common Sub-expression

F = SAL > 35K

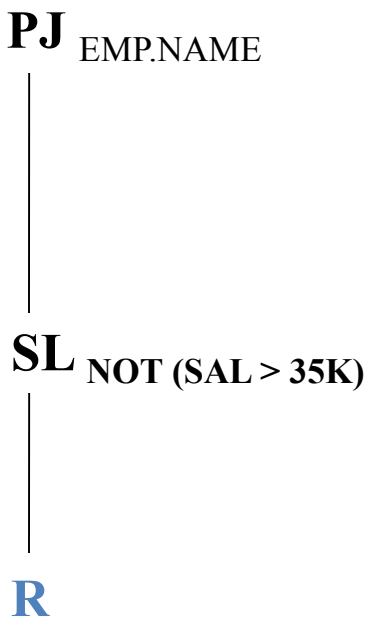


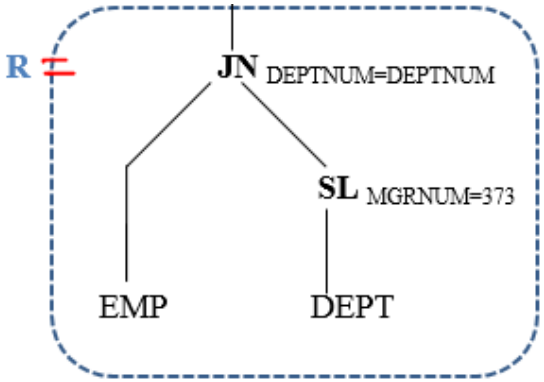


- R DF SL_F R ↔ $SL_{NOT\,F}$ R

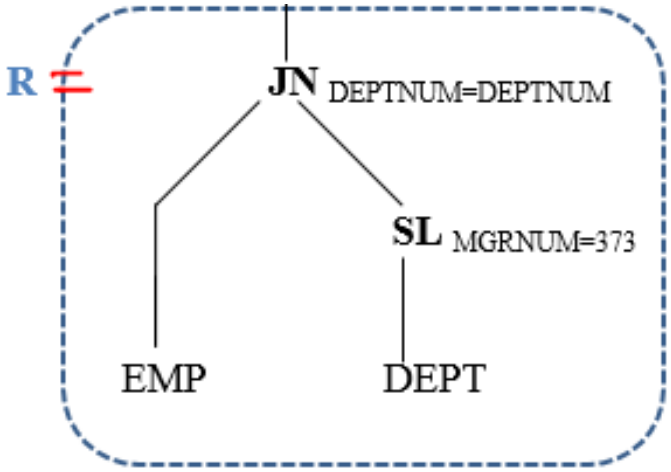
Removing Common Sub-expression

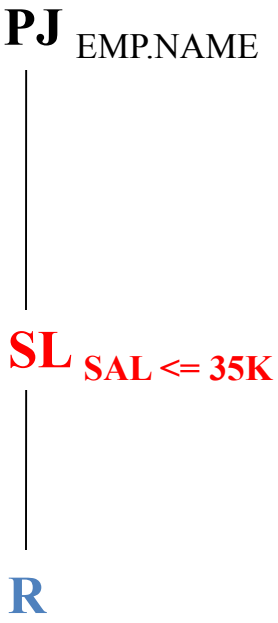
F = SAL > 35K



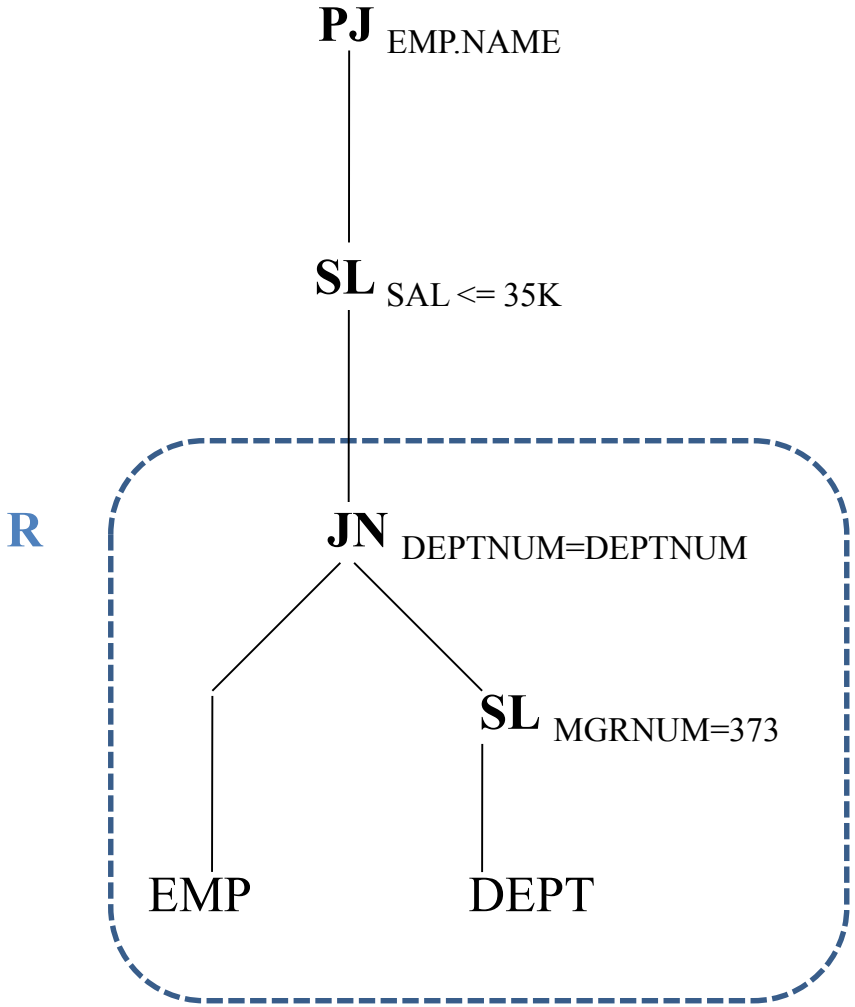


Removing Common Sub-expression





Removing Common Sub-expression



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

Example 3

Practise

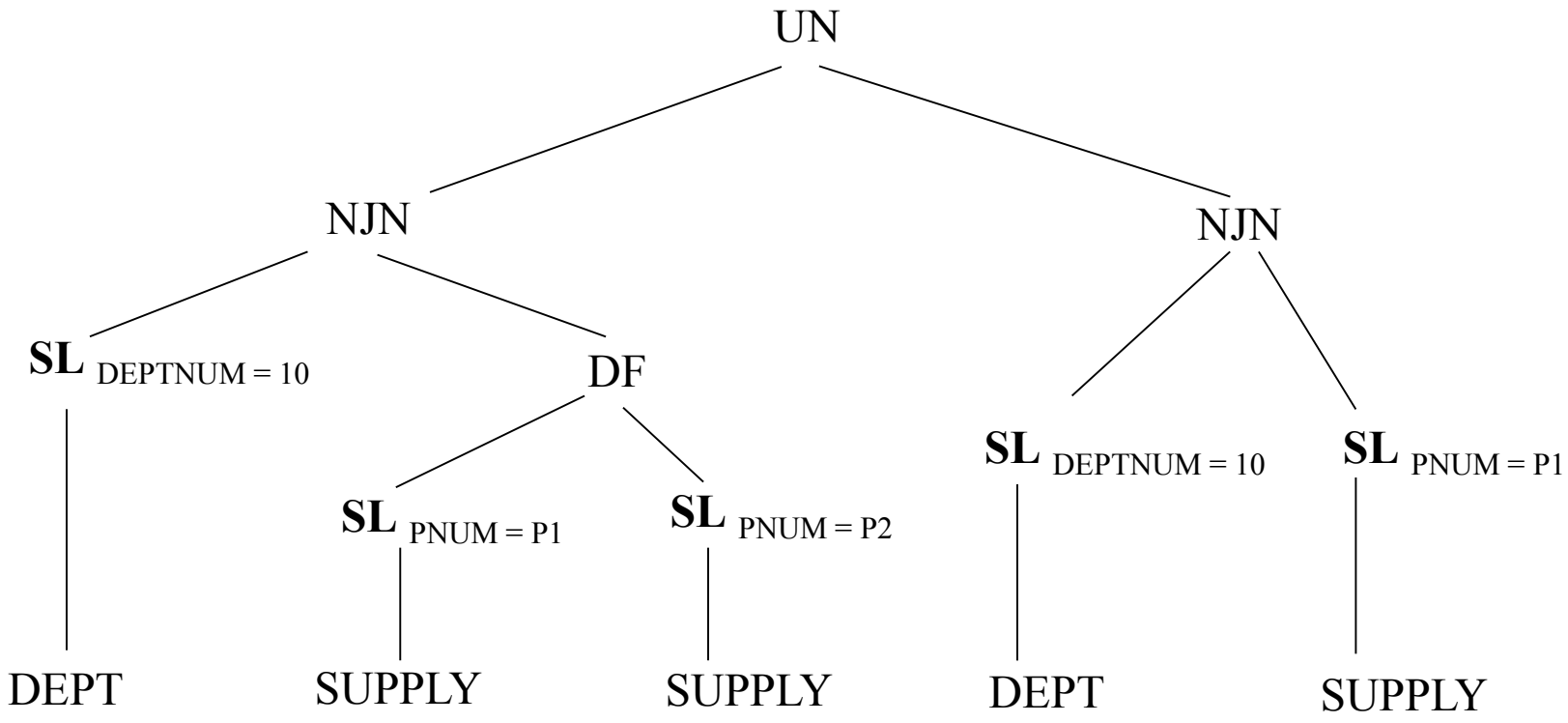
* Draw Operator Tree for the following queries:

