CSE 4125: Distributed Database Systems
Chapter – 5

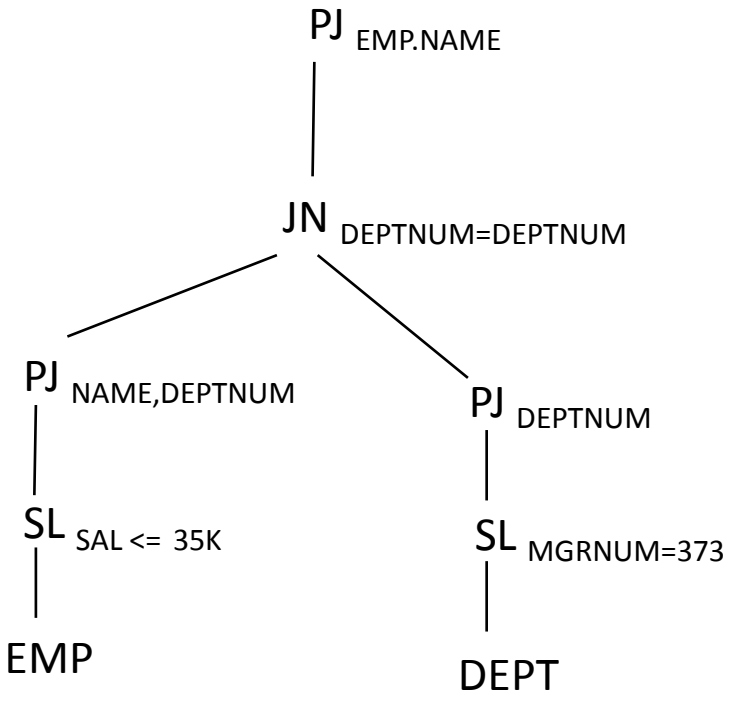
Translation of Global Queries to Fragment Queries.

(part – B)

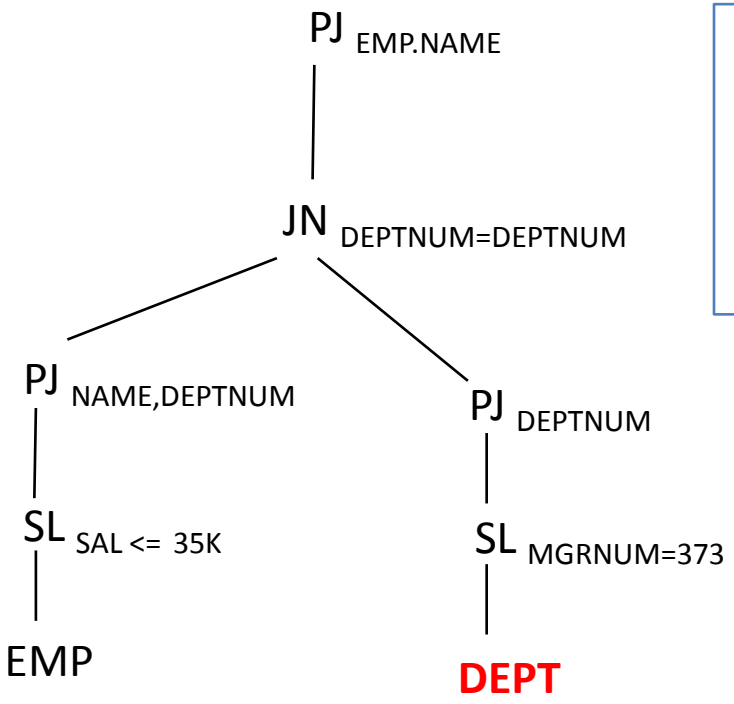
Outline

- Qualified relations.
- Algebra of qualified relations.
- Simplification of horizontally and vertically fragmented relations.
- Simplification of joins between horizontally fragmented relations.
- Criterion – 3, 4 and 5.
- Parametric queries and their simplifications.

Qualified Relation



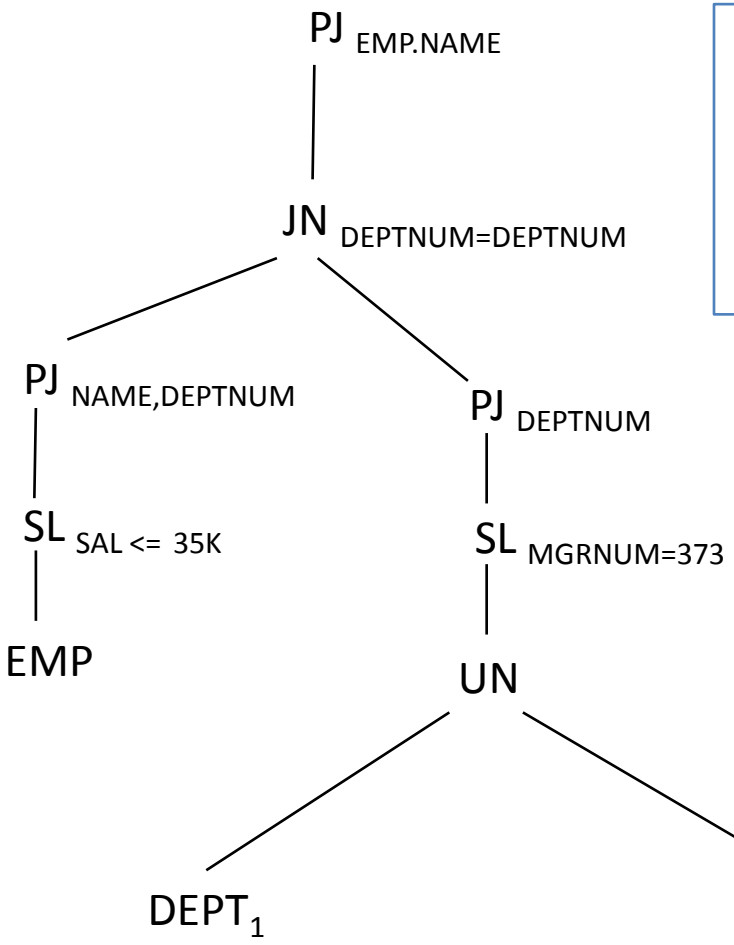
Qualified Relation (contd.)



