CSE 4125: Distributed Database Systems
Chapter – 5

Translation of Global Queries to Fragment Queries.

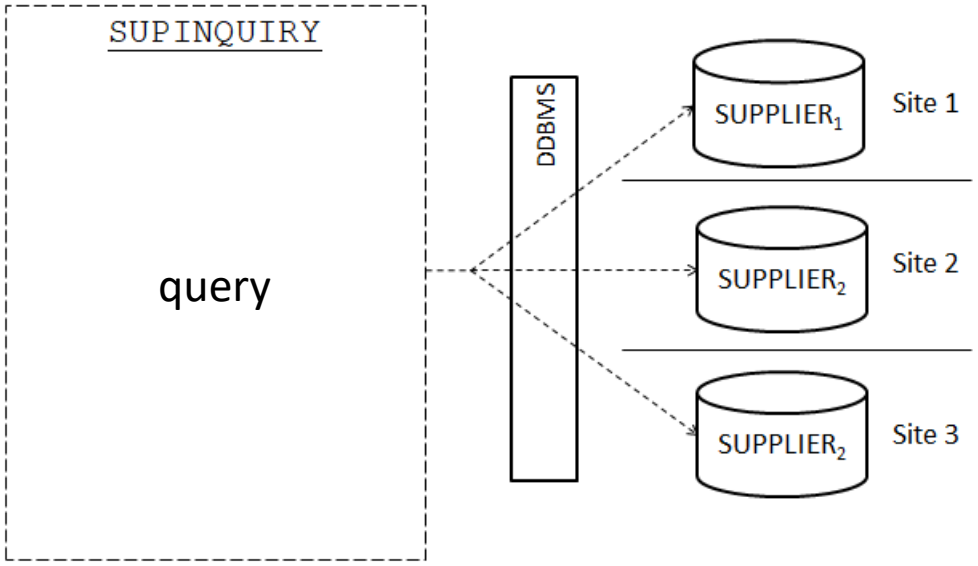
(part – A)

Outline

- Query Transform.
- Operator Tree.
- Simplification of Operator Tree.
- Equivalence Transformation for Queries (step-by-step).
- Transforming Global Queries into Fragment Queries.

Query Transform

In case of Level – 1 Transparency –



Global Query → Fragment Query

Query Transform (contd.)

Steps for Global Query → Fragment Query:

1. *Non-distributed* (Equivalence Transformation for Queries):
   - Query → Operator Tree.
   - Operator Tree → Simplified Operator Tree.
2. *Distributed:*
   - Global Query → Fragment Query.

Equivalent Expressions of Queries

- **Q_1:** PJ $_{\text{NAME, DEPTNUM}}$ SL $_{\text{DEPTNUM = 15}}$ *EMP*
- **Q_2:** SL $_{\text{DEPTNUM = 15}}$ PJ $_{\text{NAME, DEPTNUM}}$ *EMP*

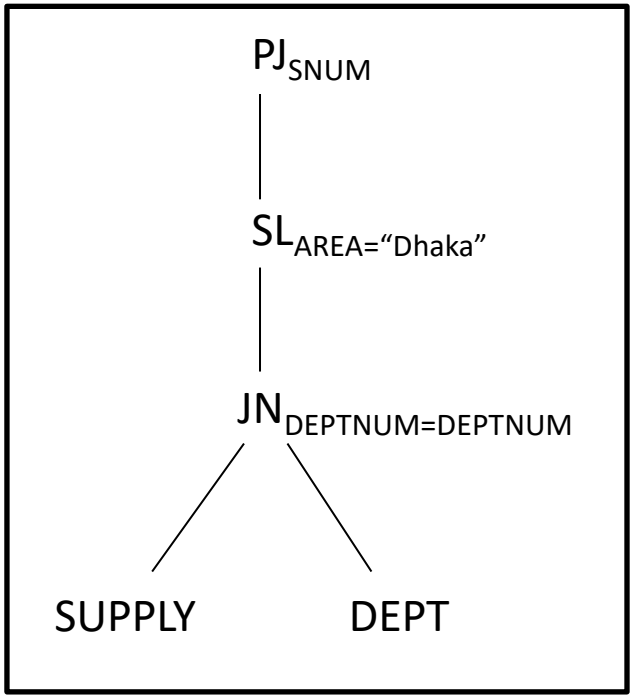
$$Q_1 \leftrightarrow Q_2$$

Operator Tree

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

Q1: $PJ_{SNUM}\ SL_{AREA=\text{“Dhaka”}}\ (SUPPLY\ JN_{DEPTNUM=DEPTNUM}\ DEPT)$

- Operator Tree for Q1:



Simplification of Operator Tree

Criteria to simplify:

Criterion – 1:

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

- To get rid of unnecessary attributes.

Criterion – 2:

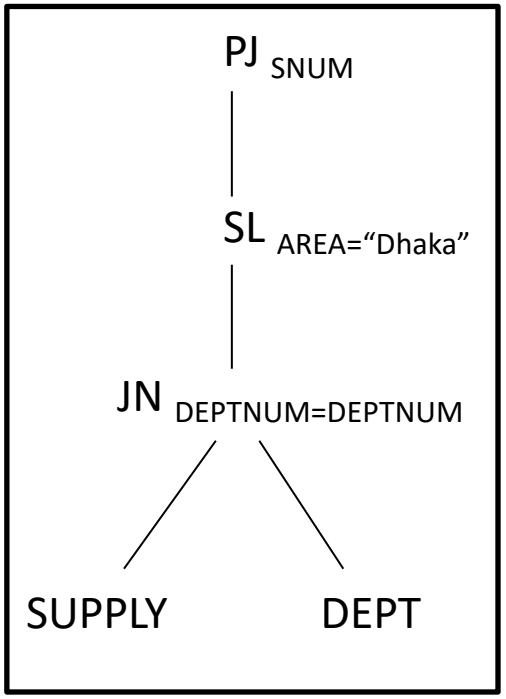
Push *SL* and *PJ* as down as possible in the tree.

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

Simplification of Operator Tree (contd.)

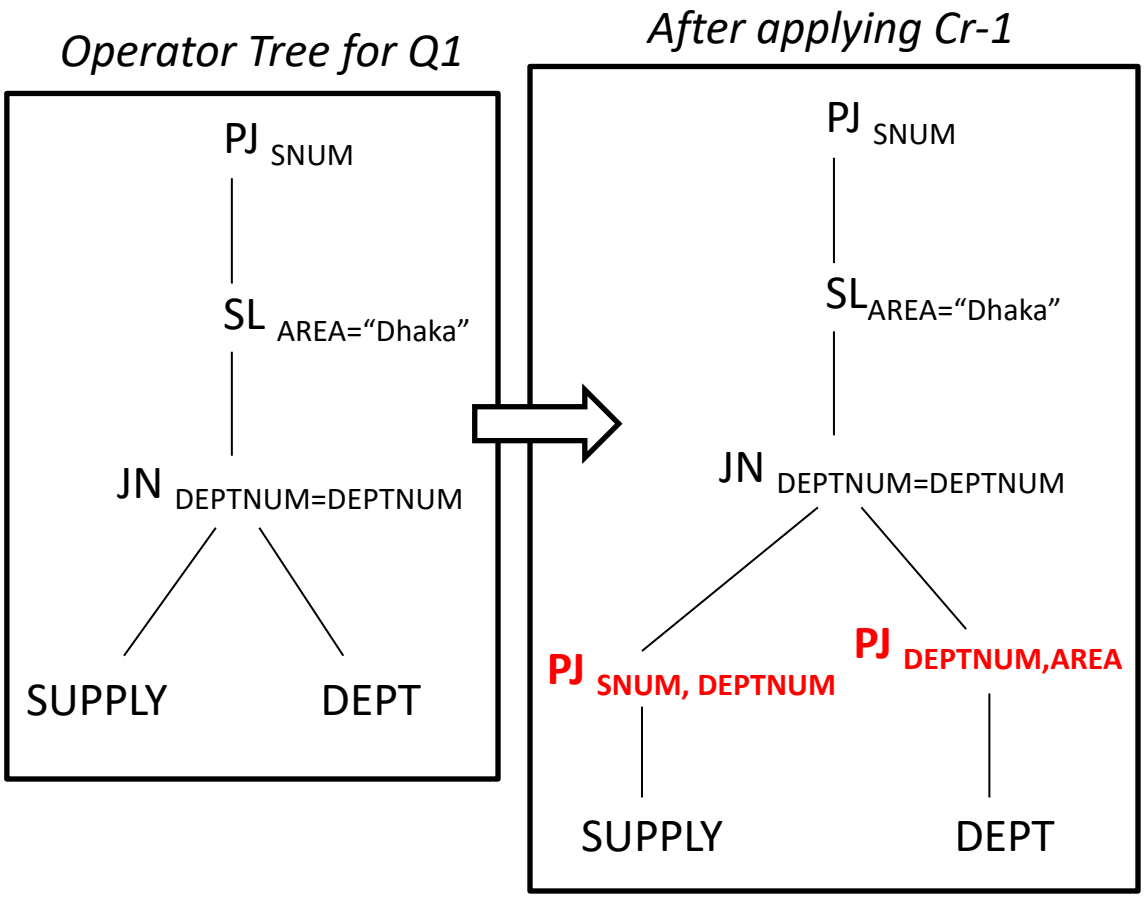
- *SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*
- *DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

Operator Tree for Q1



Simplification of Operator Tree (contd.)

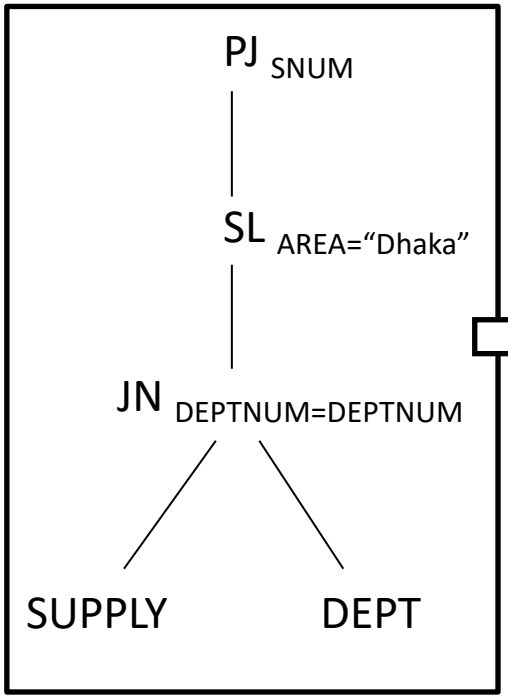
- *SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*
- *DEPT (DEPTNUM, NAME, AREA, MGRNUM)*



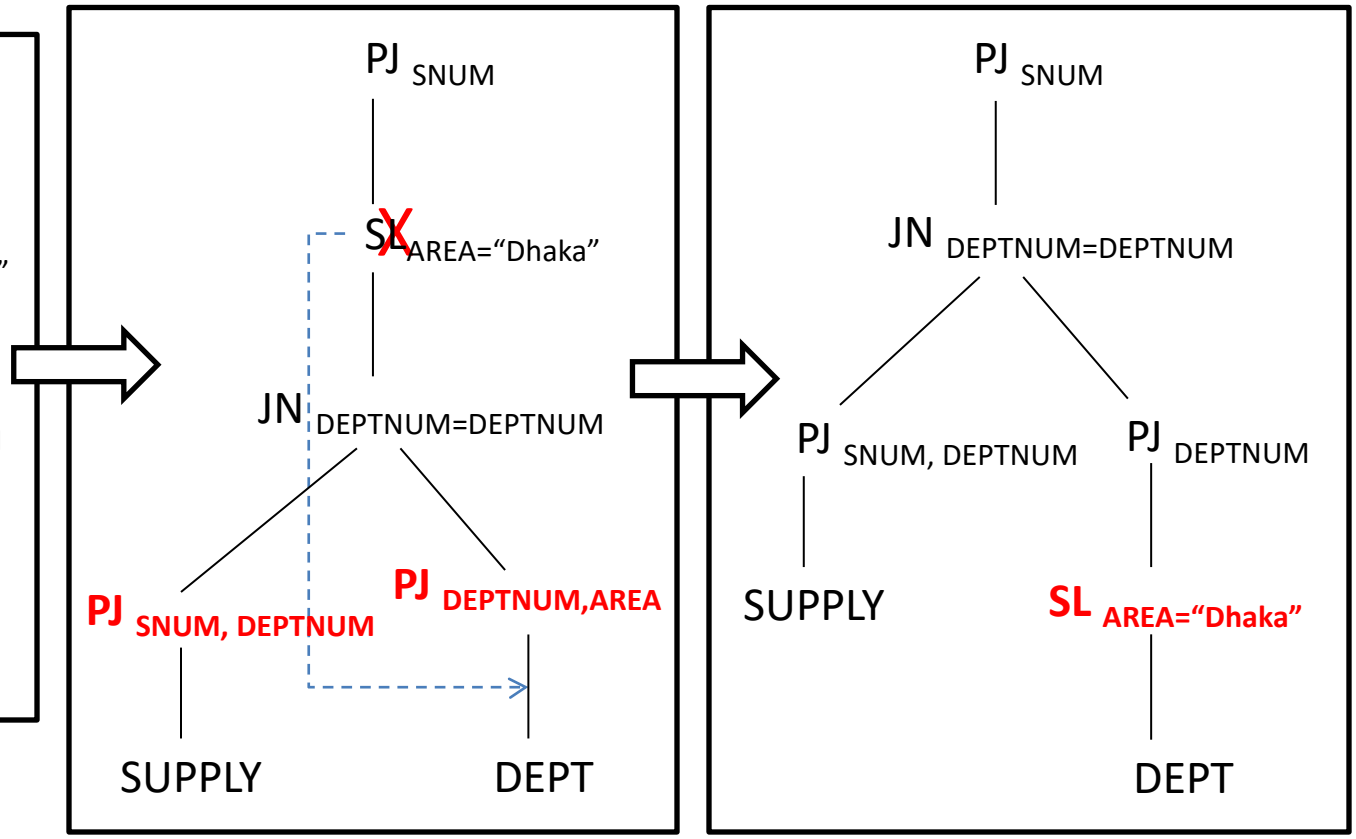
Simplification of Operator Tree (contd.)