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

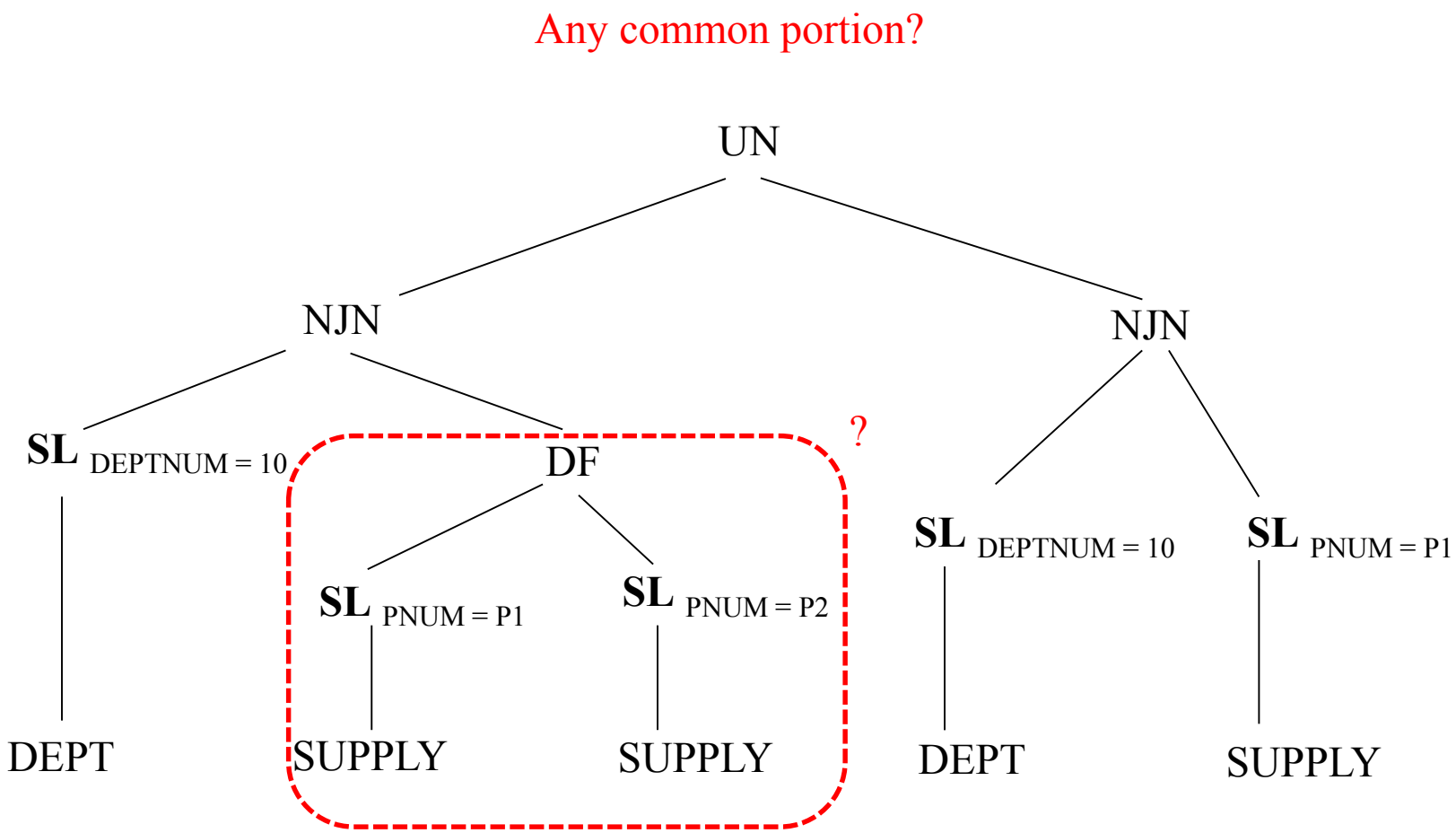
Query:

$(SL_{DEPTNUM=10}$ DEPT NJN $(SL_{PNUM="P1"}$ SUPPLY DF $SL_{PNUM="P_2"}$ SUPPLY)) UN $(SL_{DEPTNUM=10}$ DEPT NJN $SL_{PNUM="P_1"}$ SUPPLY)