Say, DEPT has 2 fragments: $DEPT_1$ and $DEPT_2$.

$DEPT_1$: SL $_{deptnum <= 10}$ DEPT
$DEPT_2$: SL $_{deptnum > 10}$ DEPT

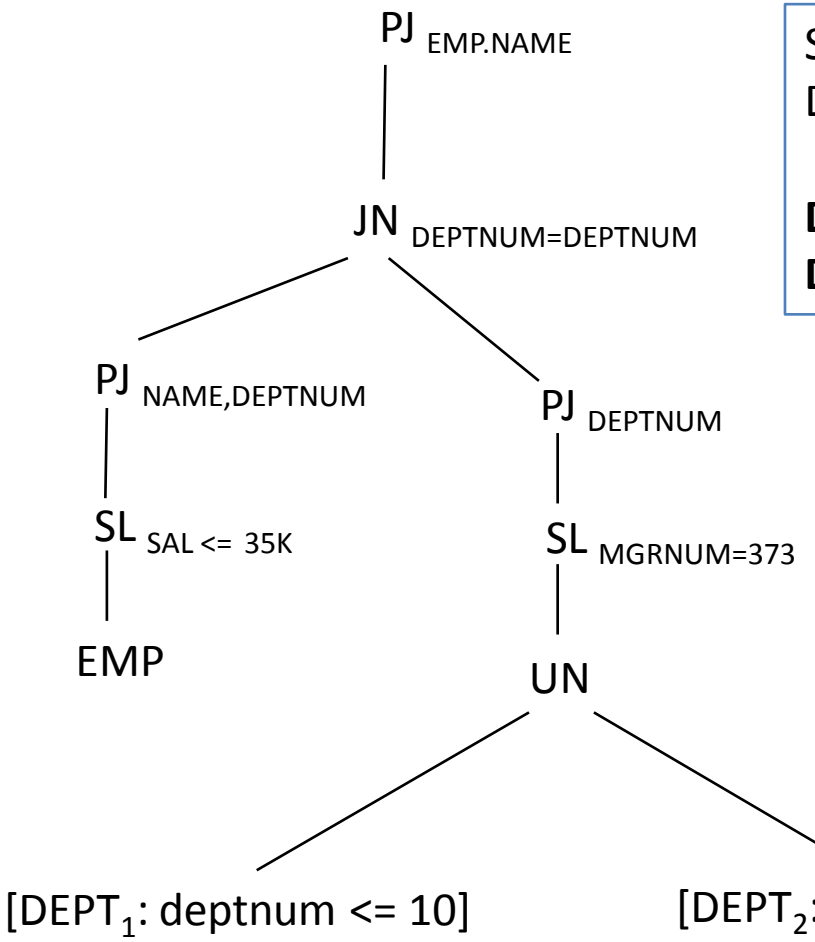
Qualified Relation (contd.)



Say, DEPT has 2 fragments: $DEPT_1$ and $DEPT_2$.