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

Operator Tree for Q1



After applying Cr-2



Practice Session

Q2: $PJ_{EMP.NAME}$ ((*EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*) DF ($SL_{SAL > 35K}$ *EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*))

- Operator Tree for Q2?

Practice Session (answer)

Q2: $PJ_{EMP.NAME}$ (($\underline{EMP}$ $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ $\underline{DEPT}$) DF ($SL_{SAL > 35K}$ $\underline{EMP}$ $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ $\underline{DEPT}$))

EMP

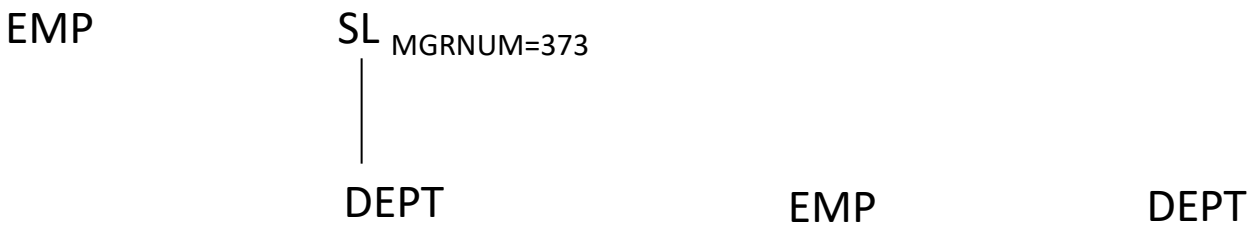
DEPT

EMP

DEPT

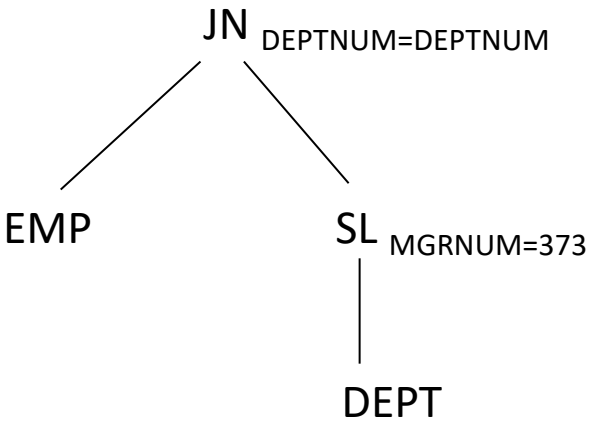
Practice Session (answer)

Q2: $PJ_{EMP.NAME}$ ((*EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*) DF ($SL_{SAL > 35K}$ *EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*))



Practice Session (answer)

Q2: $PJ_{EMP.NAME}$ ((*EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*) DF ($SL_{SAL > 35K}$ *EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*))

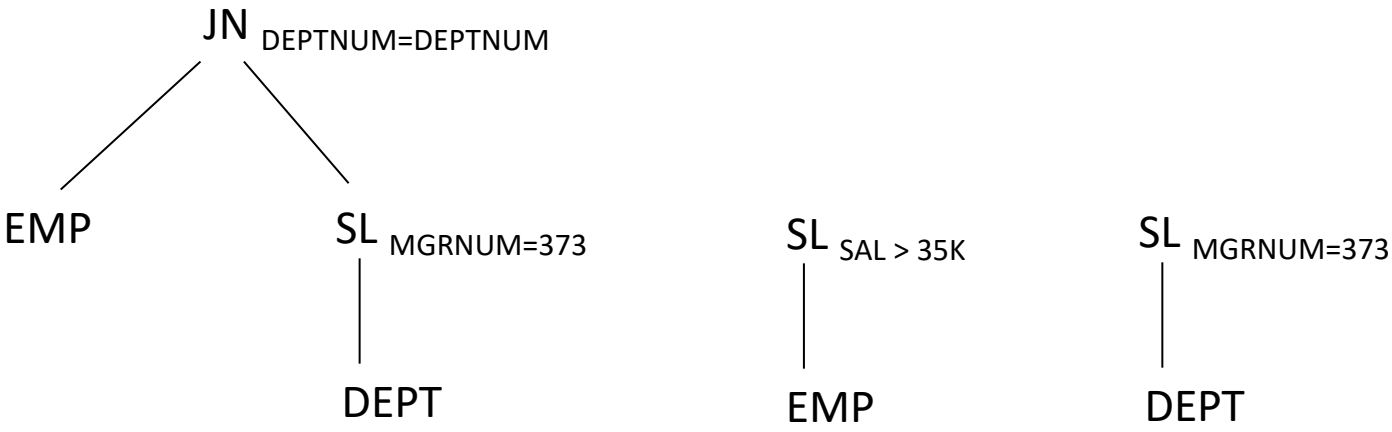


EMP

DEPT

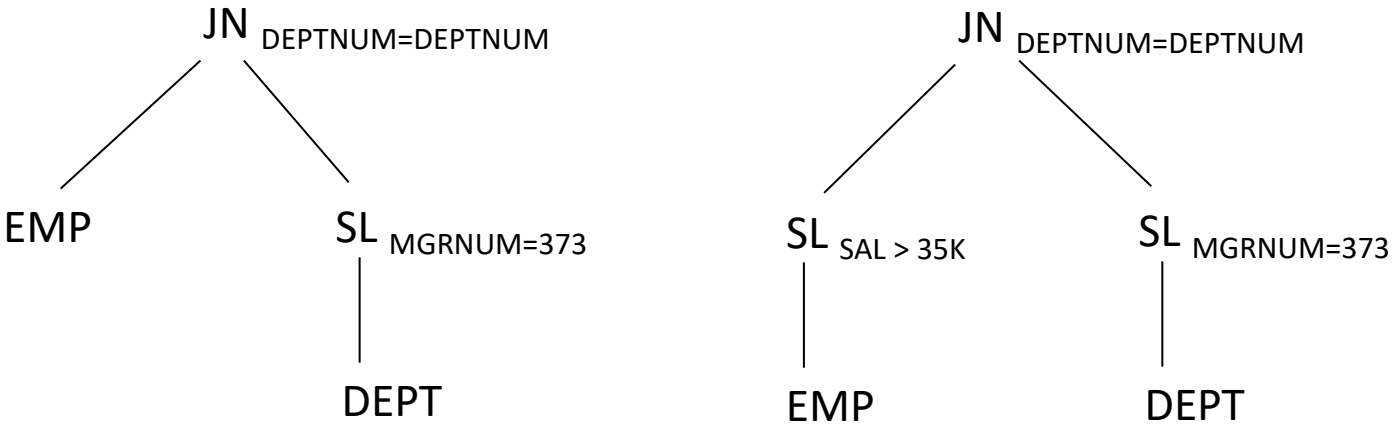
Practice Session (answer)

Q2: $PJ_{EMP.NAME}$ ((*EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*) DF ($SL_{SAL > 35K}$ *EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*))



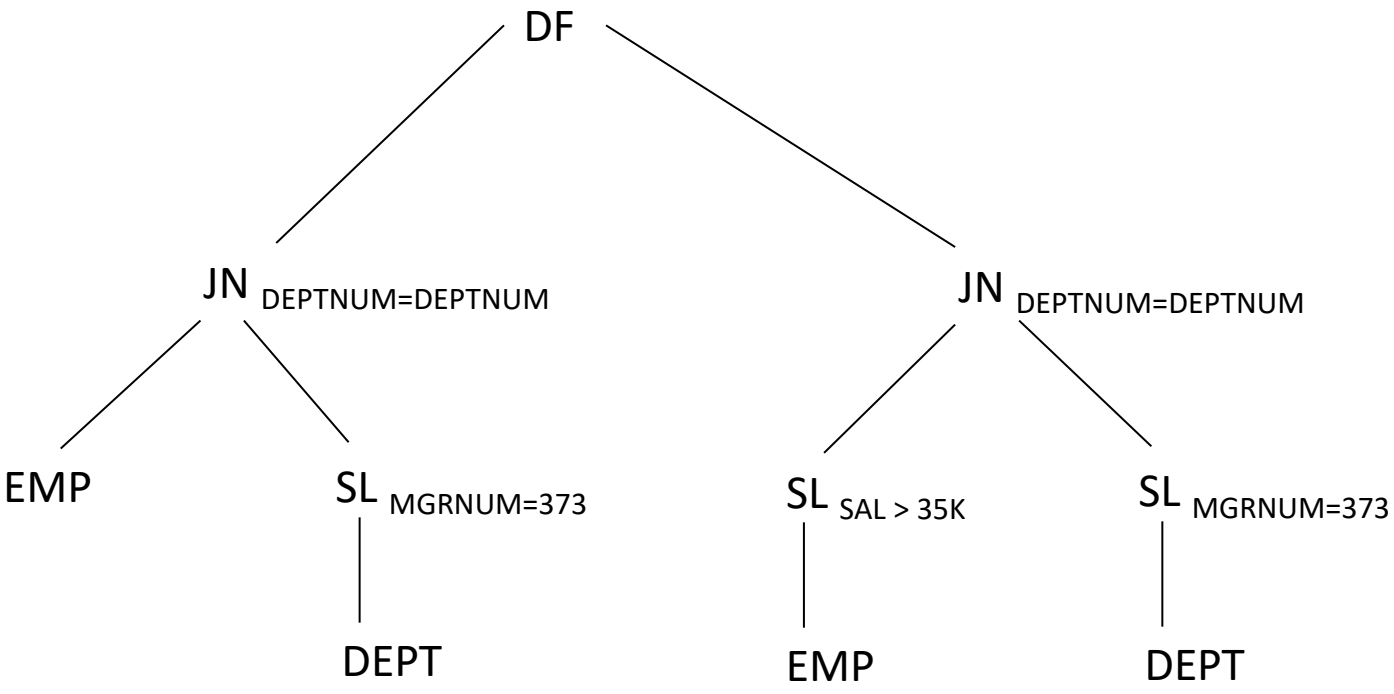
Practice Session (answer)

Q2: $PJ_{EMP.NAME}$ ((*EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*) DF ($SL_{SAL > 35K}$ *EMP* $\underline{JN}_{\underline{DEPTNUM=DEPTNUM}}$ $SL_{MGRNUM=373}$ *DEPT*))