Operator Tree

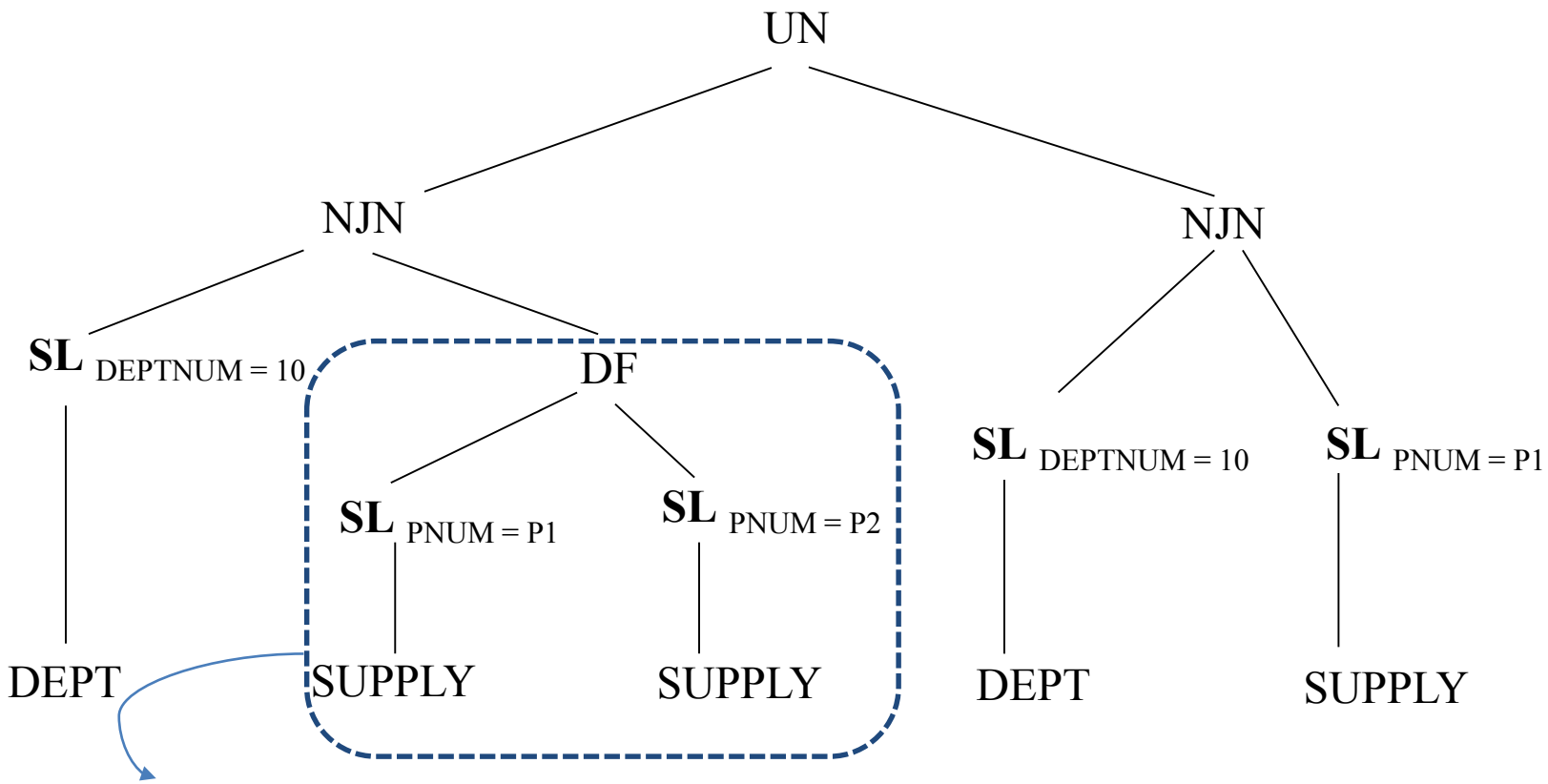


Finding Common Sub - Expression



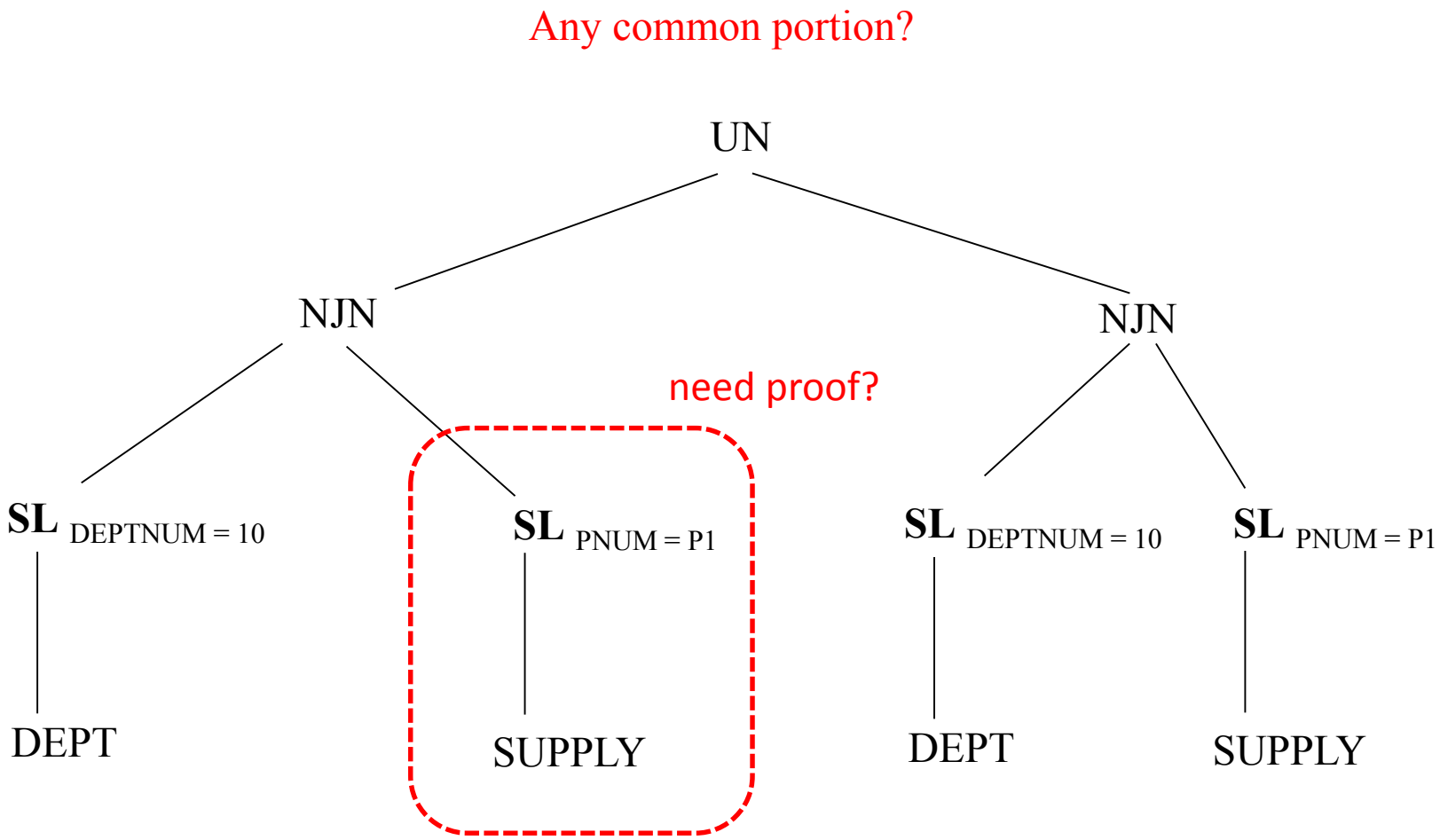
Finding Common Sub - Expression

Any common portion?



We can use Rule 9. $SL_{F1}\ R\ \ DF\ \ SL_{F2}\ R\ \ =\ \ SL_{F1\ AND\ NOT\ F2}\ R$

Finding Common Sub - Expression



Need Proof?

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 $_{PNUM = P1}$ SUPPLY

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

B

SL $_{PNUM = P2}$ 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

-

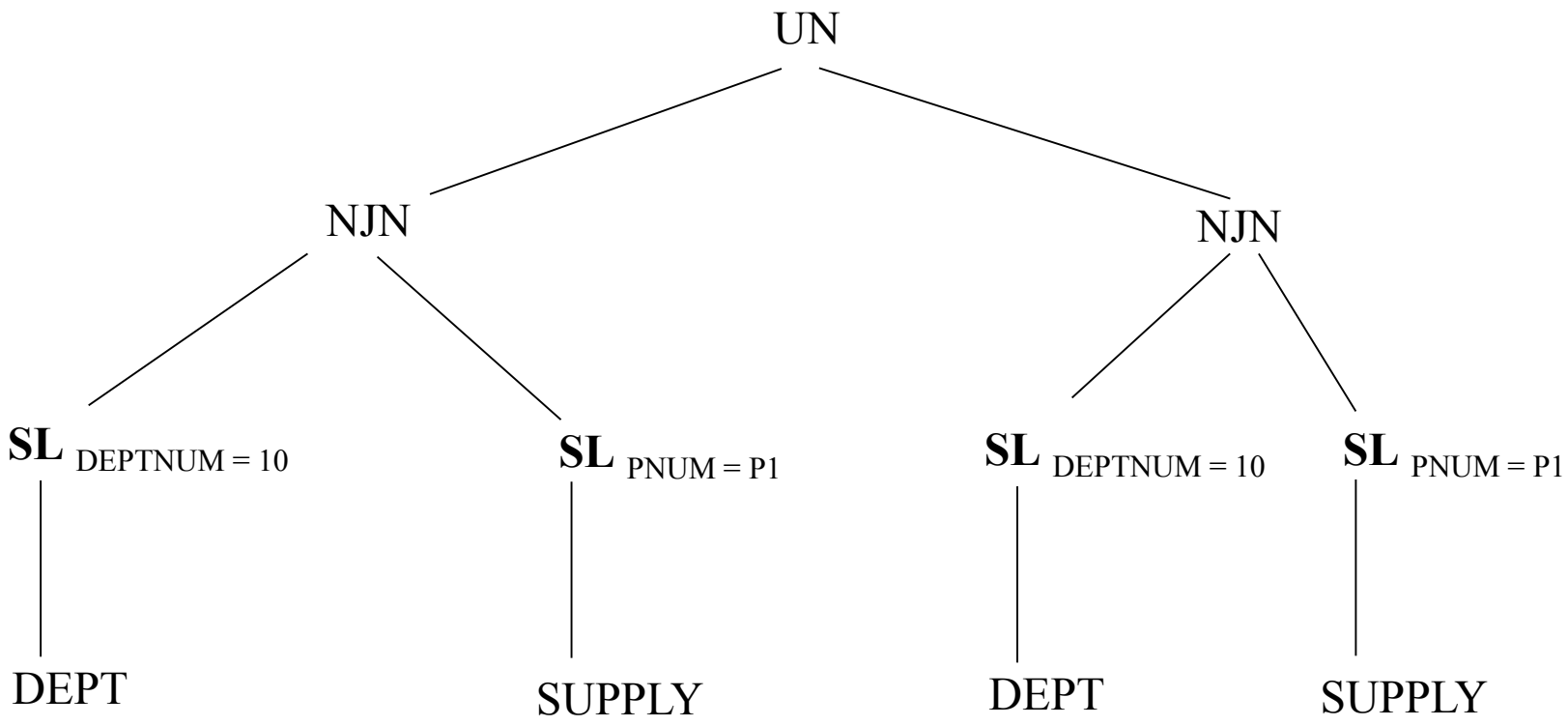
SNUM	PNUM	DEPT NUM	QUAN
2	P2	2	20
4	P2	1	40
6	P2	1	60

=

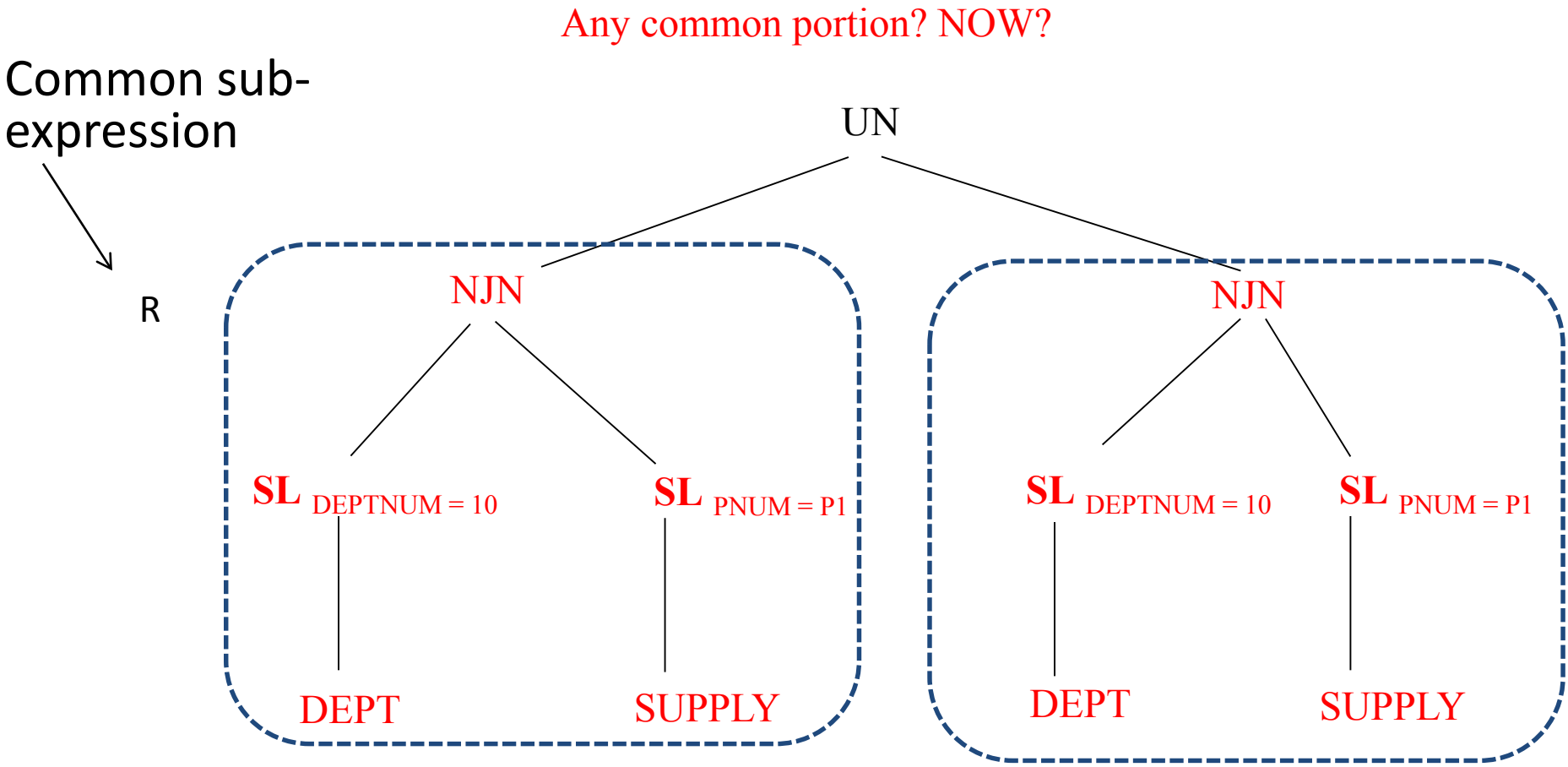
SNUM	PNUM	DEPT NUM	QUAN
1	P1	1	10
3	P1	1	30
5	P1	2	50