$DEPT_1$: SL $_{deptnum <= 10}$ DEPT
$DEPT_2$: SL $_{deptnum > 10}$ DEPT

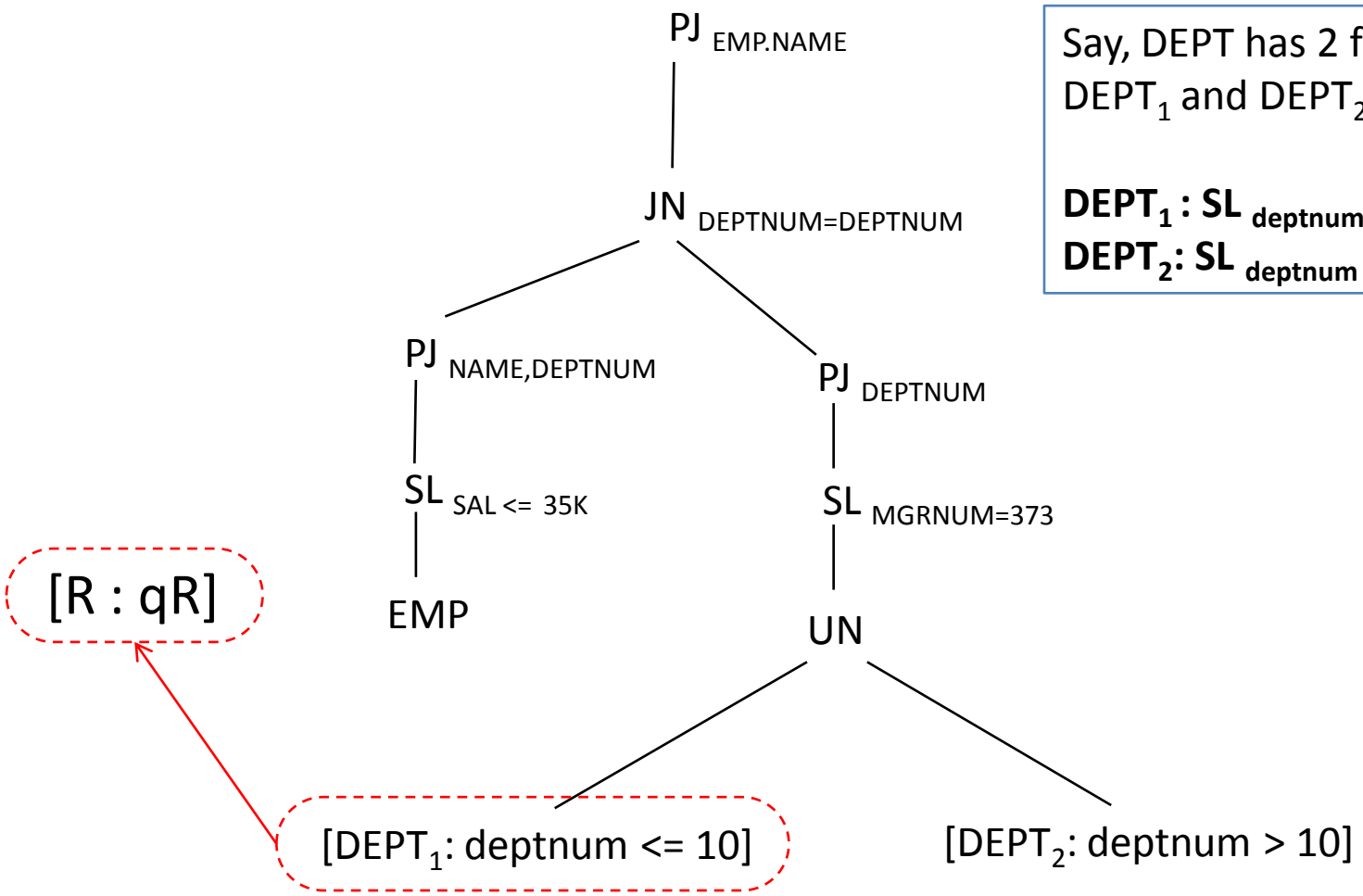
Qualified Relation (contd.)



Say, DEPT has 2 fragments: $DEPT_1$ and $DEPT_2$.

$DEPT_1$: SL $_{deptnum <= 10}$ DEPT
$DEPT_2$: SL $_{deptnum > 10}$ DEPT

Qualified Relation (contd.)



Say, DEPT has 2 fragments: $DEPT_1$ and $DEPT_2$.

$DEPT_1$: SL $_{deptnum <= 10}$ DEPT
$DEPT_2$: SL $_{deptnum > 10}$ DEPT

Qualified Relation (contd.)

- A Qualified relation –
  - Is a relation extended by a qualification.
  - Is denoted as a pair **[R : qR]**, where *R* is a relation called **body** and *qR* is a predicate called **qualification**.
    - Qualifications can be seen as an intentional property possessed by all the tuples of the relation. For example all the tuples in **R** satisfies **qR**.

Algebra of Qualified Relation

- We know **relational algebra** uses **relations** as operands.
  - For example, SL_F **R**
- **Algebra of qualified relation** uses **qualified relations** as operands.
  - For example, SL_F **[R** *: qR*]

Rules of Algebra of Qualified Relation

Rule 1: SL_F **[R : *qR*]** → **[SL_F R : *F and qR*]**

[**ACCOUNT $_1$:** *ID < 5*]

ID	*NAME*	*CITY*
1	a	dhk
2	b	dhk
3	c	ctg
4	d	ctg

$SL_{CITY = dhk}$ [**ACCOUNT $_1$:** *ID < 5*]

ID	*NAME*	*CITY*
1	a	dhk
2	b	dhk

[**$SL_{CITY = dhk}$ ACCOUNT $_1$:** *ID < 5 and CITY = dhk*]

ID	*NAME*	*CITY*
1	a	dhk
2	b	dhk

Rules of Algebra of Qualified Relation

Rule 1: SL_F [R : *qR*] → [SL_F R : *F and qR*]

Rule 2: PJ_A [R : *qR*] → [PJ_A R : *qR*]

Rule 3: [R : *qR*] CP [S : *qS*] → [R CP S: *qR and qS*]

Rule 4: [R : *qR*] DF [S : *qS*] → [R DF S: *qR*]

Rules of Algebra of Qualified Relation (contd.)