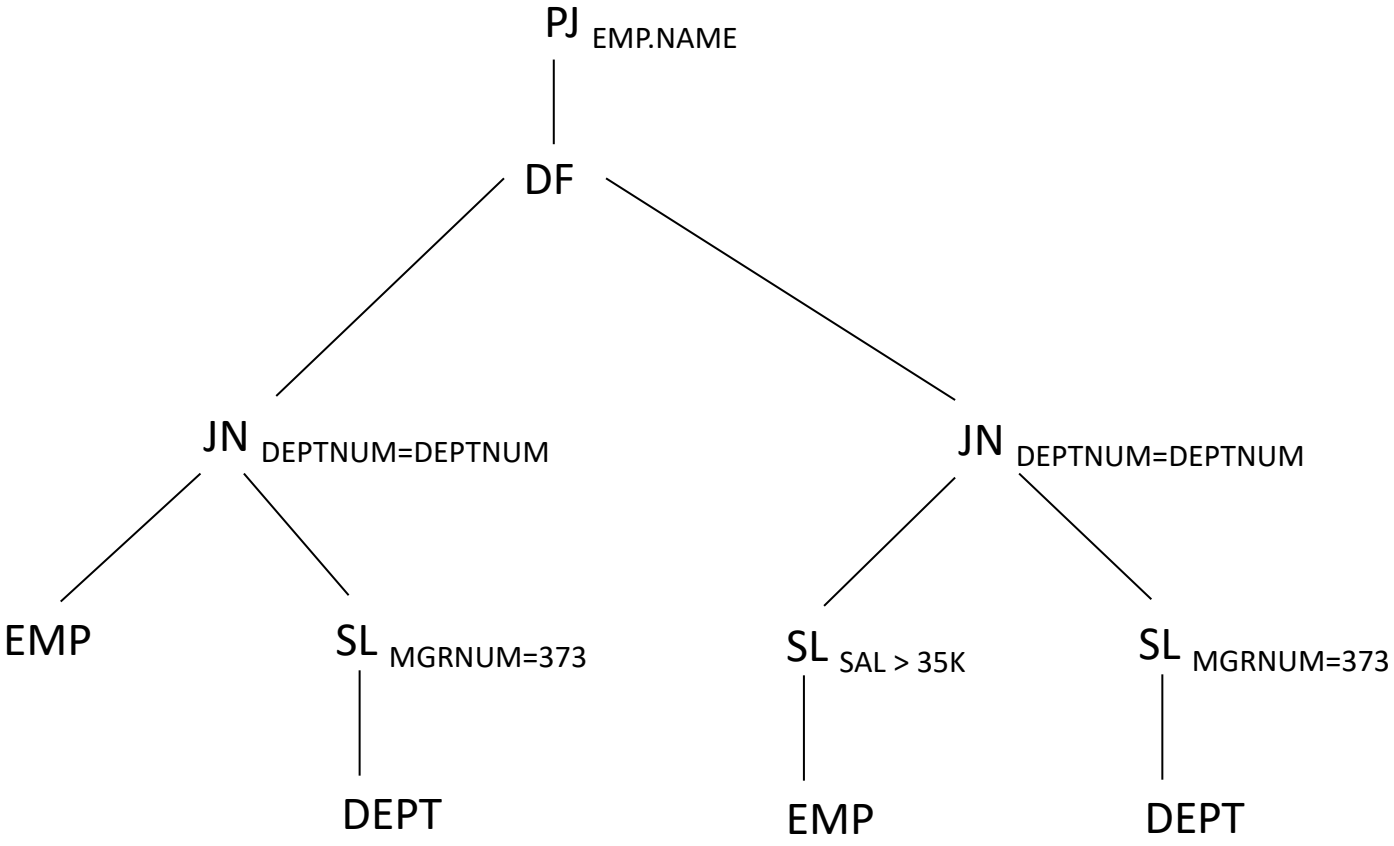
Practice Session (answer)

Q2: $PJ_{EMP.NAME}$ ((*EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*) DF ($SL_{SAL > 35K}$ *EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*))



Practice Session (answer)

Q2: $\underline{PJ_{EMP.NAME}}$ ((*EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*) DF ($SL_{SAL > 35K}$ *EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*))

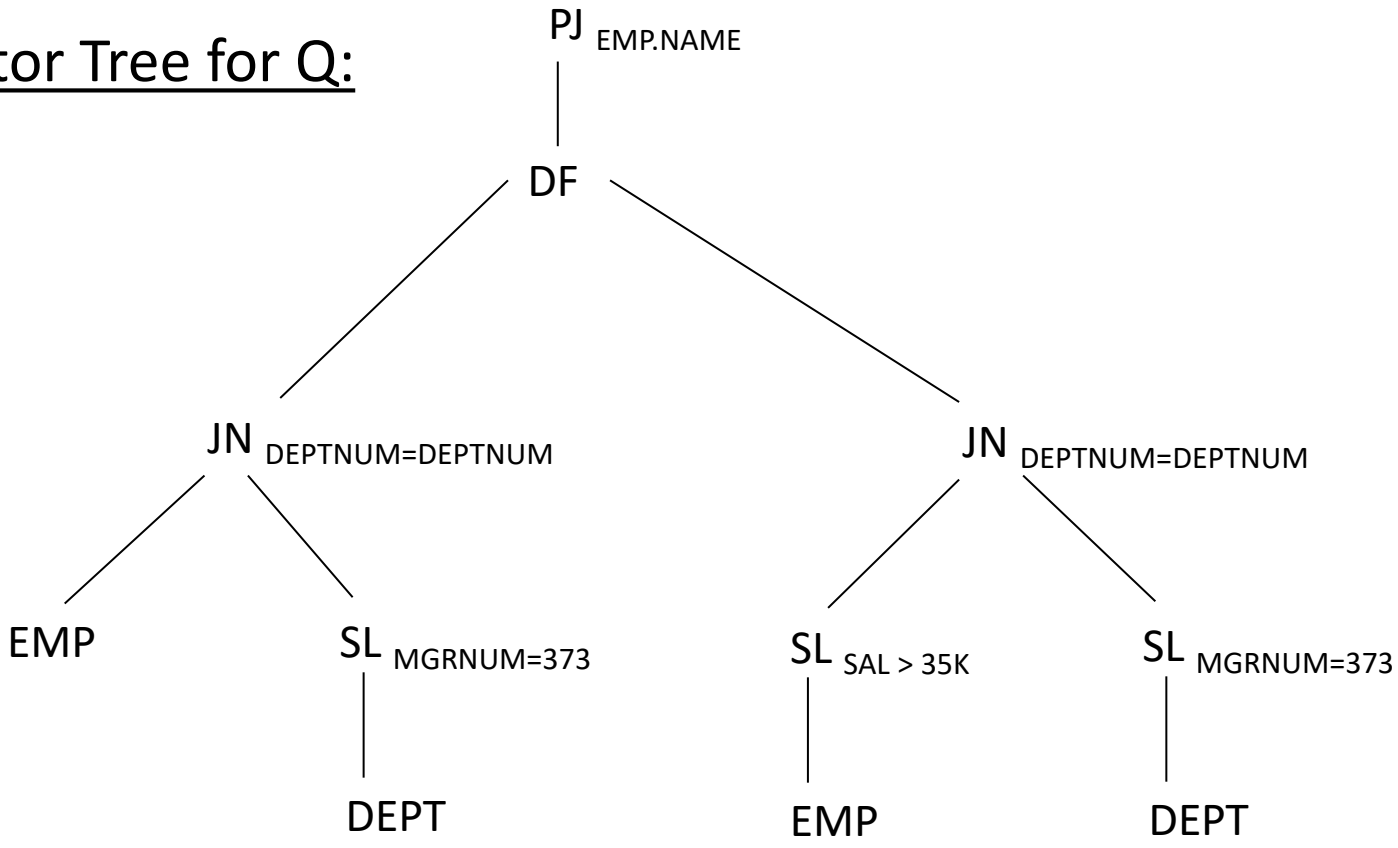


Equivalence Transformation for Queries (step-by-step)

Given query and Operator Tree

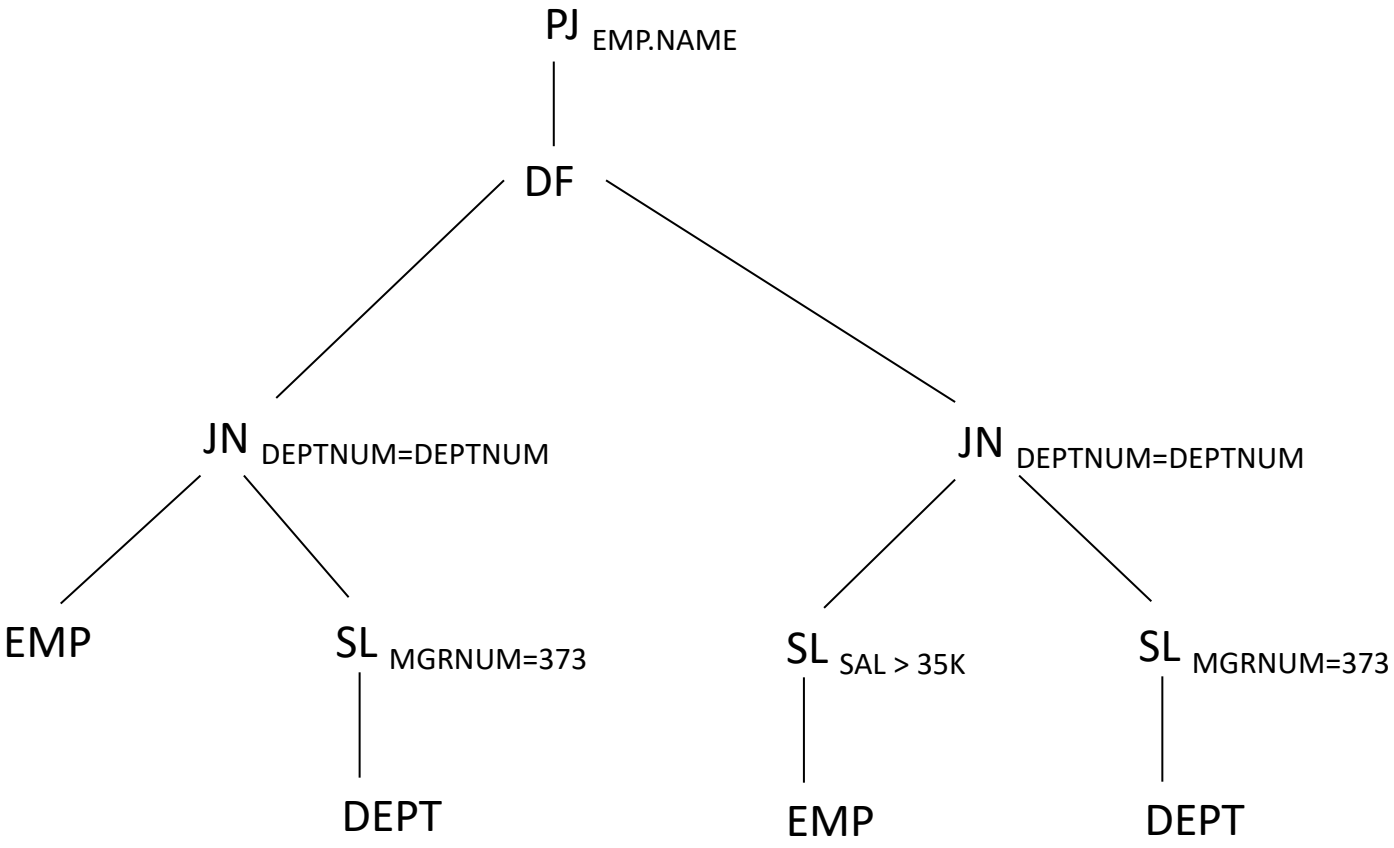
Q: $PJ_{EMP.NAME}$ ((EMP $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ $DEPT$) DF ($SL_{SAL > 35K}$ EMP $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ $DEPT$))

Operator Tree for Q:



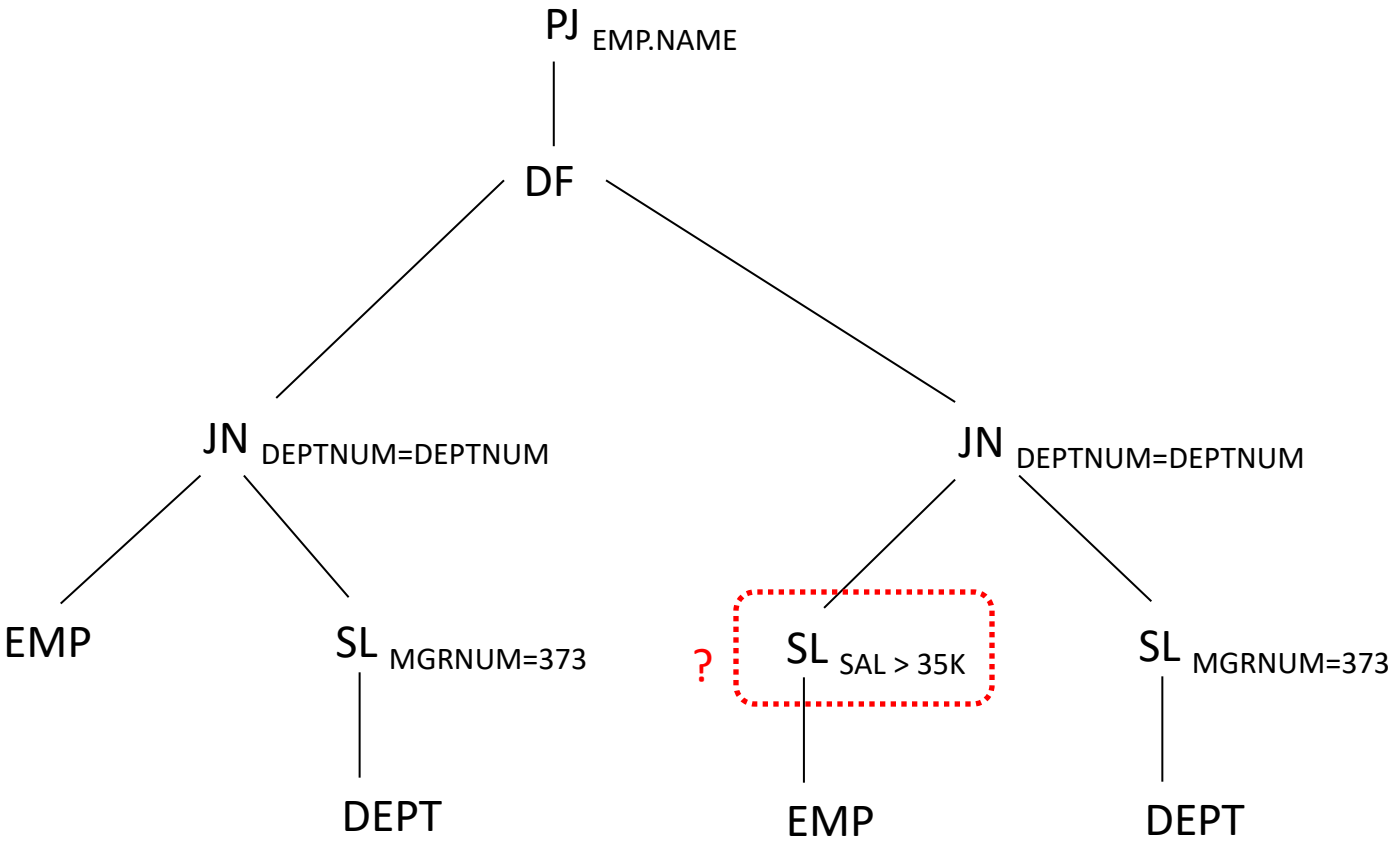
Finding Common Sub-expression

- Any common portion?



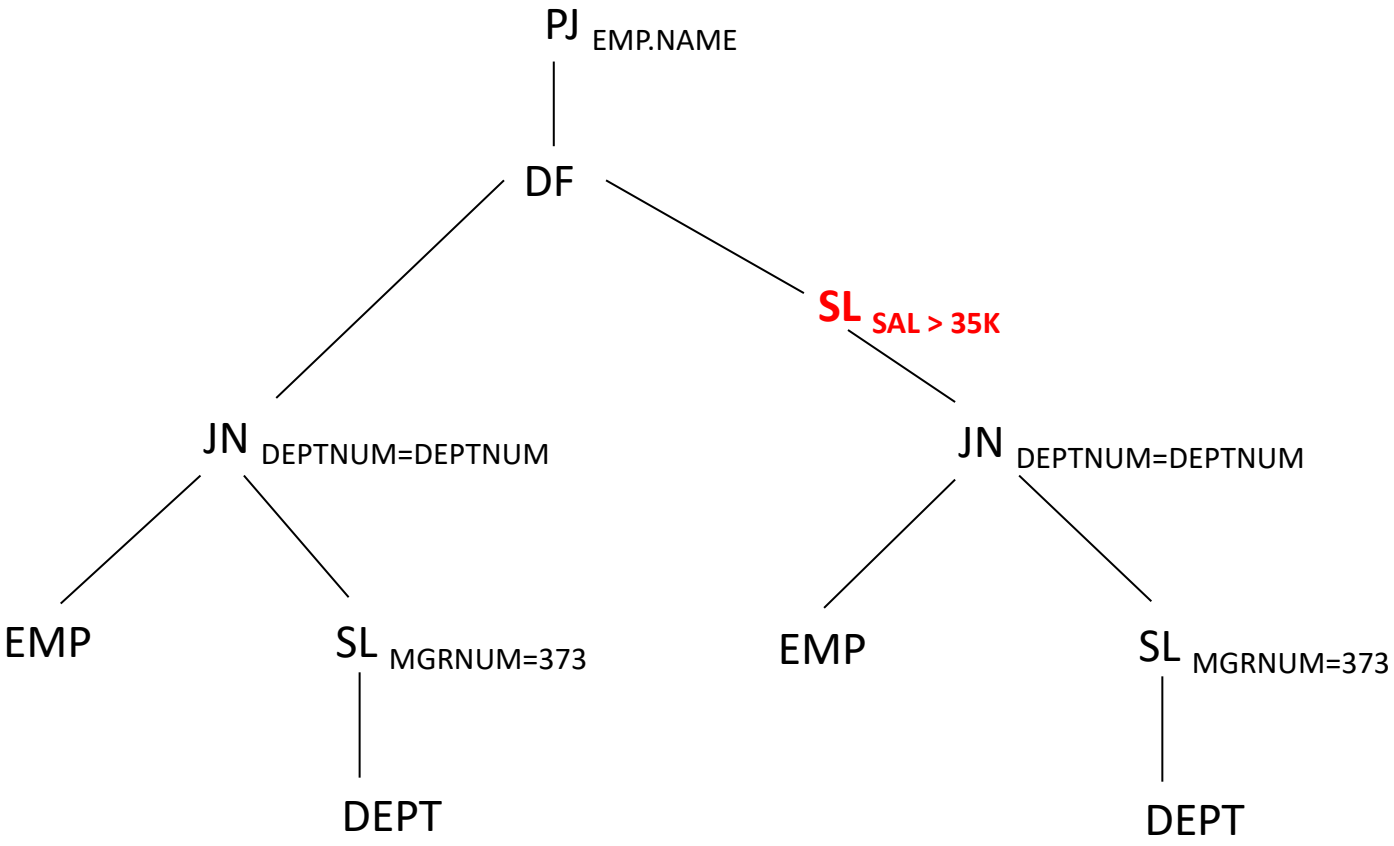
Finding Common Sub-expression

- Any common portion?

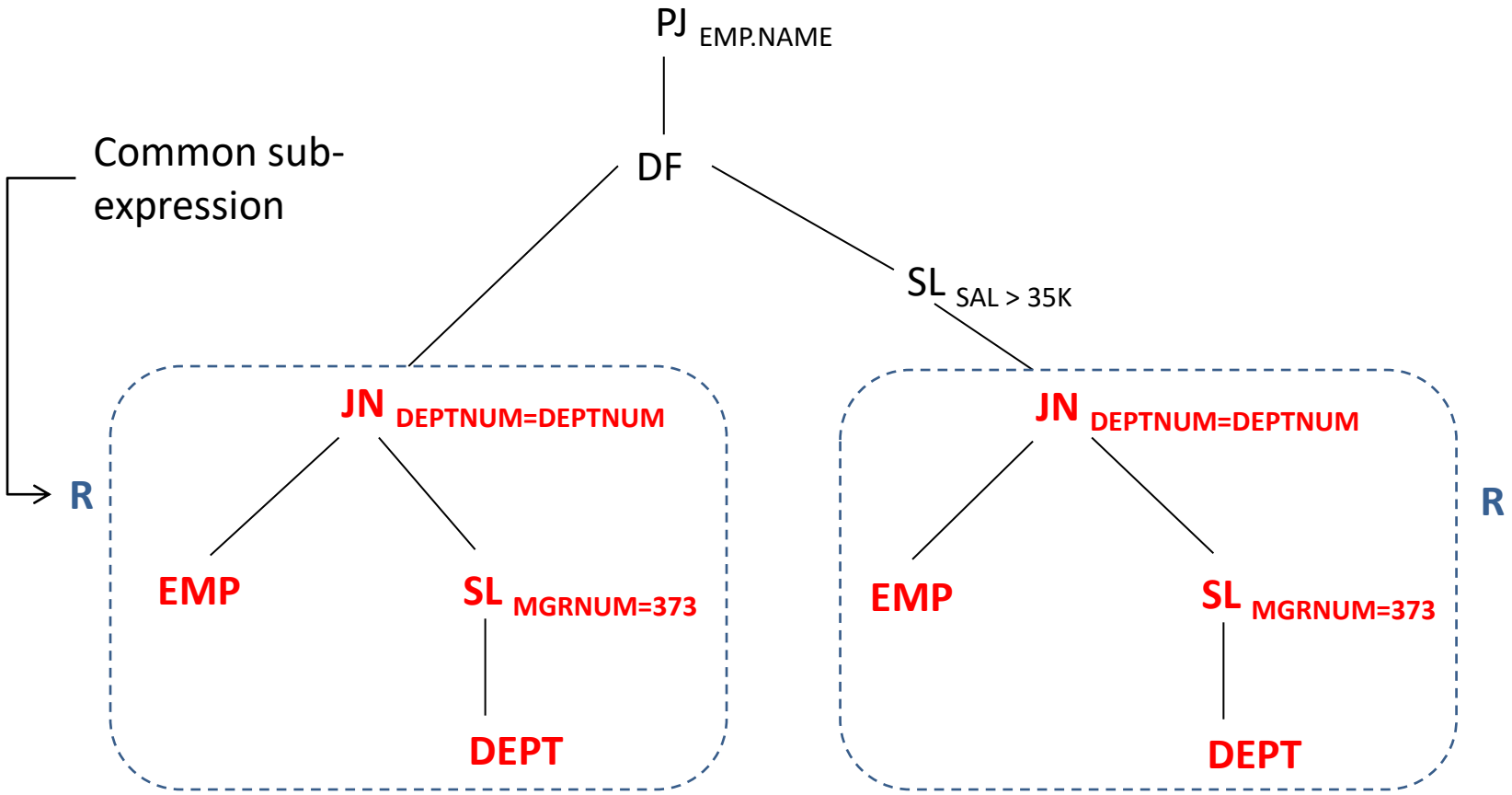


Finding Common Sub-expression

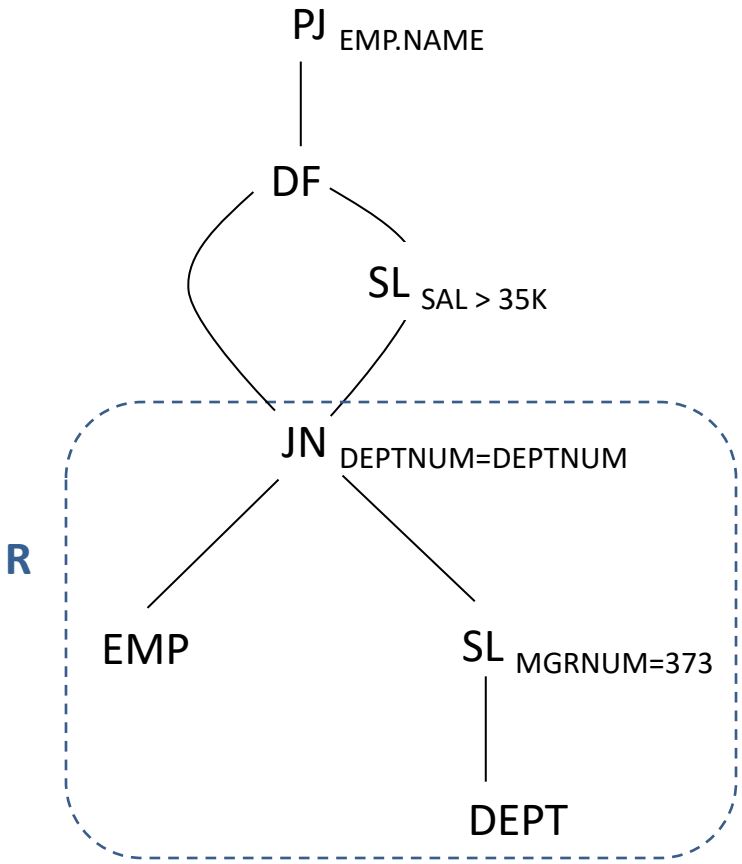
- Any common portion? NOW?