Finding Common Sub - Expression

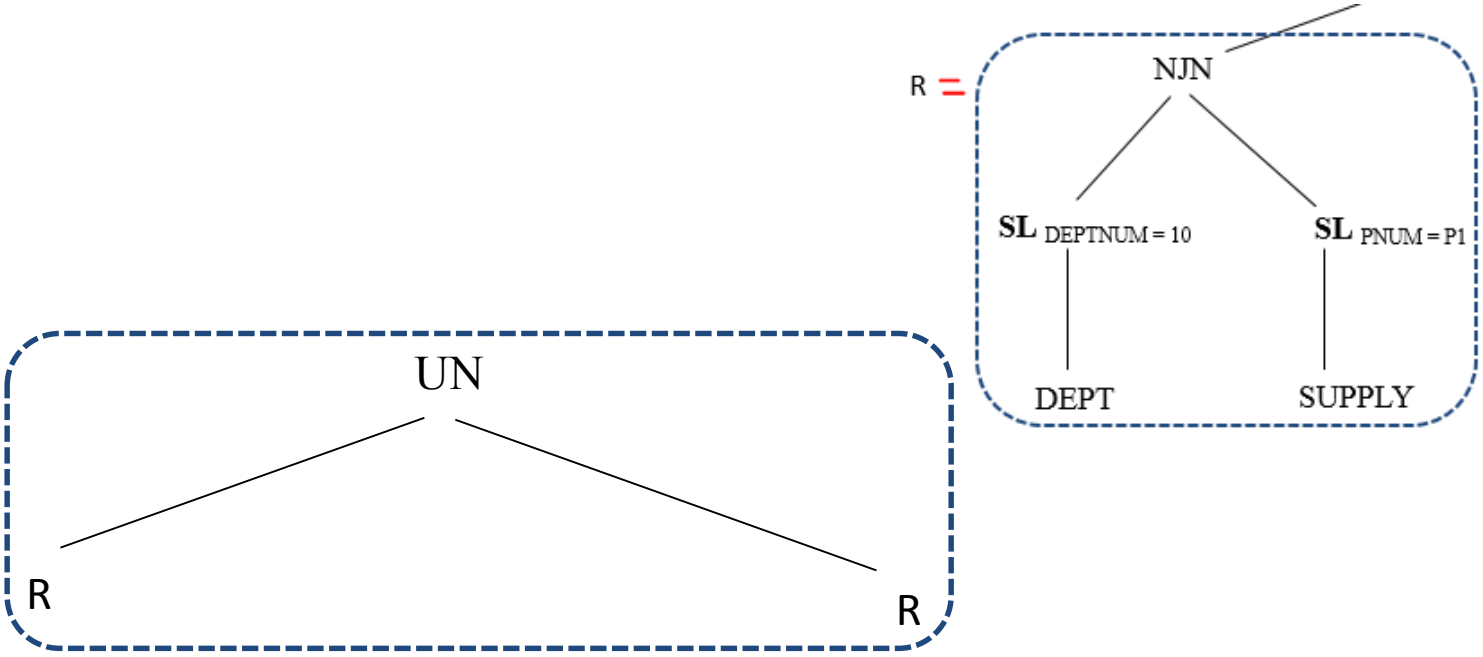
Any common portion?