Rule 5: [R : *qR*] UN [S : *qS*] → [R UN S: *qR or qS*]

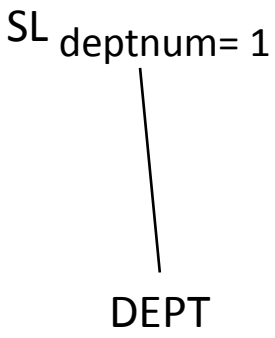
Rule 6: [R : *qR*] JN_F [S : *qS*] → [R JN_F S: *qR and qS and F*]

Rule 7: [R : *qR*] SJ_F [S : *qS*] → [R SJ_F S: *qR and qS and F*]

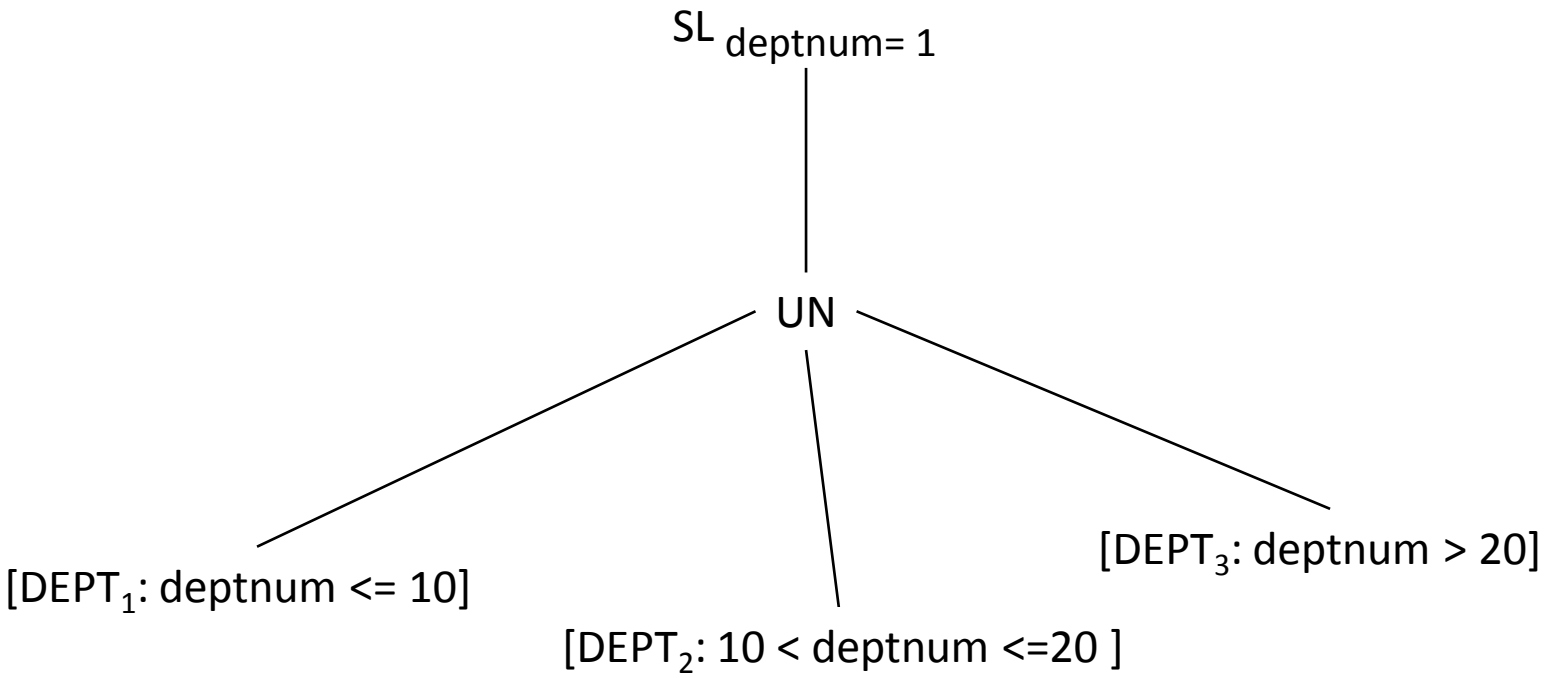
Simplification of Horizontally Fragmented Relations