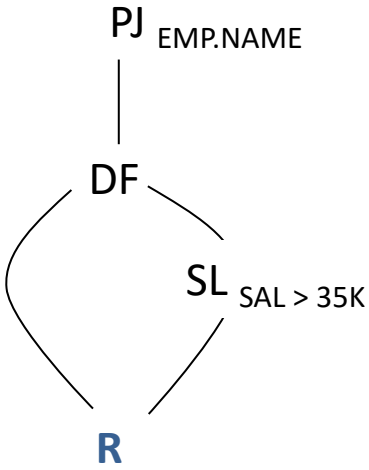
Finding Common Sub-expression



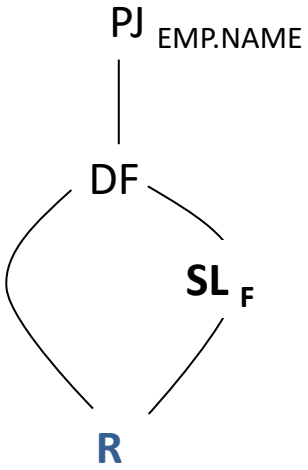
Removing Common Sub-expression



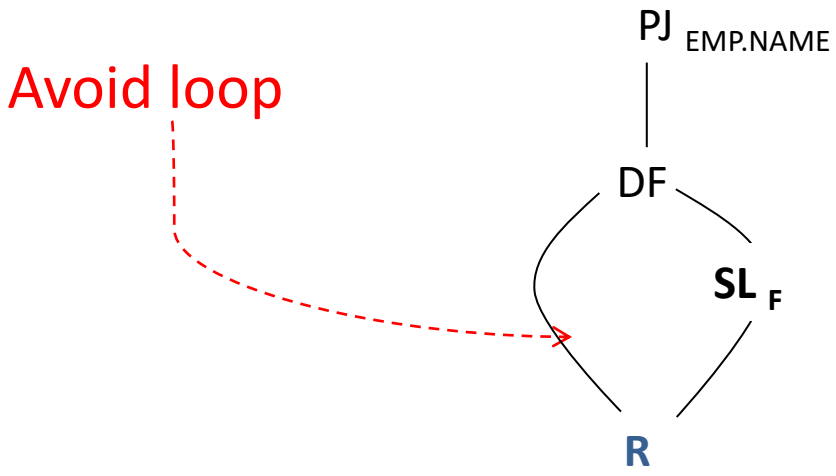
Removing Common Sub-expression



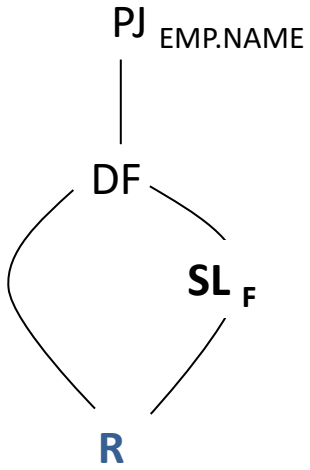
Removing Common Sub-expression



Removing Common Sub-expression



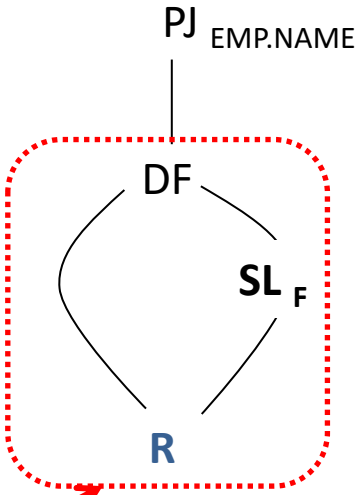
Removing Common Sub-expression



Properties

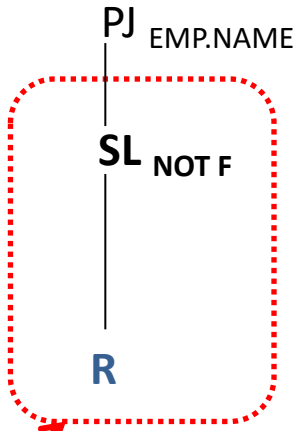
- R NJN R $\leftrightarrow$ R
- R UN R $\leftrightarrow$ R
- R DF R $\leftrightarrow$ 0
- R NJN SL_F R $\leftrightarrow$ SL_F R
- R UN SL_F R $\leftrightarrow$ R
- R DF SL_F R $\leftrightarrow$ $SL_{NOT\ F}$ R
- (SL_{F1} R) NJN (SL_{F2} R) $\leftrightarrow$ $SL_{F1\ AND\ F2}$ R
- (SL_{F1} R) UN (SL_{F2} R) $\leftrightarrow$ $SL_{F1\ OR\ F2}$ R
- (SL_{F1} R) DF (SL_{F2} R) $\leftrightarrow$ $SL_{F1\ AND\ NOT\ F2}$ R

Removing Common Sub-expression



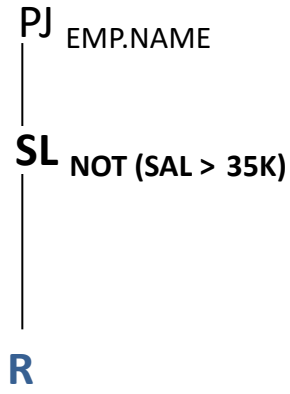
- $R\ DF\ SL_F\ R \leftrightarrow SL_{NOT\ F}\ R$

Removing Common Sub-expression



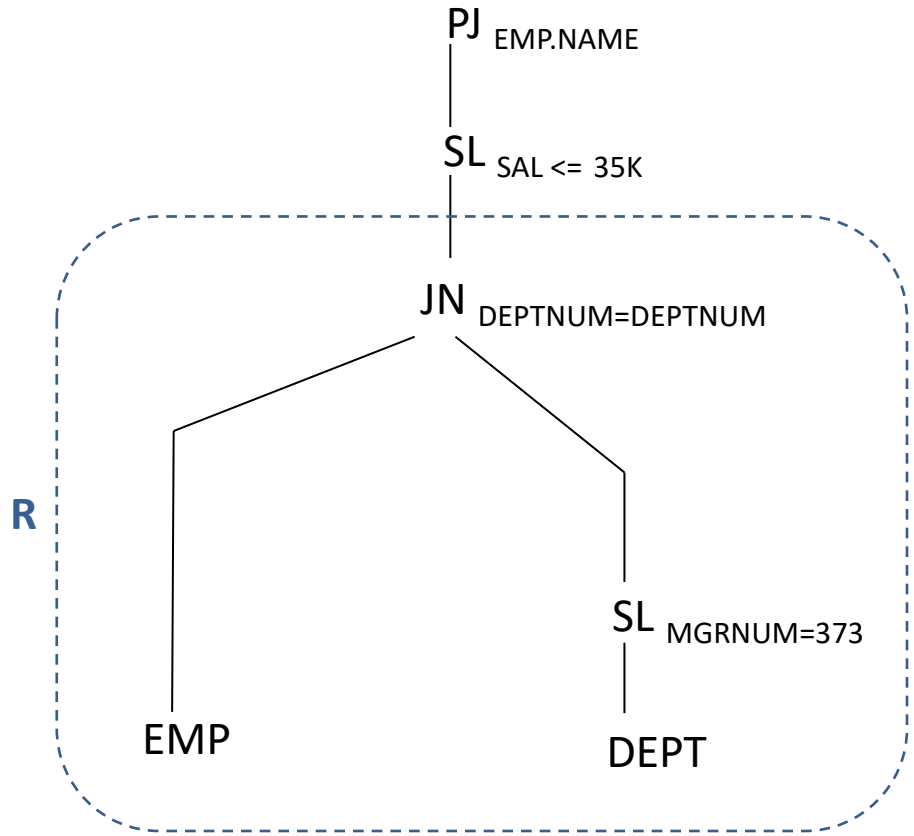
- $R\ DF\ SL_F\ R \leftrightarrow SL_{NOT\ F}\ R$