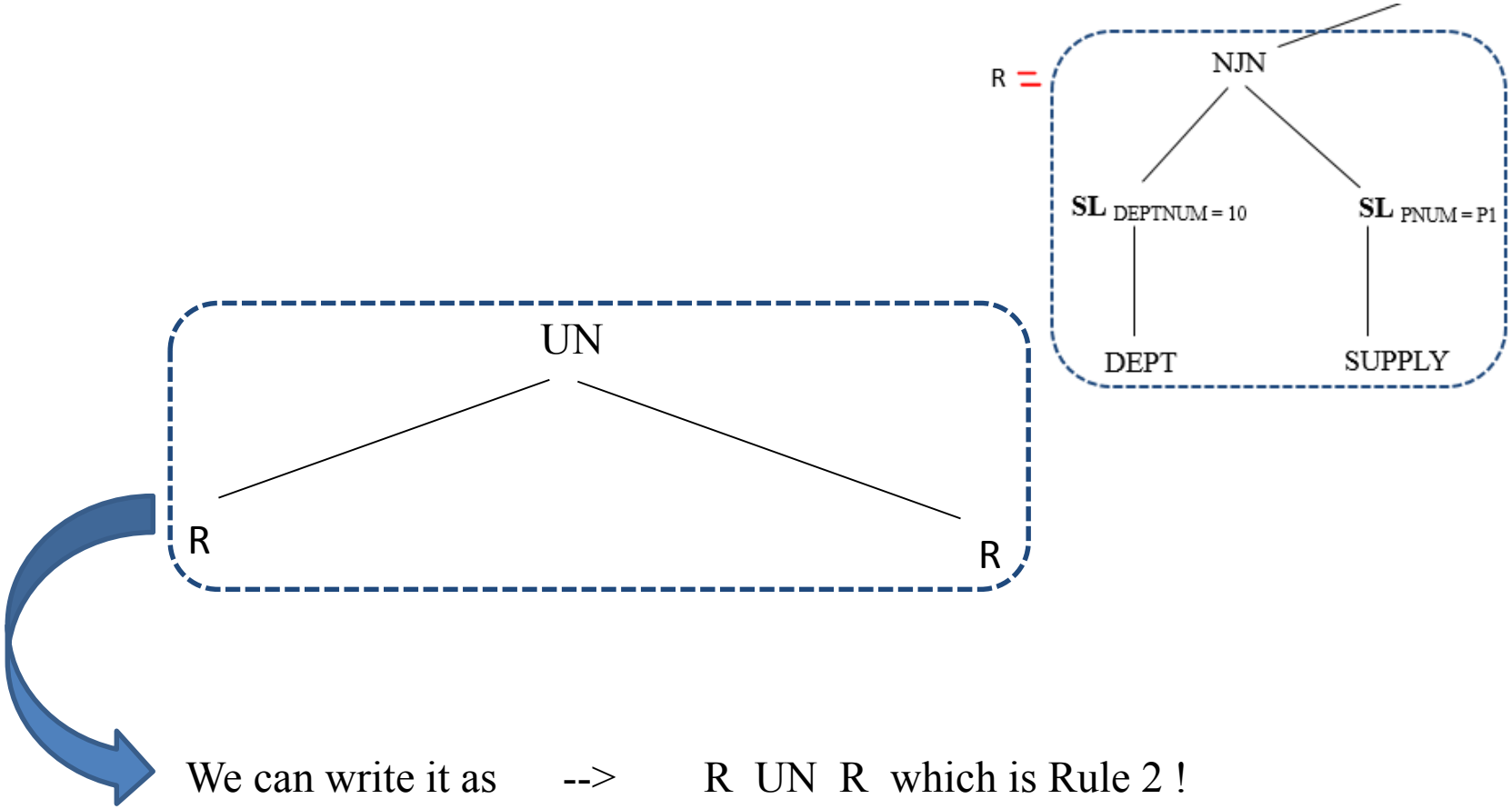
Finding Common Sub - Expression



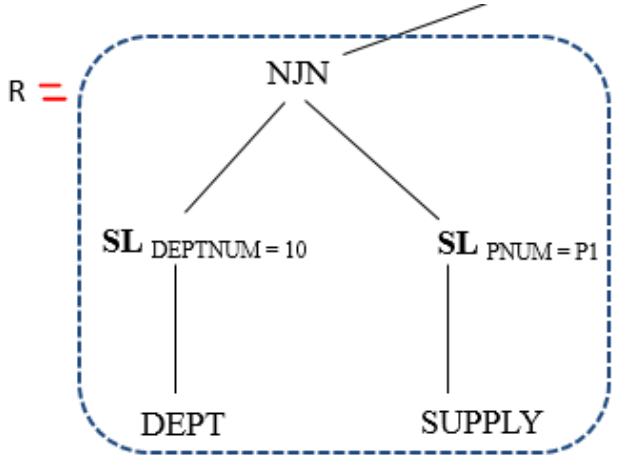
Finding Common Sub - Expression

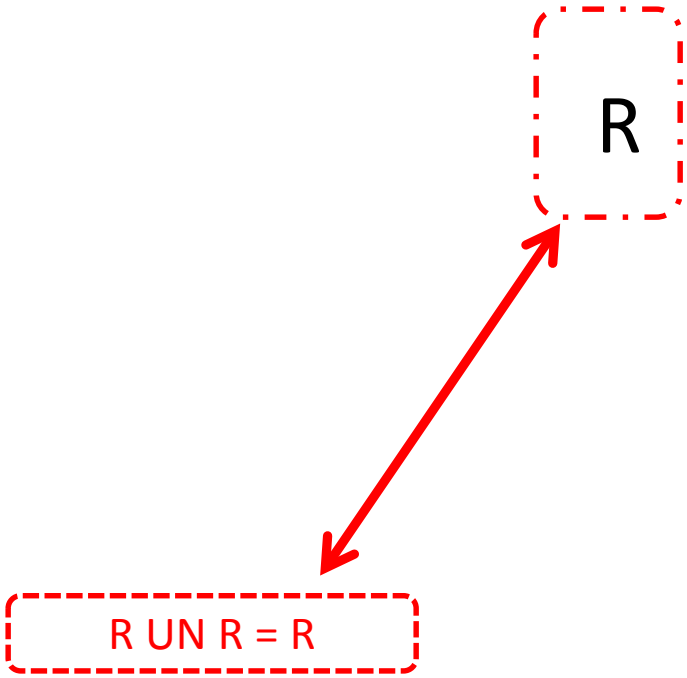


Finding Common Sub - Expression

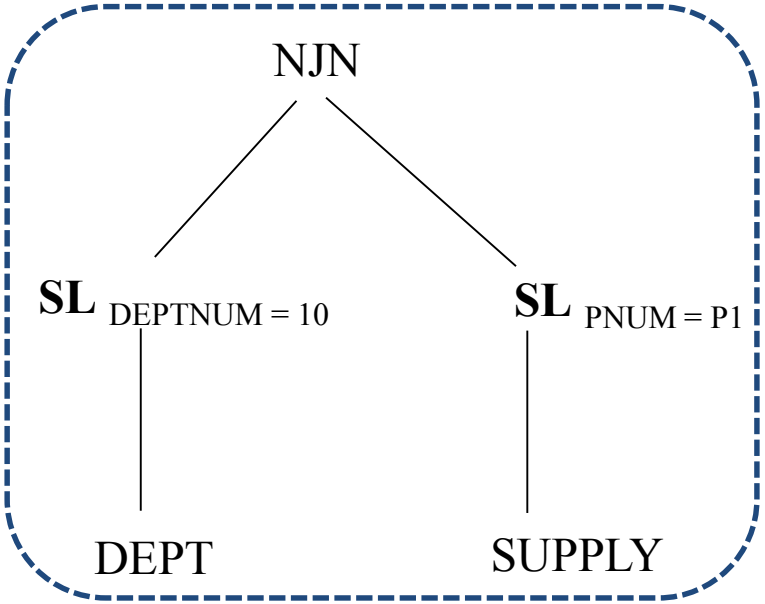


Removing Common Sub - Expression





Removing Common Sub - Expression



Do we need to apply criteria 1 and/or 2? No, Already Simplified

Last Example

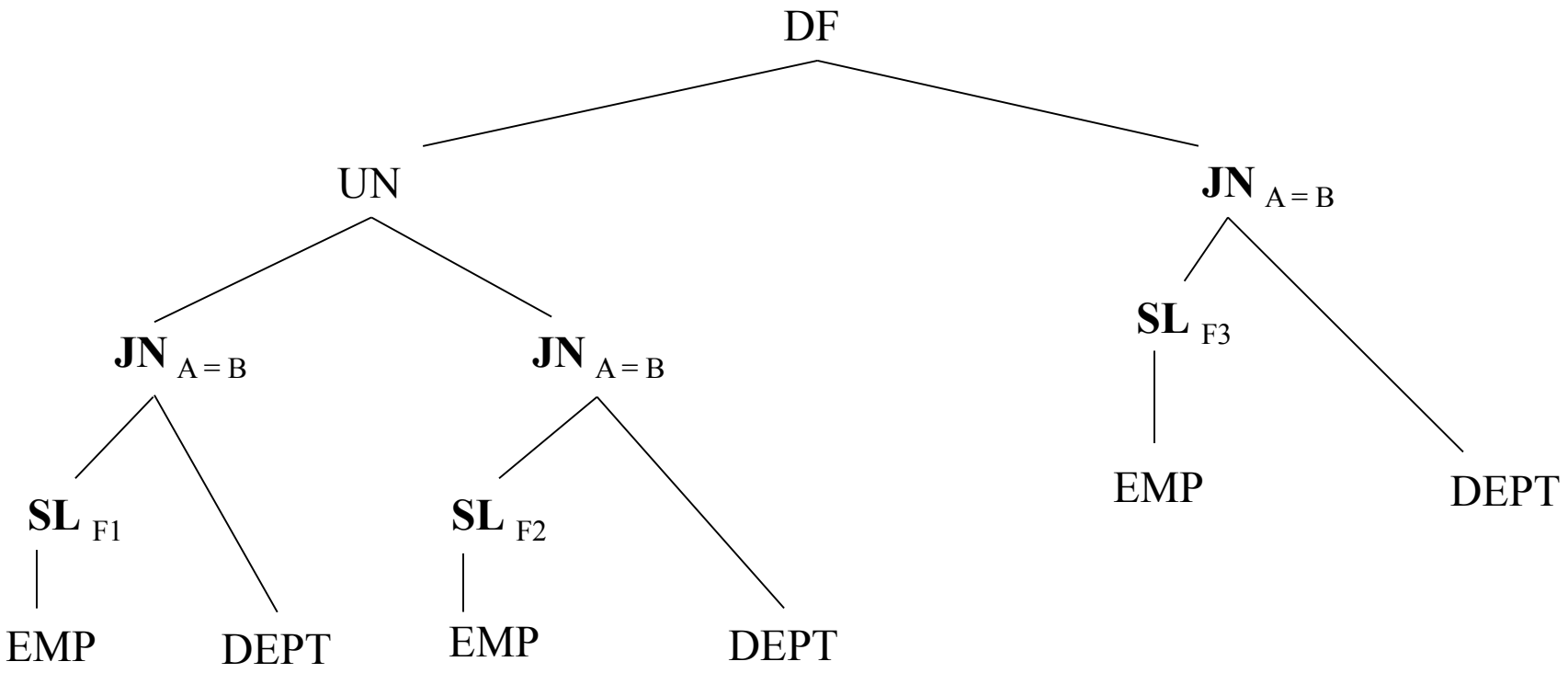
EMP (EMPNUM, DEPTNUM, NAME, SAL, AGE)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Consider the following global query:

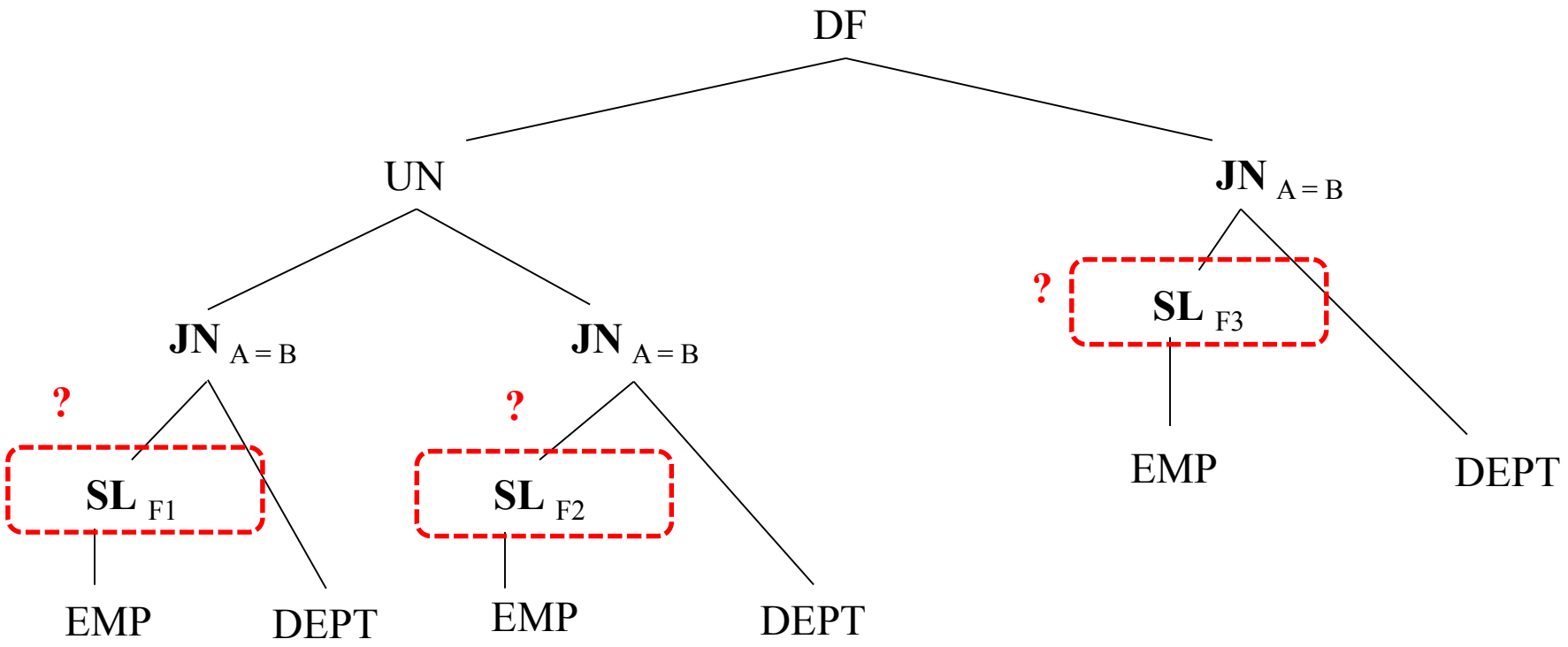
$$\left((SL_{F1}\ EMP\ \ JN_{A=B}\ \ DEPT)\ \ UN\ \ (SL_{F2}\ EMP\ \ JN_{A=B}\ \ DEPT)\right)\ \ DF\ \ (SL_{F3}\ \ EMP\ \ JN_{A=B}\ \ DEPT)$$