Example

Q: $\sigma_{\text{deptnum} = 1}$ **DEPT**

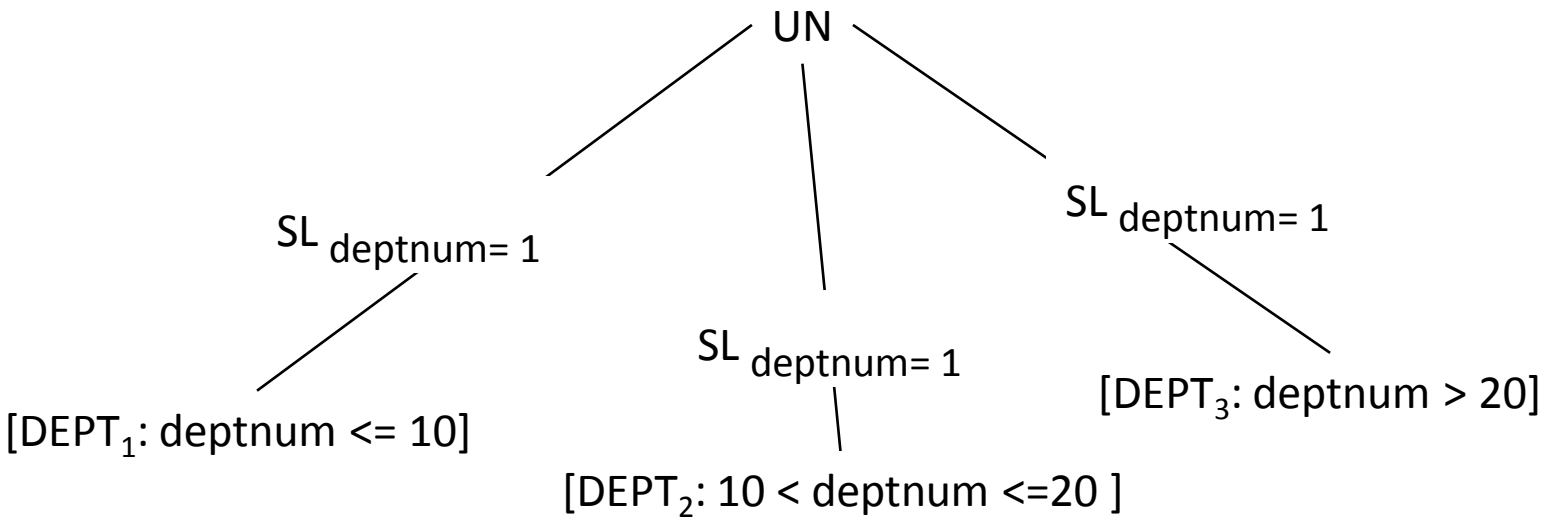


Example (contd.)



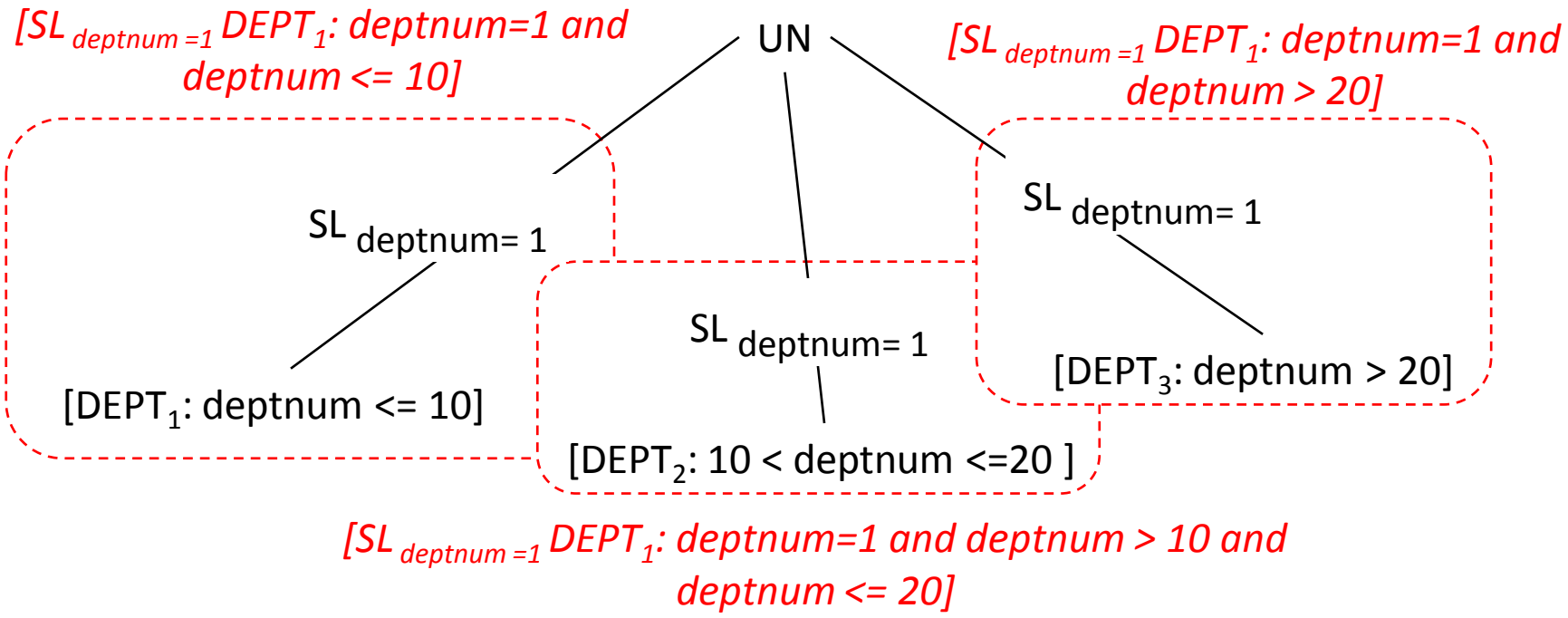
Example (contd.)

- Now, apply algebra of qualified relation.