Removing Common Sub-expression



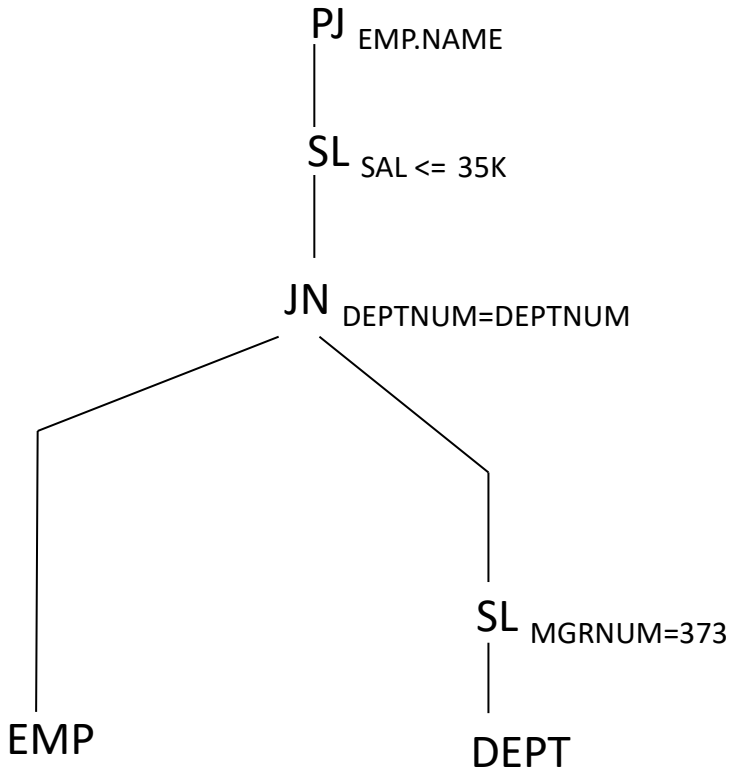
Removing Common Sub-expression

Removing Common Sub-expression



Simplification

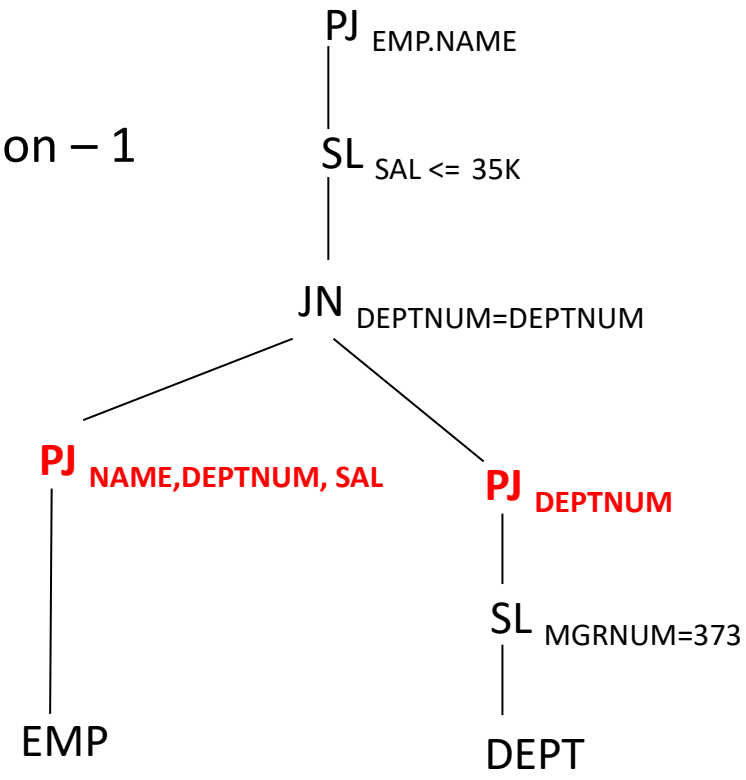
- Now apply criterion – 1 and 2



Simplification

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

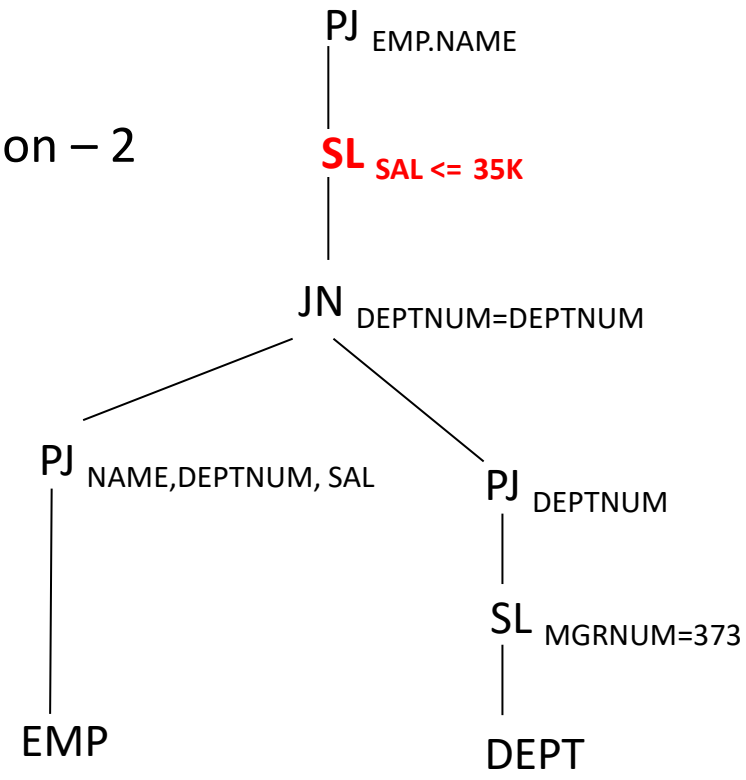
- Applying Criterion – 1



Simplification

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

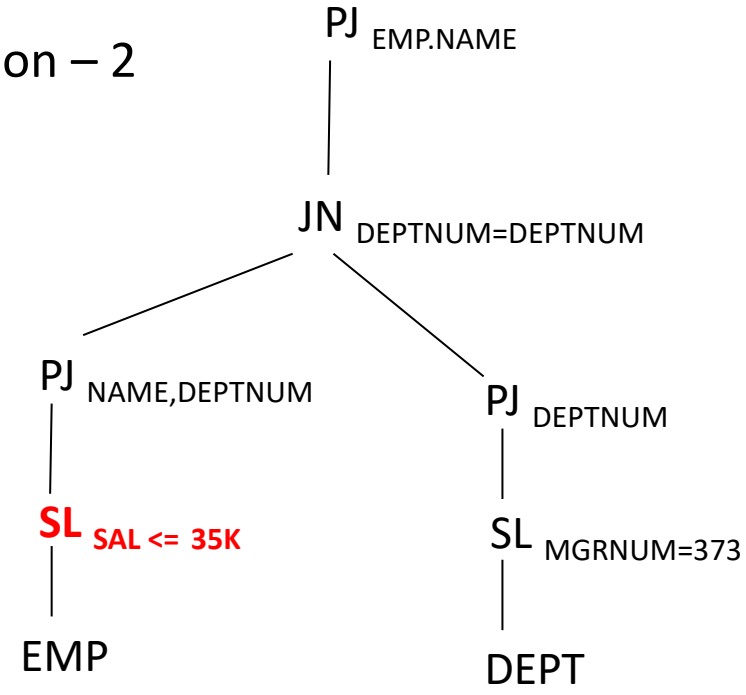
- Applying Criterion – 2



Simplification

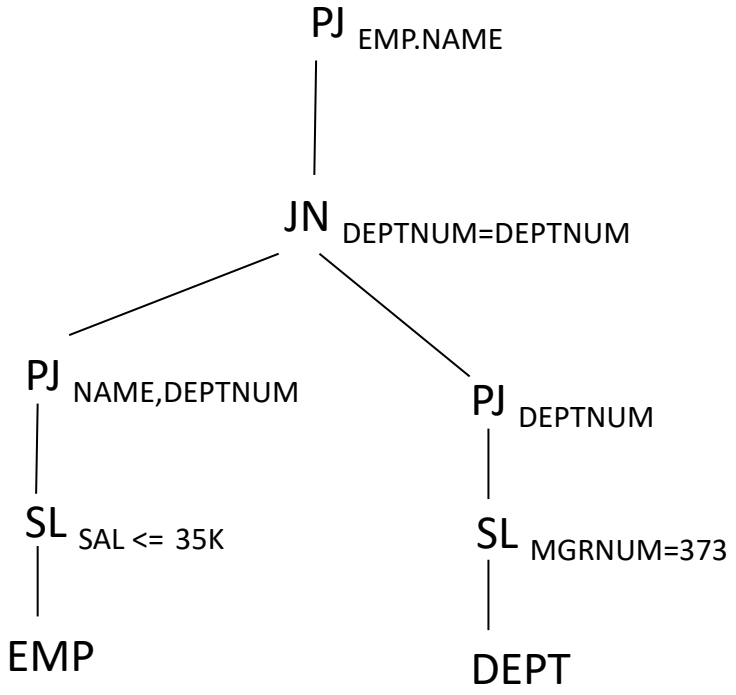
- *EMP (EMPNUM, NAME, SAL, TAX, MGRNUM, DEPTNUM)*
- *DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

- Applying Criterion – 2



Transformed Query

$\mathbf{Q_T}$: $PJ_{EMP.NAME}$ (($PJ_{NAME,DEPTNUM}$ $SL_{SAL<=35K}$ *EMP*) $JN_{DEPTNUM=DEPTNUM}$ ($PJ_{DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*))



Transformed Query

Output:

Q_T: $PJ_{EMP.NAME}$ (($PJ_{NAME,DEPTNUM}$ $SL_{SAL<=35K}$ *EMP*) $JN_{DEPTNUM=DEPTNUM}$ ($PJ_{DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*))

Input:

Q: $PJ_{EMP.NAME}$ ((*EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*) DF ($SL_{SAL > 35K}$ *EMP* $JN_{DEPTNUM=DEPTNUM}$ $SL_{MGRNUM=373}$ *DEPT*))

$$Q \leftrightarrow Q_T$$

Summary

Equivalence Query transformation steps:

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 $T_{transformed}$.
5. Write the query $Q_{transformed}$ from $T_{transformed}$.

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

Transforming Global Queries into Fragment Queries

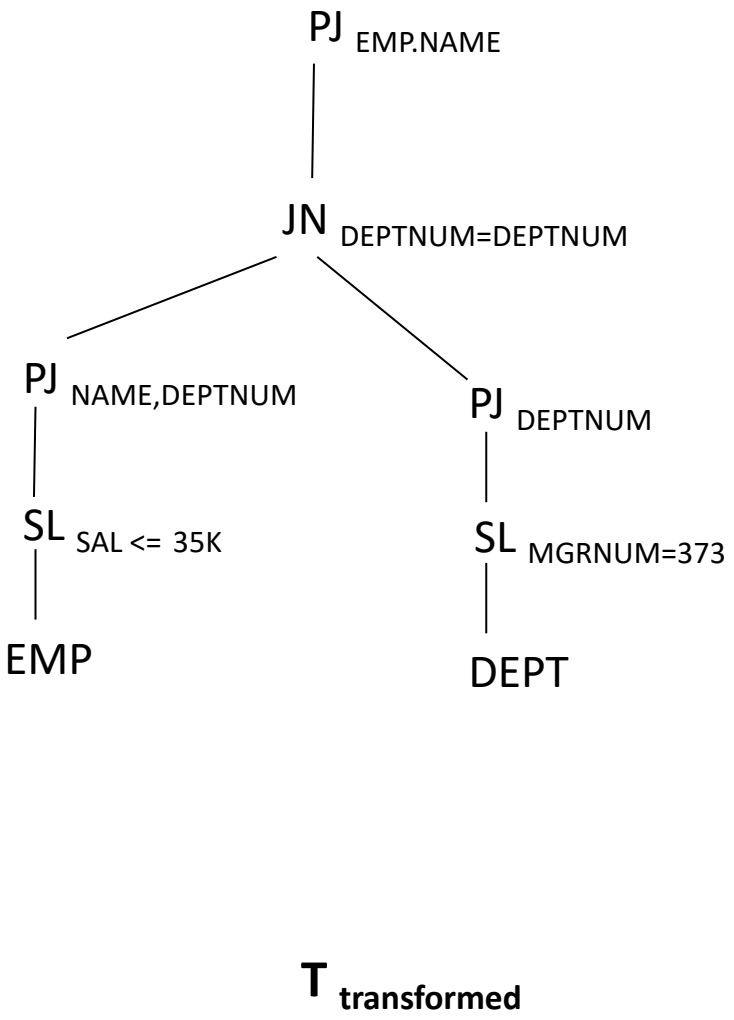
Why transformation?

- Previous $Q_{transformed}$ only works on global relations (i.e. DEPT) , but what about the fragments (i.e. $DEPT_1$, $DEPT_2$)?
- Need to map the query over the global schema ($Q_{transformed}$) to a query over the fragmentation schema ($Q_{fragments}$).

Why transformation?

- Previous $Q_{transformed}$ only works on global relations (i.e. DEPT) , but what about the fragments (i.e. $DEPT_1$, $DEPT_2$)?
- Need to map the query over the global schema ($Q_{transformed}$) to a query over the fragmentation schema ($Q_{fragments}$).
  - We take the final tree ($T_{tranformed}$) from the previous steps and transform it, so that it works on the fragments.

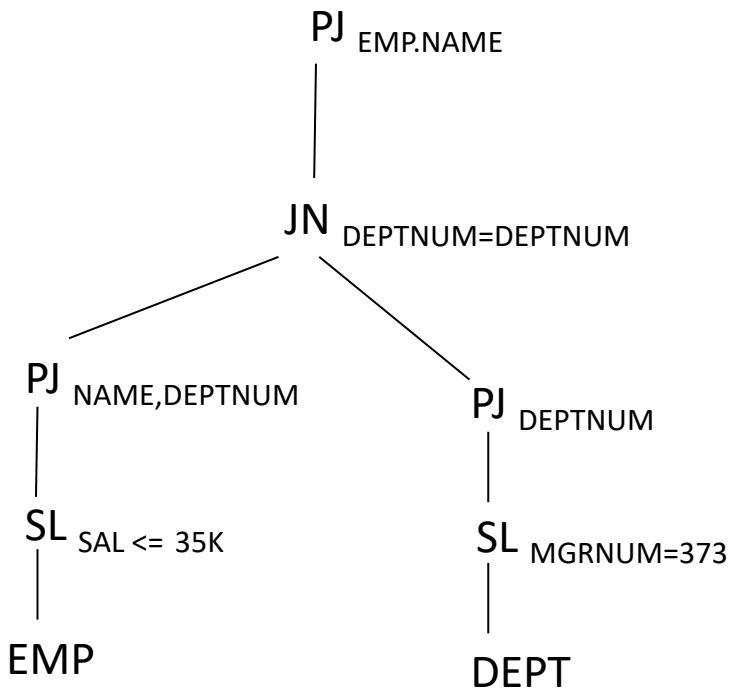
Canonical Expression of a Fragment Query



T transformed

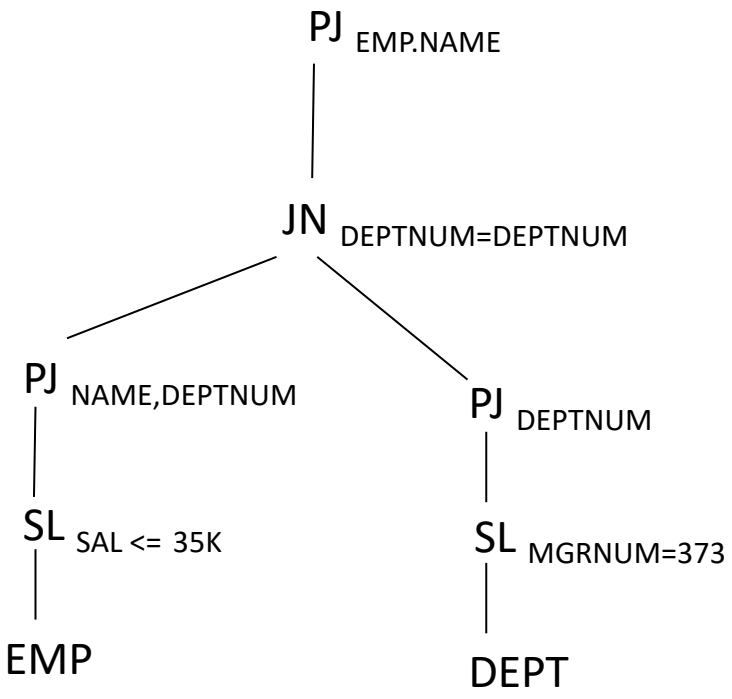
Canonical Expression of a Fragment Query

- Say, DEPT has 2 horizontal fragments: $DEPT_1$ and $DEPT_2$.
- How to convert this tree so that the leaves becomes the fragment?