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

Operator Tree

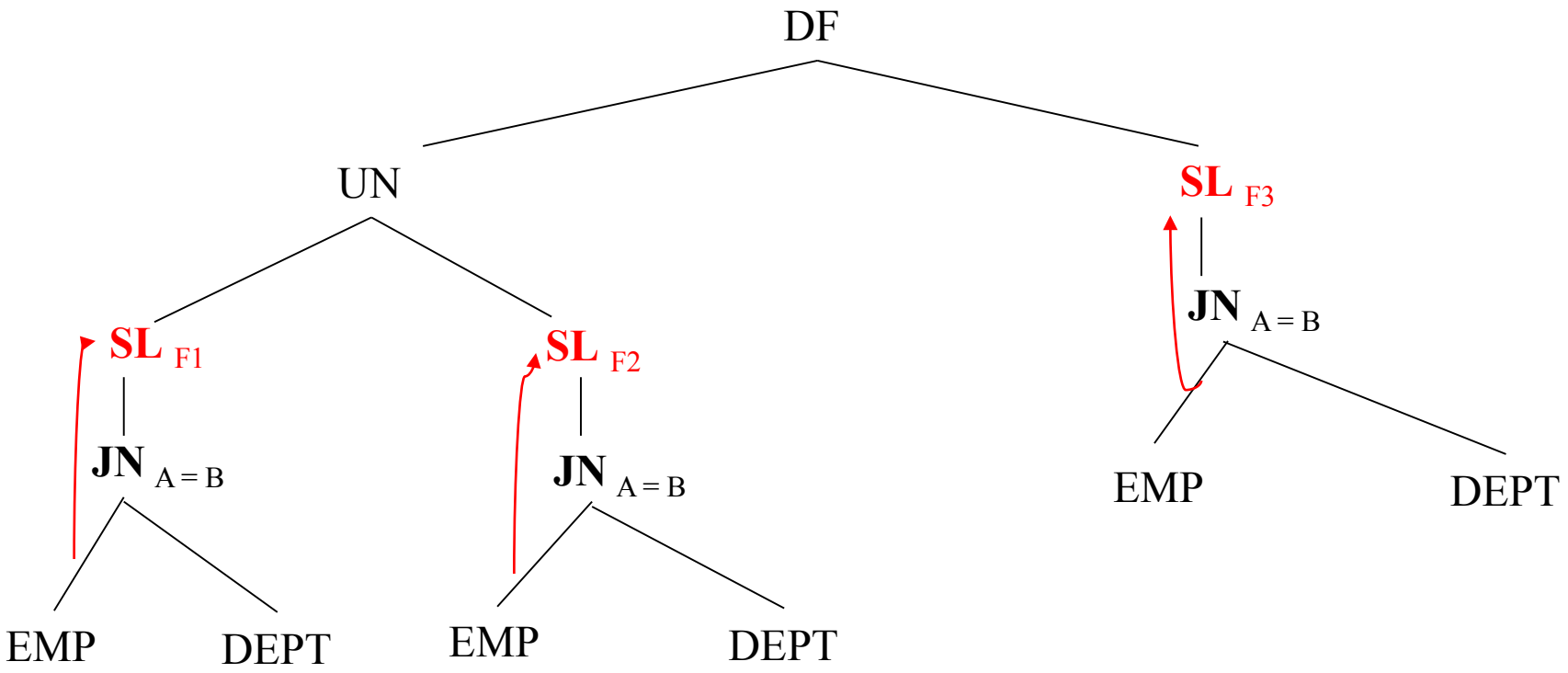


Finding Common Sub-Expression

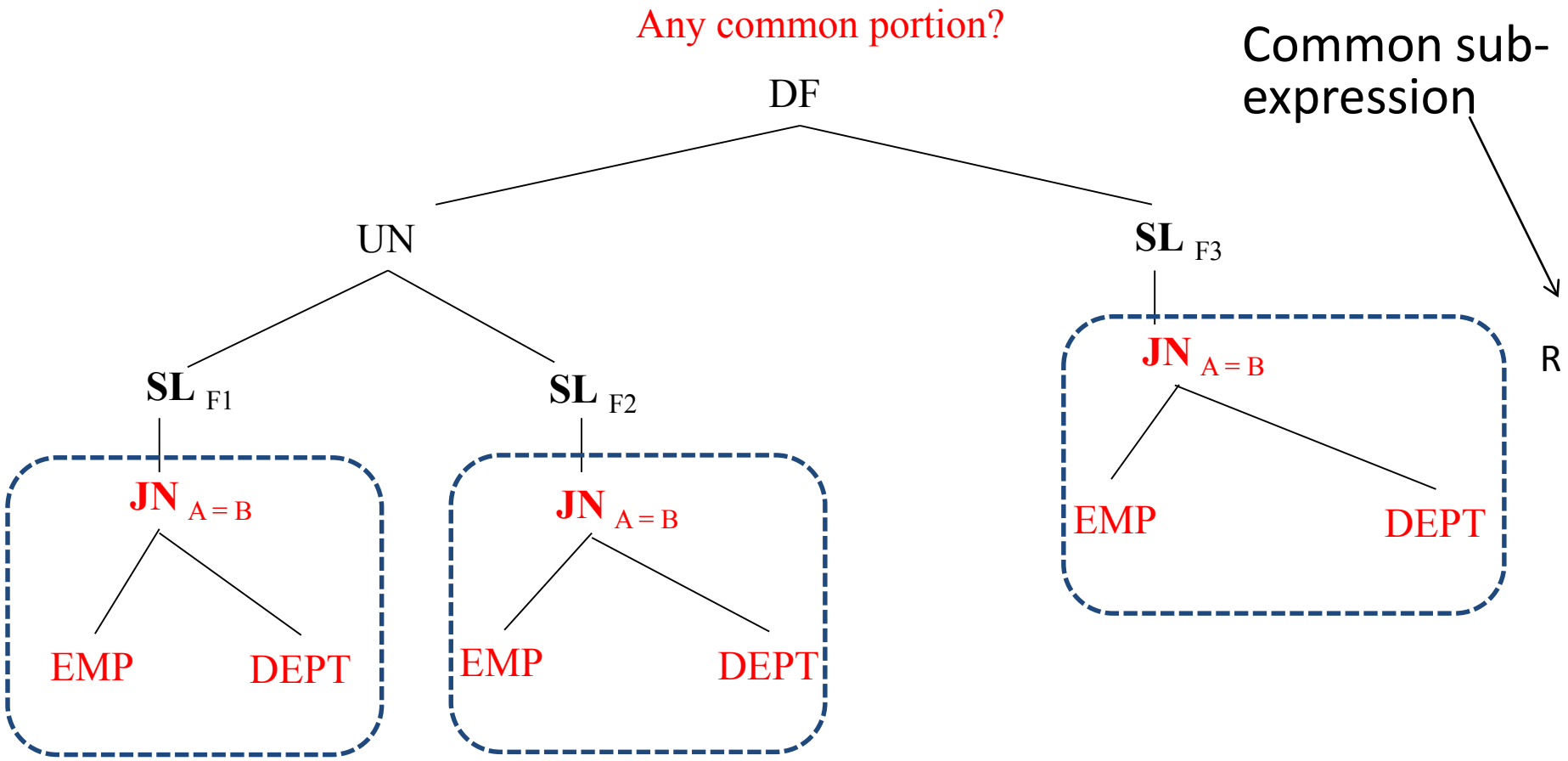


Finding Common Sub-expression

Any common portion?