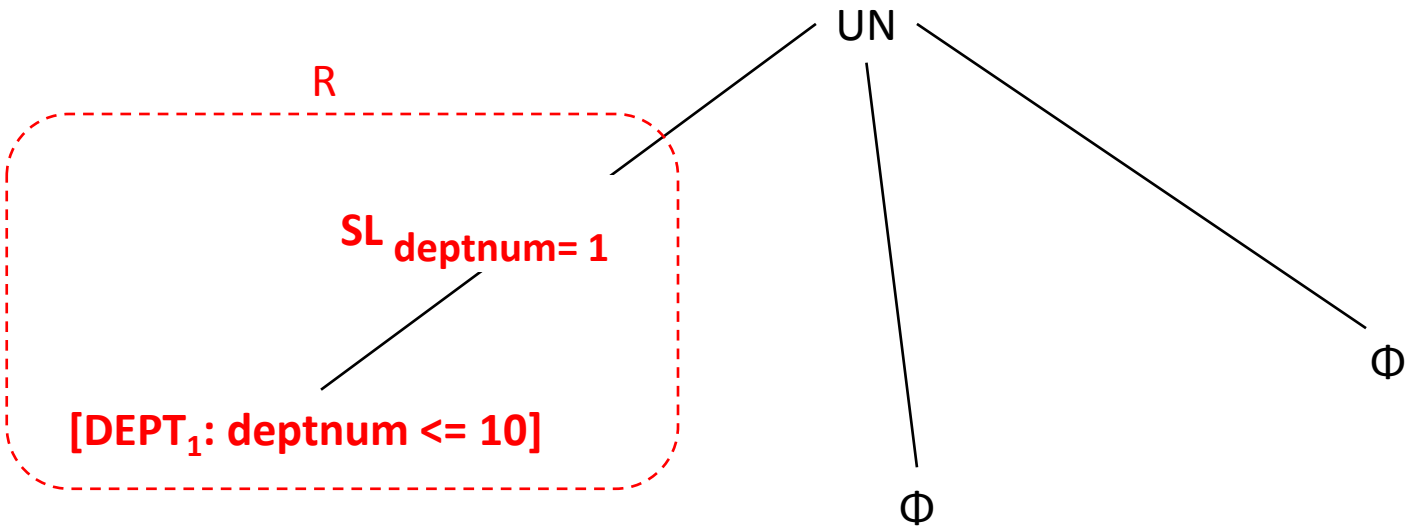


Example (contd.)

- Check if any contradiction.



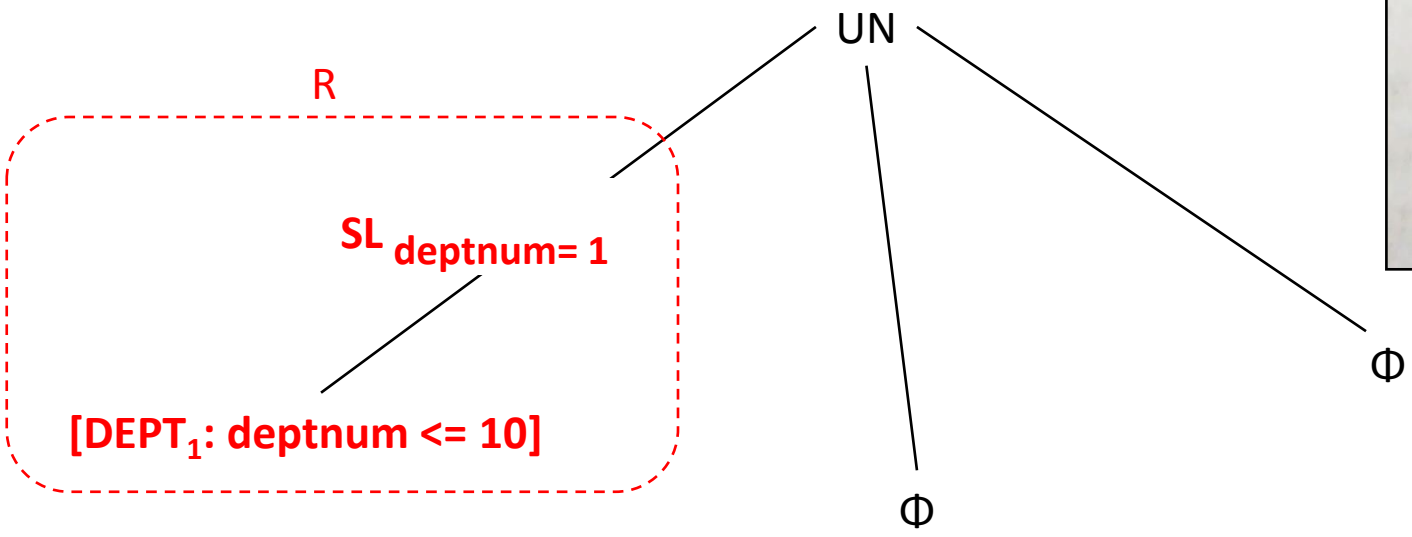
Example (contd.)



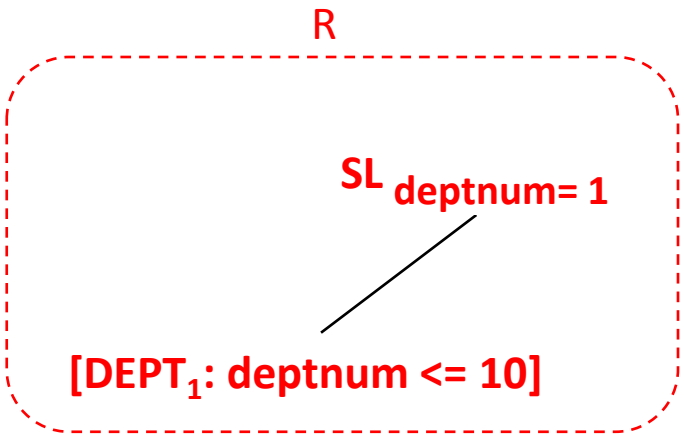
Example (contd.)