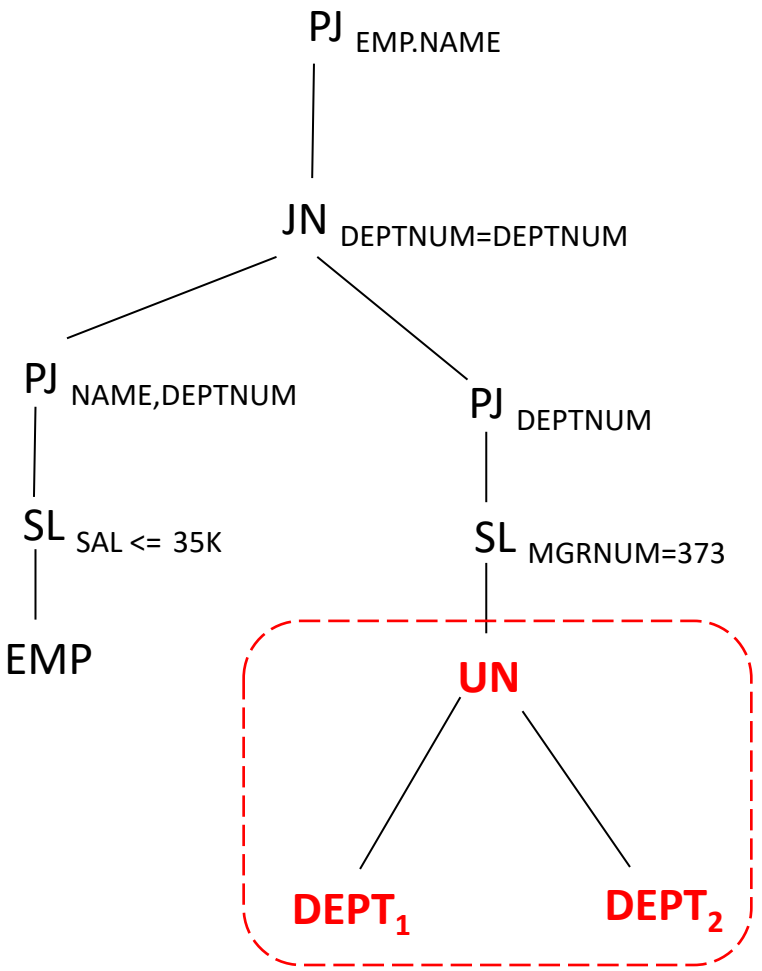
Canonical Expression of a Fragment Query

- Say, 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 expression.
  - Replace the leaf with the sub-tree of the reconstruction expression.



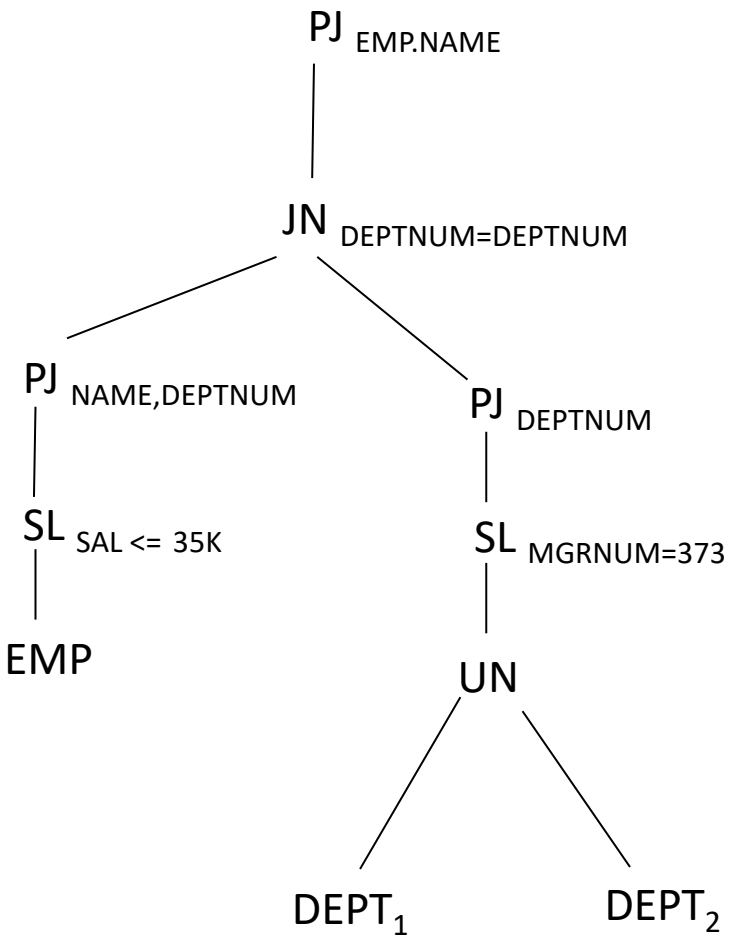
Canonical Expression of a Fragment Query

- Say, 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 expression.



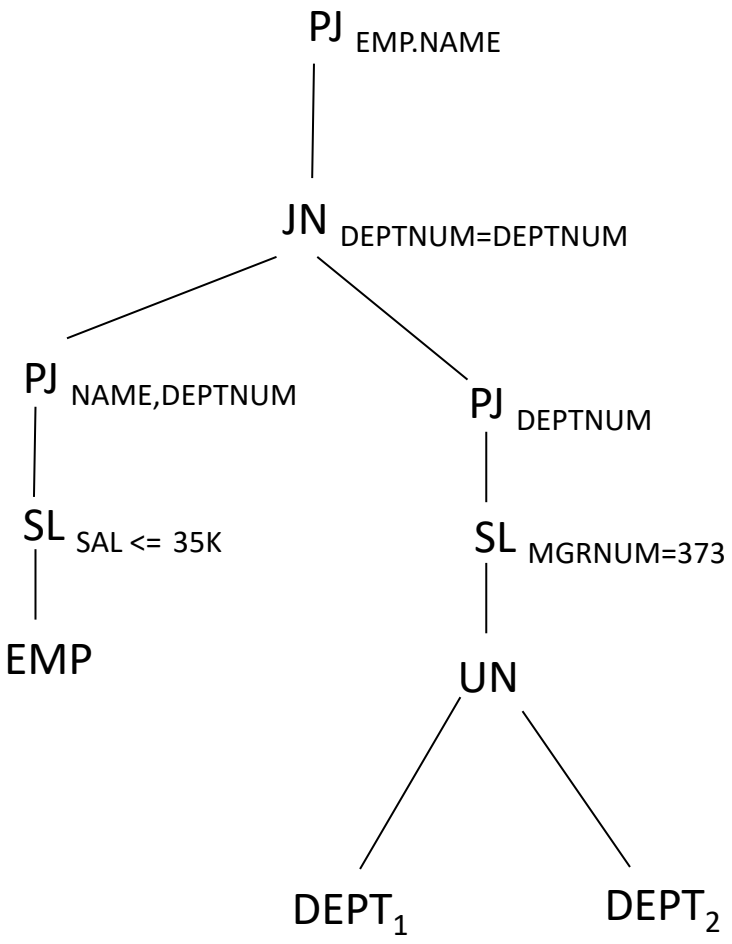
Canonical Expression of a Fragment Query

- Do you think it is still ok?