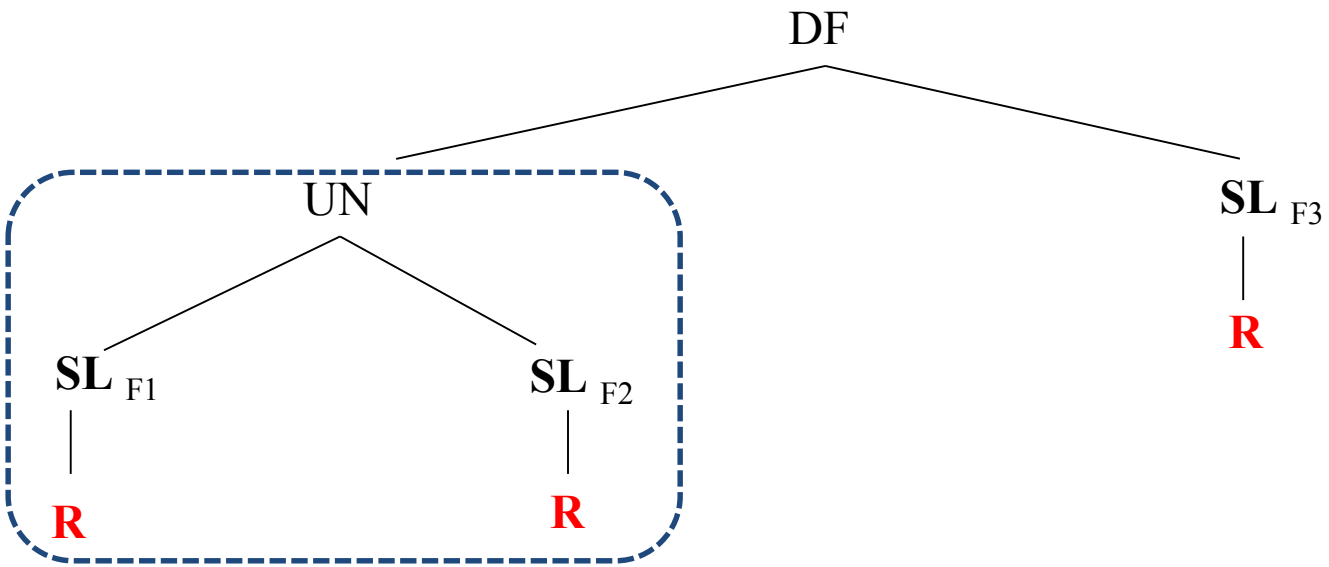


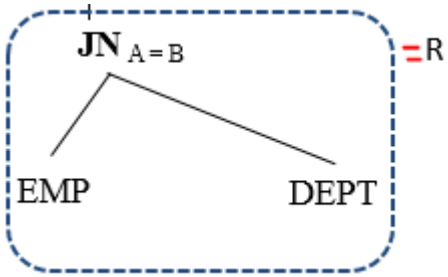
Finding Common Sub-expression



Finding Common Sub-expression

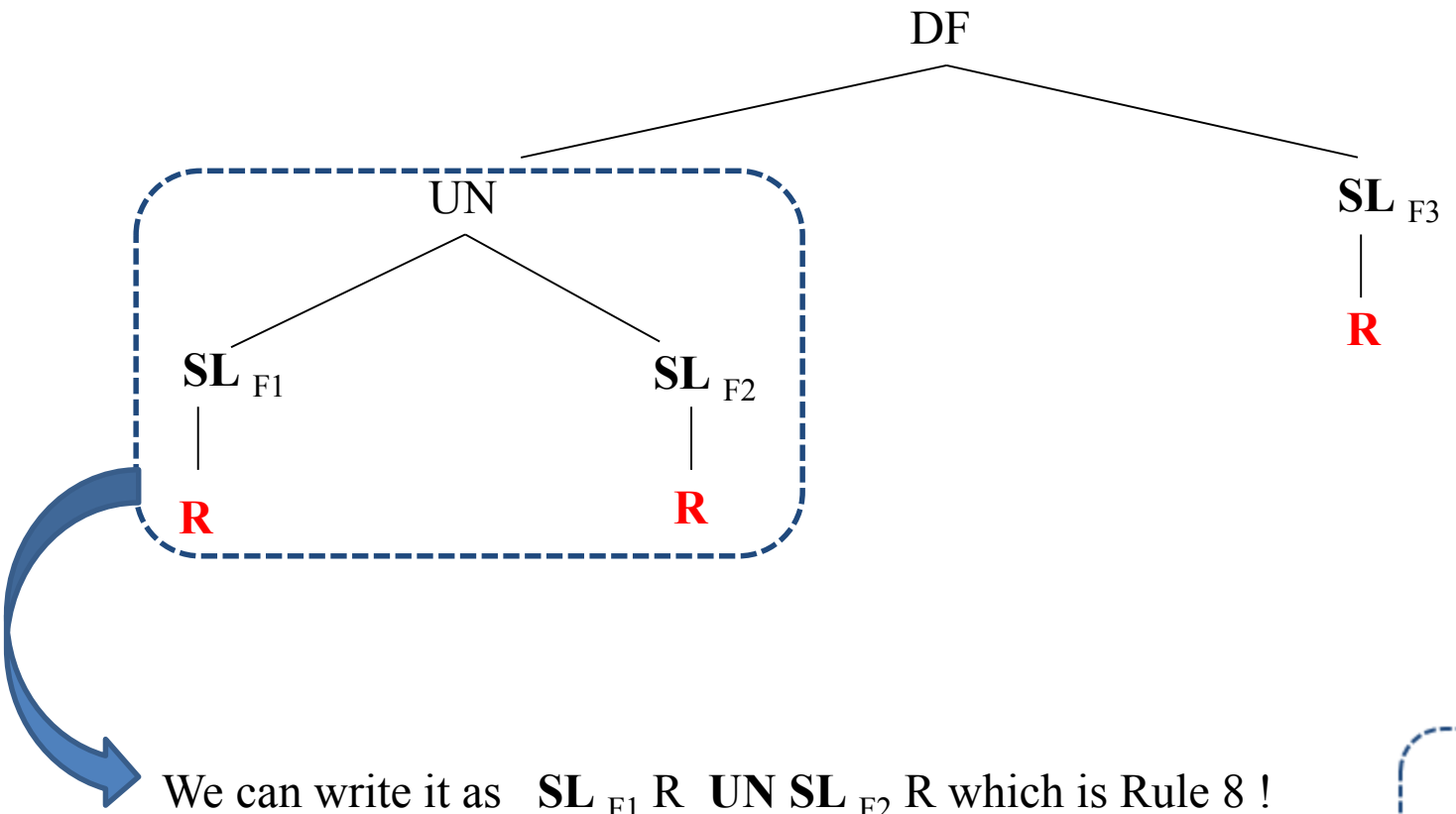
Any common portion?



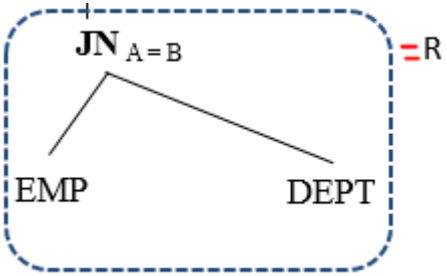


Finding Common Sub-expression

Any common portion?

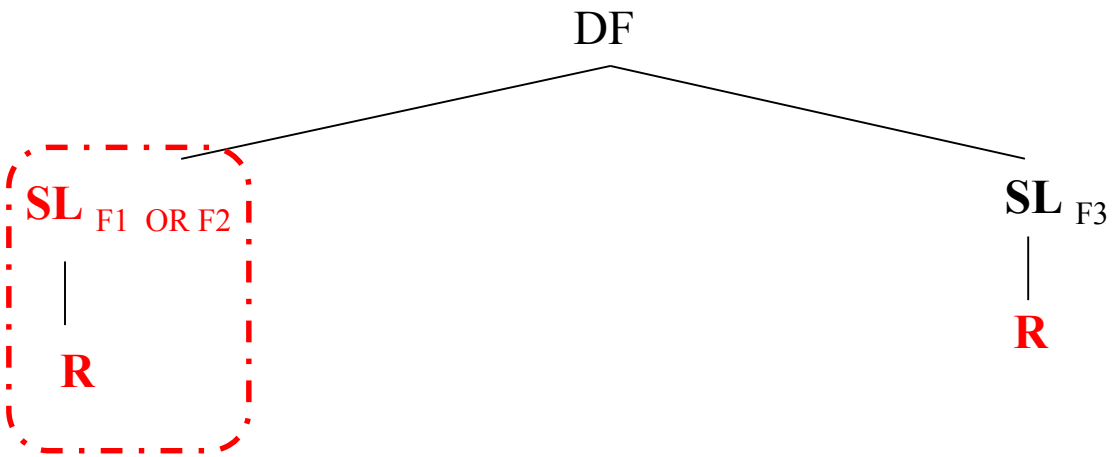


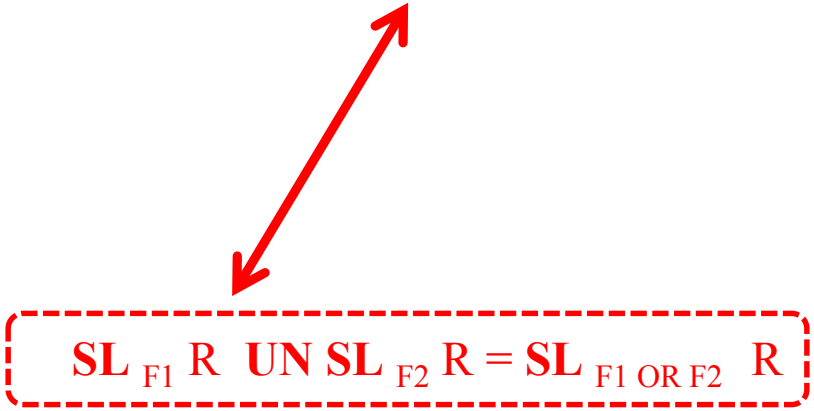
We can write it as **SL** $_{F1}$ R **UN SL** $_{F2}$ R which is Rule 8 !

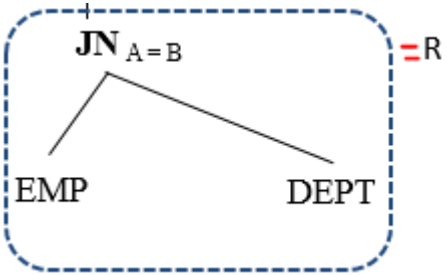


Removing Common Sub-expression

Any common portion?

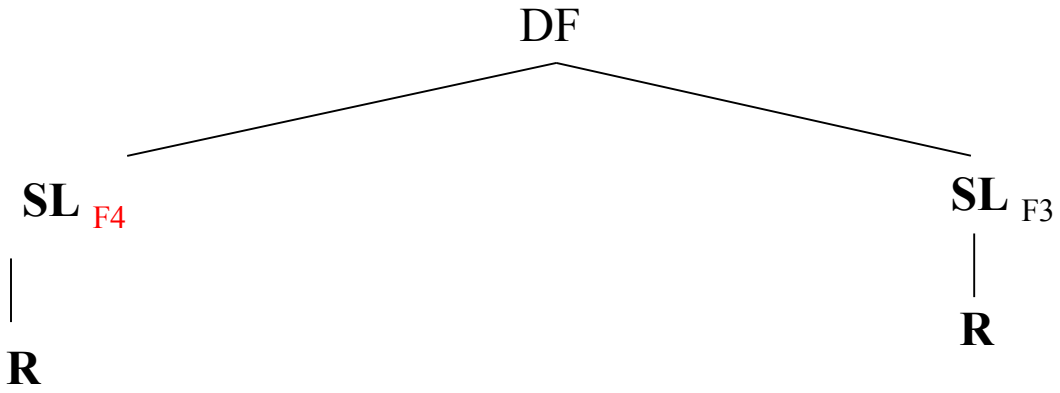




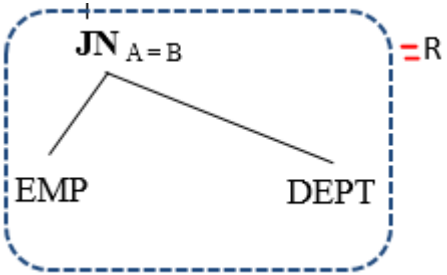


Finding Common Sub-expression

Any common portion?