- Apply the predefined equivalence transformations.

$$\mathrm{SL}_F(\emptyset) \leftrightarrow \emptyset$$
$$\mathrm{PJ}_A(\emptyset) \leftrightarrow \emptyset$$
$$R\,\mathrm{CP}\,\emptyset \leftrightarrow \emptyset$$
$$R\,\mathrm{UN}\,\emptyset \leftrightarrow R$$
$$R\,\mathrm{DF}\,\emptyset \leftrightarrow R$$
$$\emptyset\,\mathrm{DF}\,R \leftrightarrow \emptyset$$
$$R\,\mathrm{JN}_F\,\emptyset \leftrightarrow \emptyset$$
$$R\,\mathrm{SJ}_F\,\emptyset \leftrightarrow \emptyset$$
$$\emptyset\,\mathrm{SJ}_F\,R \leftrightarrow \emptyset$$



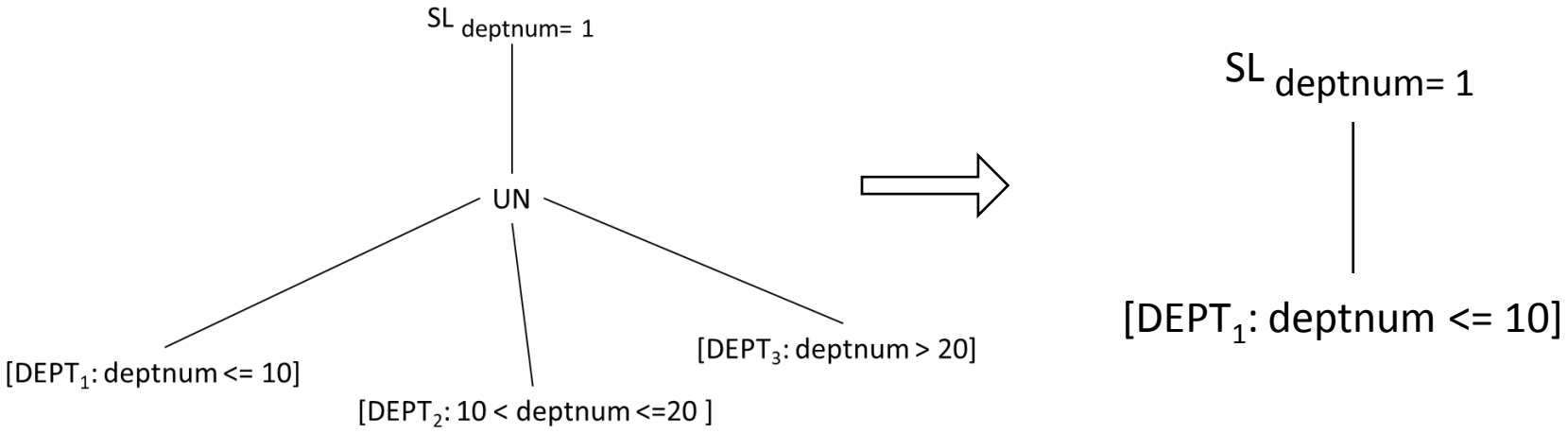
Example (contd.)



Criterion - 3

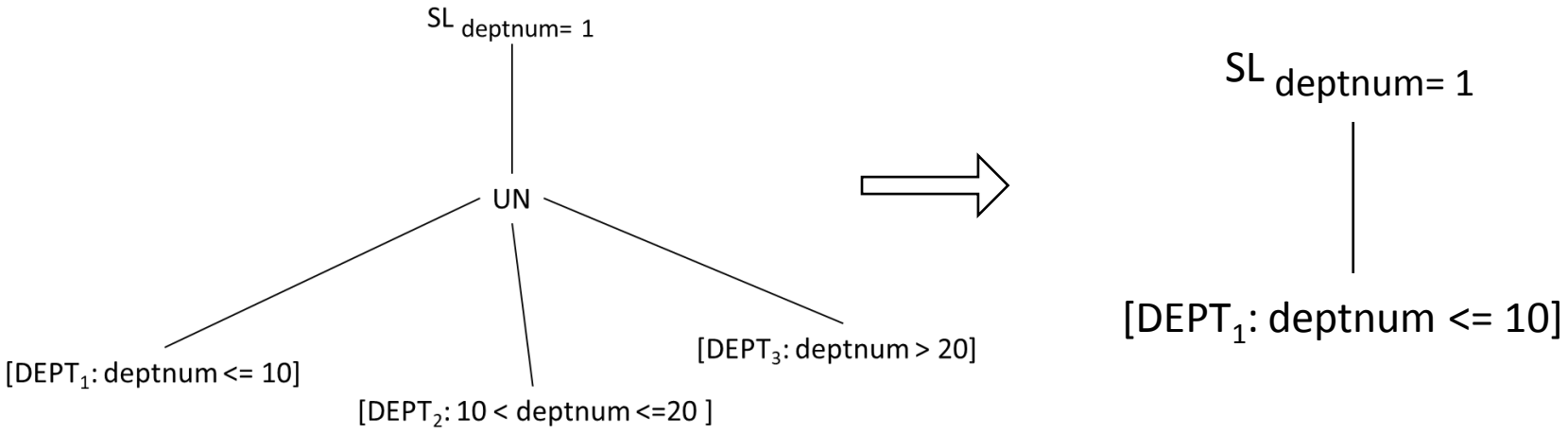
Criterion - 3:

- Push *SL* down, then apply *algebra of qualified relations*.
- Substitute the selection with *empty* if the qualifications *contradicts*.