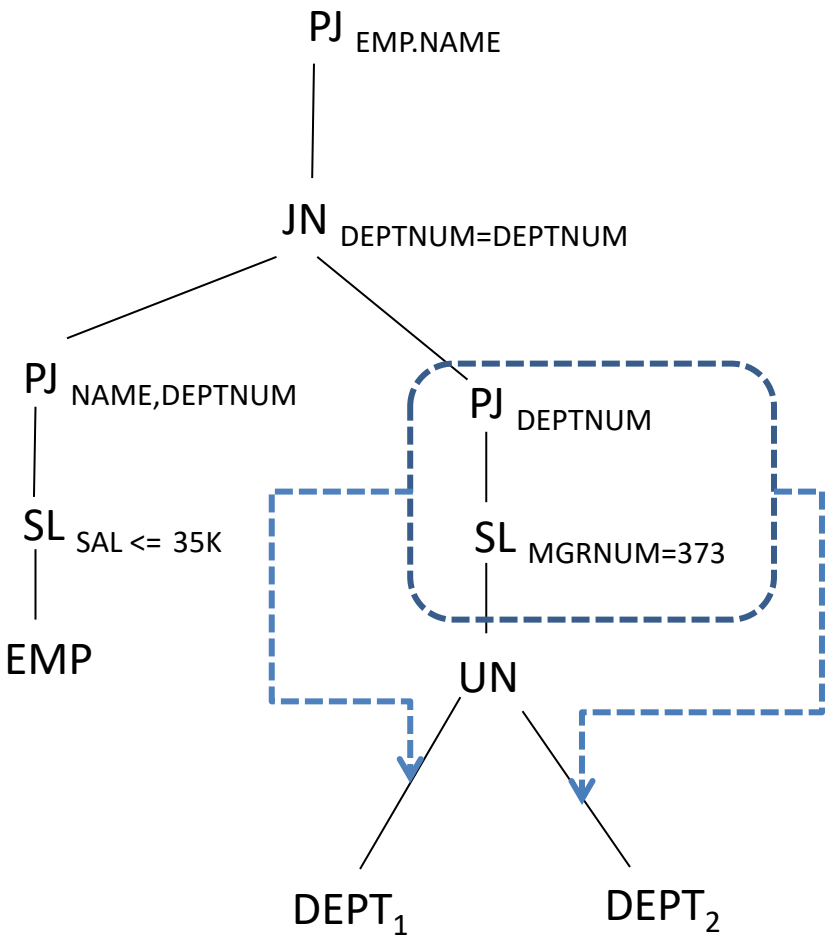


Practice Session

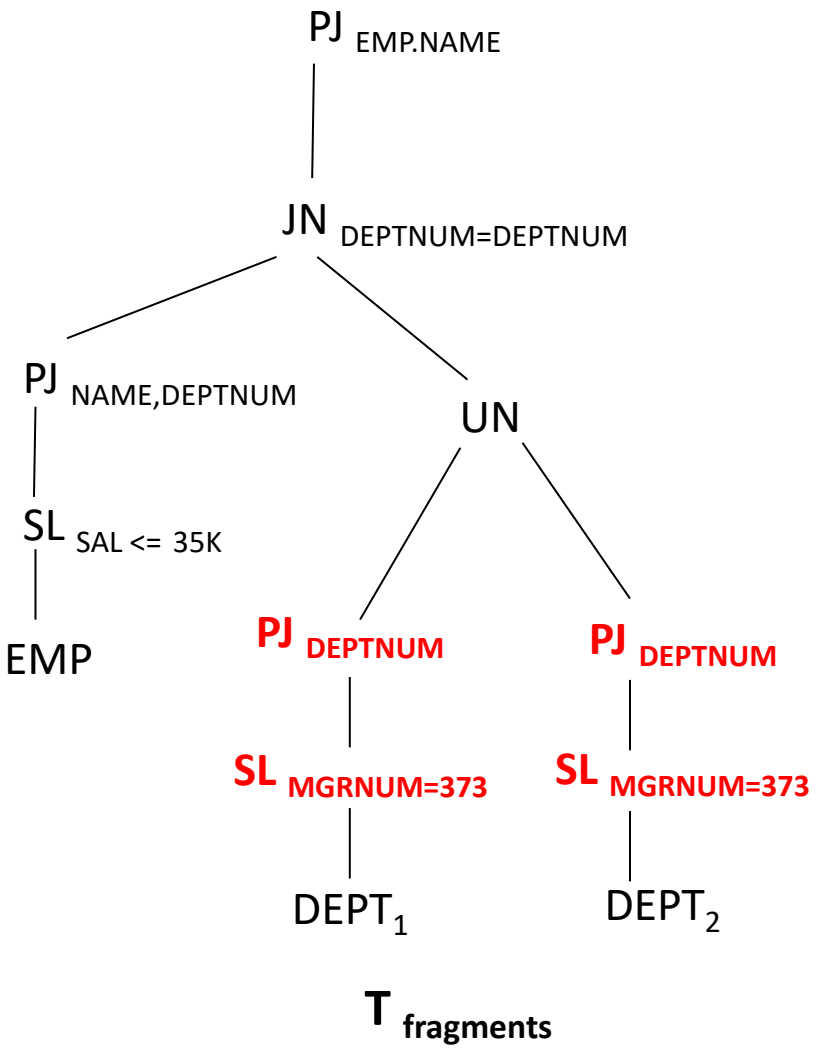
- Apply criterion – 1 and 2



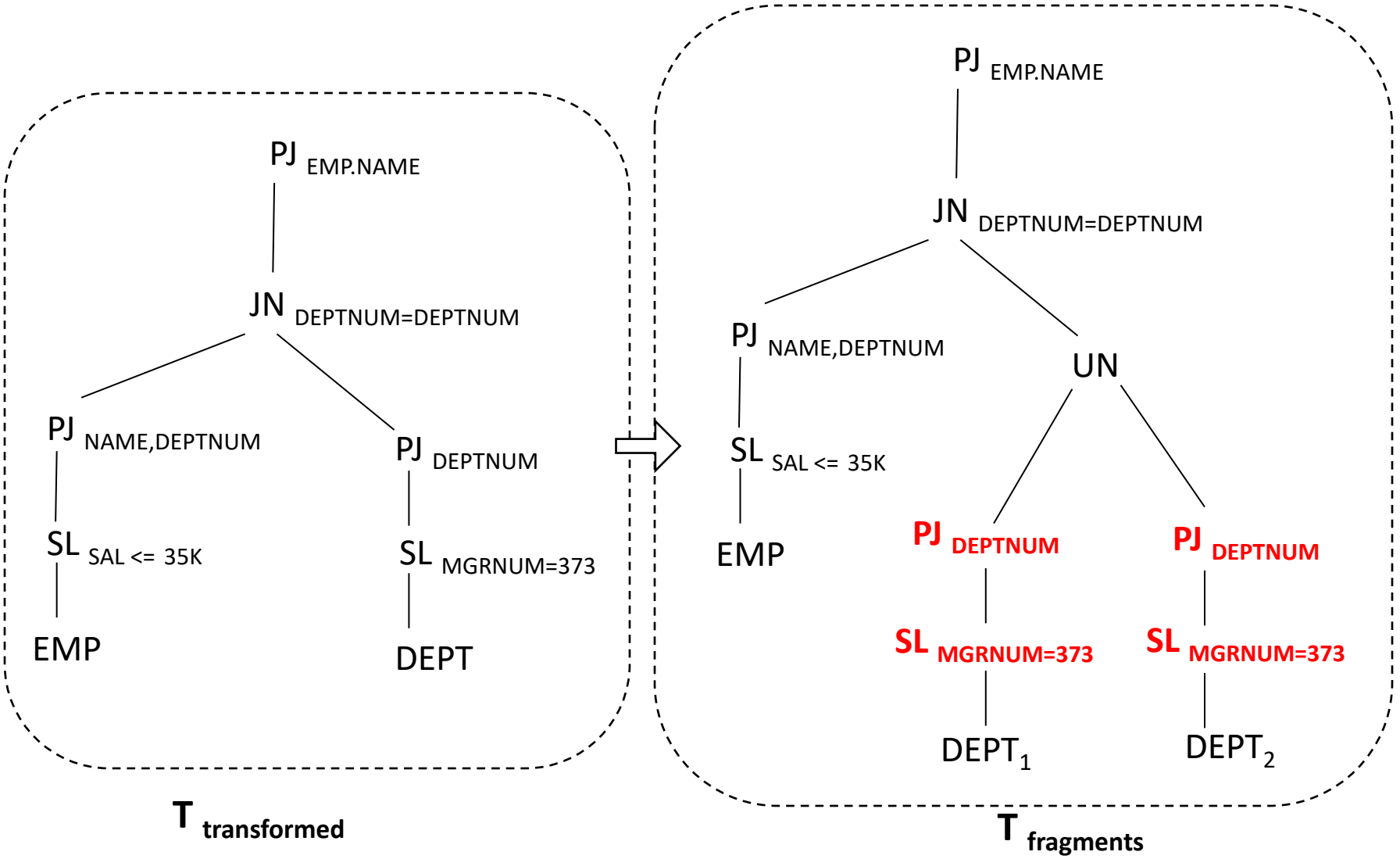
Practice Session (answer)



Practice Session (answer)

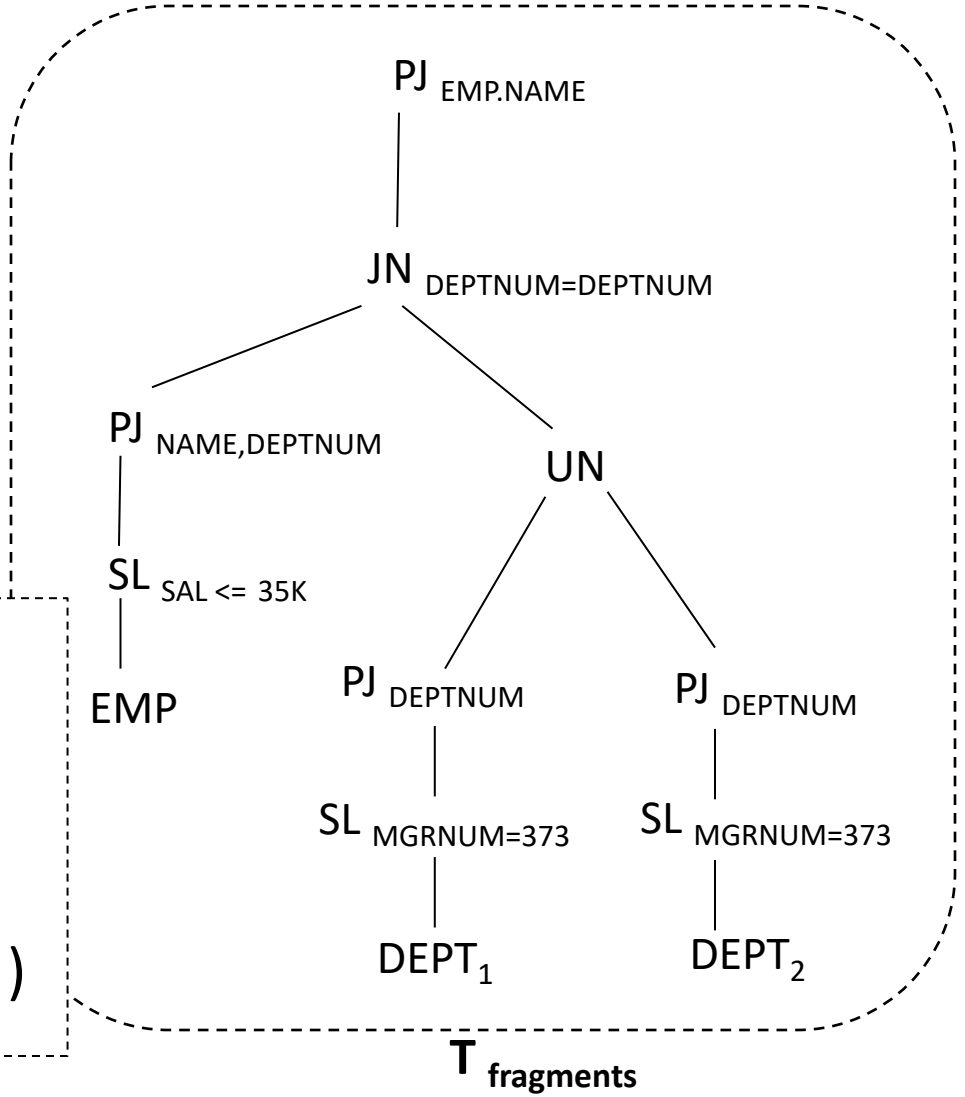


Practice Session (answer)



Practice Session

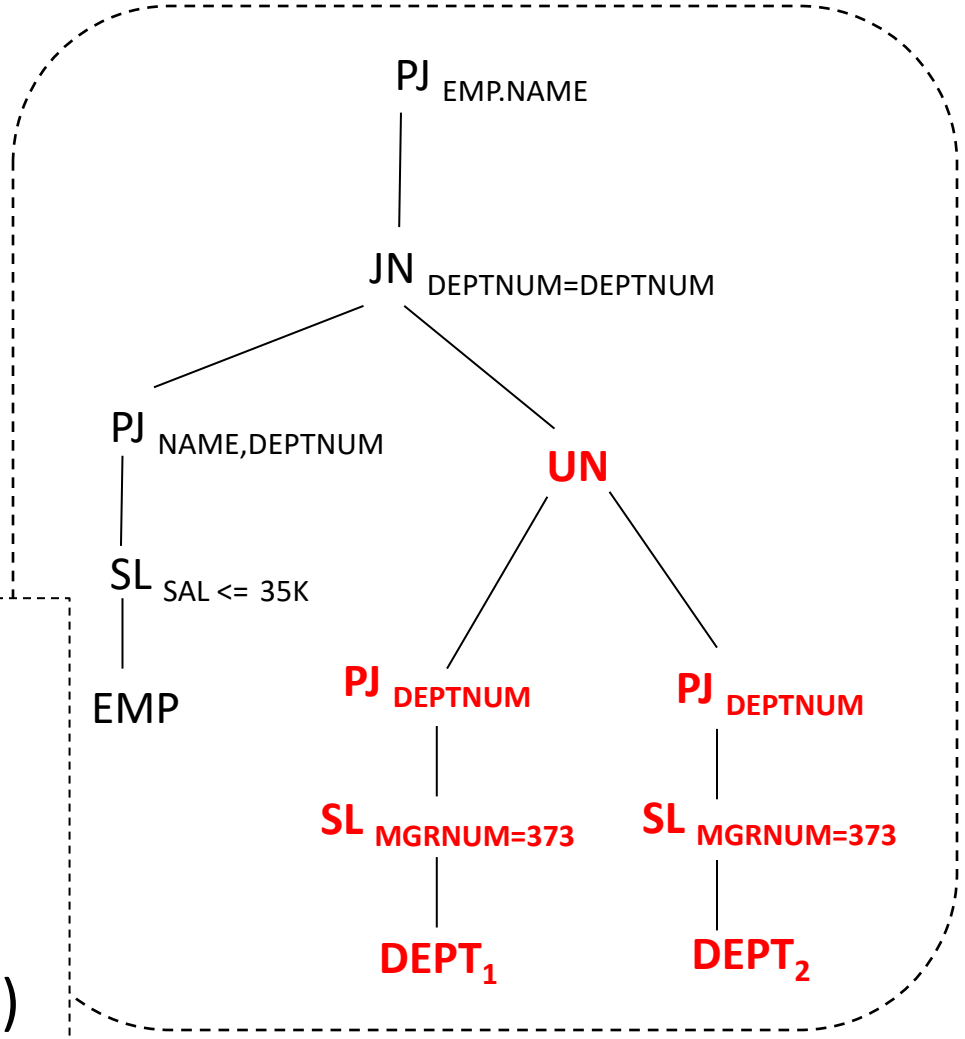
- Write the equivalent query $Q_{fragments}$ from $T_{transformed}$.



- $Q_{fragments}$:

$PJ_{EMP.NAME}((PJ_{NAME,DEPTNUM}$ $SL_{SAL<=35K}$ *EMP*) .....
.....
.....)

Practice Session (answer)



- **Q fragments**:

$PJ_{EMP.NAME}((PJ_{NAME,DEPTNUM}(SL_{SAL<=35K}\ EMP))$
$JN_{DEPTNUM=DEPTNUM}\ (PJ_{DEPTNUM}(SL_{MGRNUM=373}\ DEPT_1)\ UN\ PJ_{DEPTNUM}(SL_{MGRNUM=373}\ DEPT_2))$

Global query to Fragment Query (summary)

- Q_{global}:

$PJ_{EMP.NAME}((EMP\ JN_{DEPTNUM=DEPTNUM}\ SL_{MGRNUM=373}\ DEPT)\ DF\ (SL_{SAL>35K}\ EMP\ JN_{DEPTNUM=DEPTNUM}\ SL_{MGRNUM=373}\ DEPT))$

(non-distributed)

- $Q_{transformed}$:

$PJ_{EMP.NAME}((PJ_{NAME,DEPTNUM}\ SL_{SAL<=35K}\ EMP)\ JN_{DEPTNUM=DEPTNUM}\ (PJ_{DEPTNUM}\ SL_{MGRNUM=373}\ DEPT))$

(distributed)

- $Q_{fragments}$:

$PJ_{EMP.NAME}((PJ_{NAME,DEPTNUM}\ SL_{SAL<=35K}\ EMP)\ JN_{DEPTNUM=DEPTNUM}\ (PJ_{DEPTNUM}\ SL_{MGRNUM=373}\ DEPT_1\ UN\ PJ_{DEPTNUM}\ SL_{MGRNUM=373}\ DEPT_2))$

Steps for Global query to Fragment Query (summary)

ND

1. Generate the equivalent operator tree (T_{global}) for the given query ($\mathbf{Q_{global}}$).
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 $T_{transformed}$.

Steps for Global query to Fragment Query (summary)

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 $T_{transformed}$.

D

5. Apply canonical expression on $T_{transformed}$ to obtain the canonical form $T_{canonical}$.
6. Apply criterion- 1 and 2 on $T_{canonical}$ to obtain $T_{fragments}$.
7. Write the query **$Q_{fragments}$** from $T_{fragments}$.

Practice Problems/ Questions

- Text book: exercise 5.1