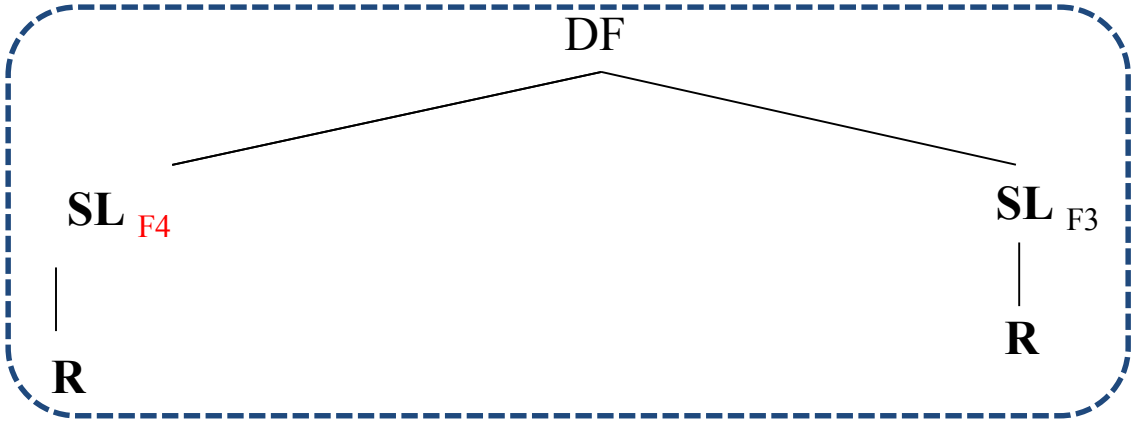


Let, F4 = F1 OR F2



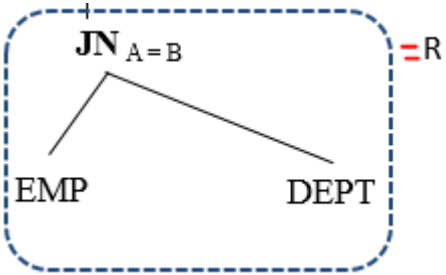
Finding Common Sub-expression

Any common portion?



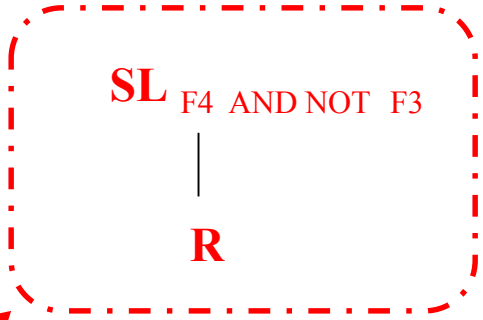
We can write it as **SL** $_{F4}$ R **DF SL** $_{F3}$ R which is Rule 9 !

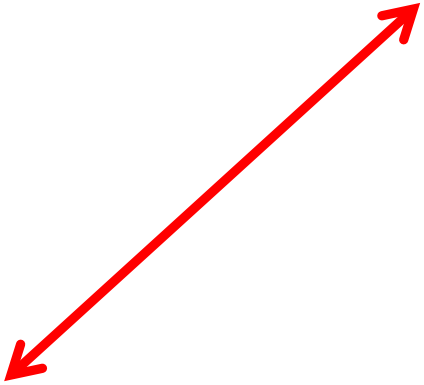
Let, F4 = F1 OR F2



Removing Common Sub-expression

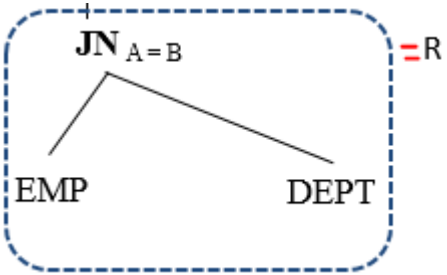
Any common portion?





SL $_{F4}$ R **DF SL** $_{F3}$ R = **SL** $_{F4\ AND\ NOT\ F3}$ R

Let, F4 = F1 OR F2

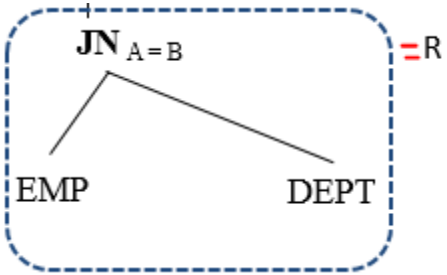


Removing Common Sub-expression

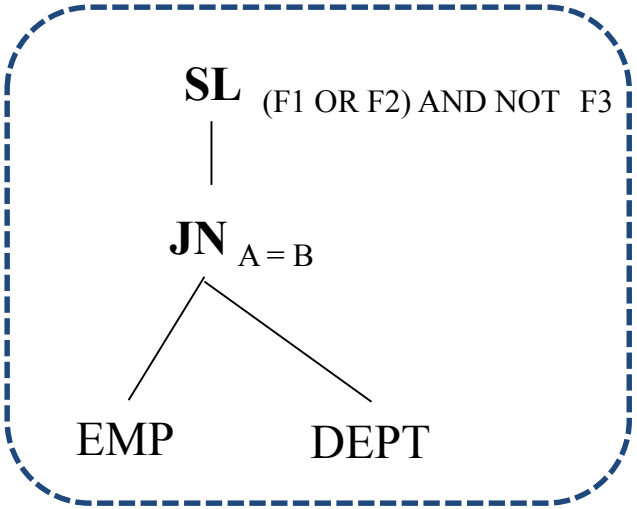
SL (F1 OR F2) AND NOT F3

|

R

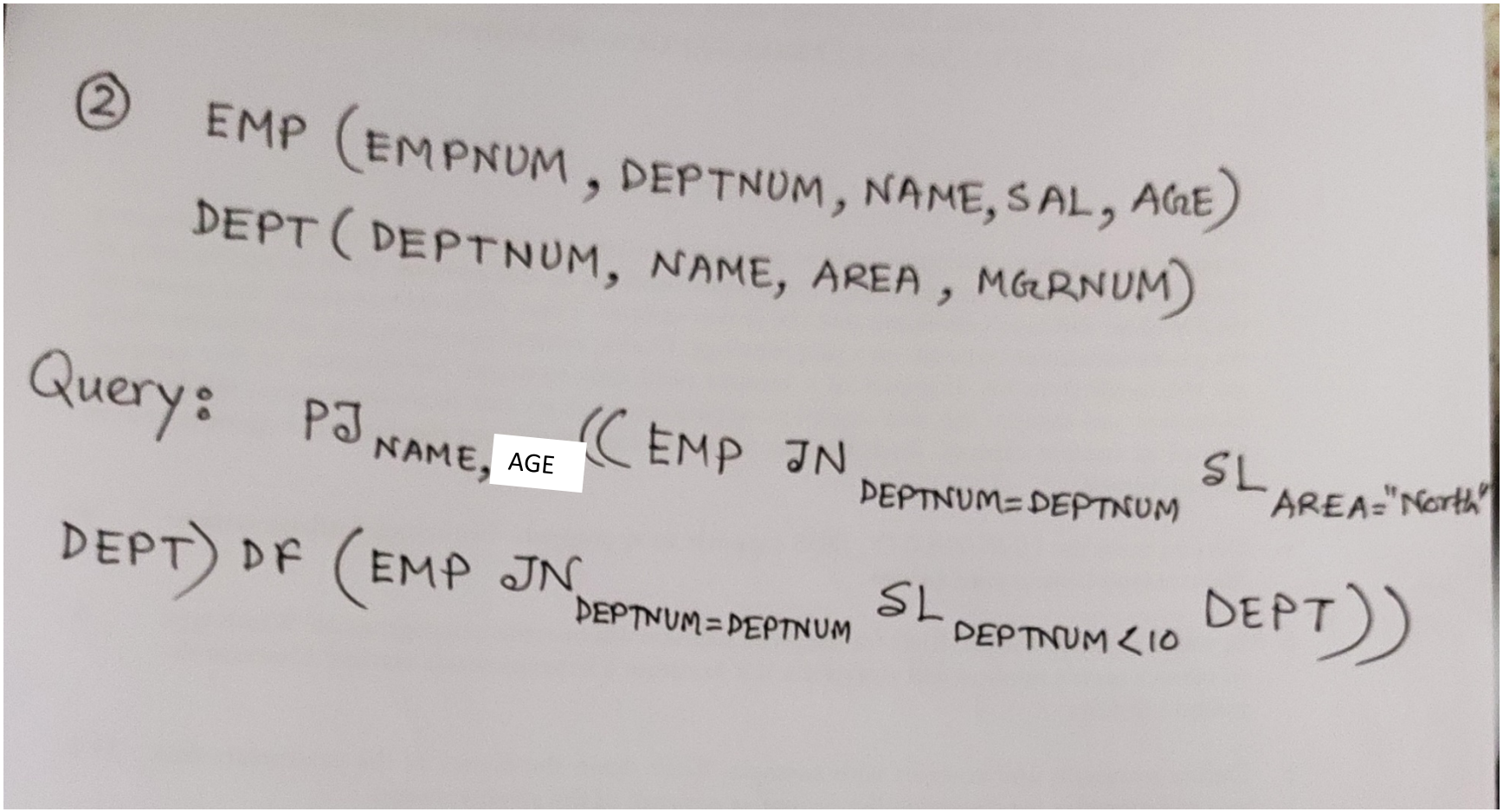


Removing Common Sub-expression



Can we apply Criterion 1 and/or 2? Yes. Criteria 2

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.

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

Exercise 3

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

Consider the following global query:

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

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

Exercise 4

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

Consider the following global query:

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

$$\left((EMP\ \ JN_{A=B}\ \ DEPT)\ \ DF\ \ (SL_{F3}\ EMP\ \ JN_{A=B}\ \ DEPT)\right)$$

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