Criterion - 3

So, from now, you also have to apply **criterion – 3** after applying **criterion – 1** and **criterion – 2**.

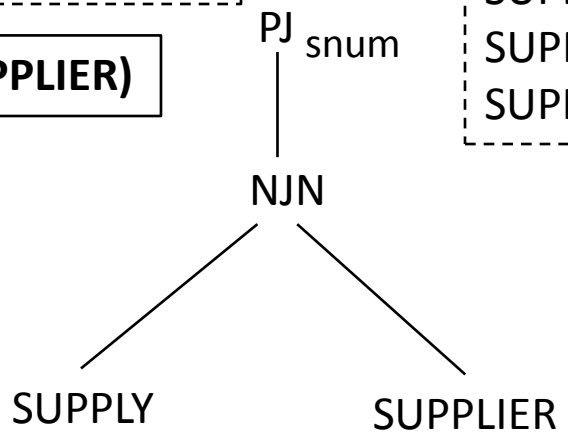


Simplification of Join between Horizontally Fragmented Relations

Example

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

Q: Π_{snum} **(SUPPLY** NJN **SUPPLIER)**



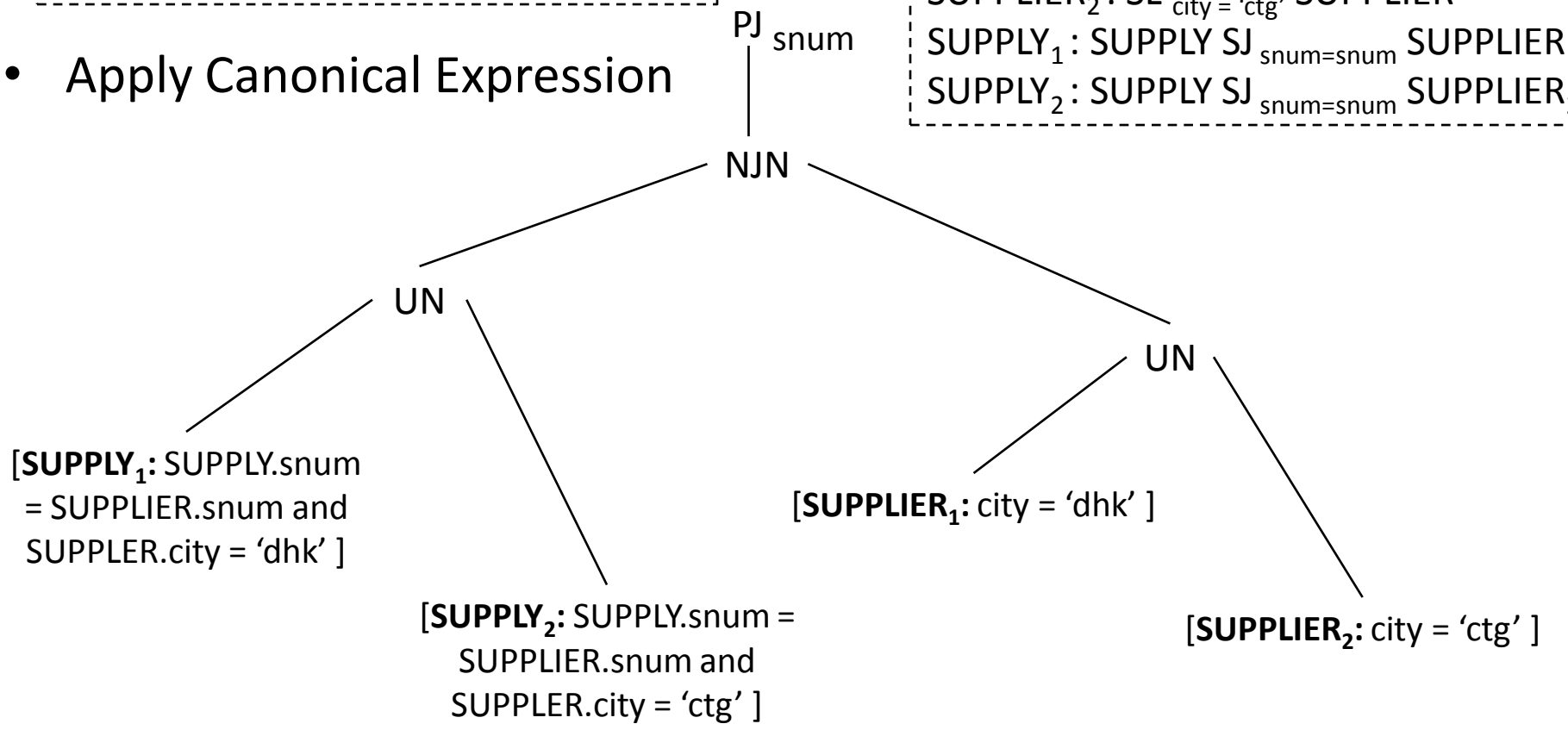
$SUPPLIER_1 : \sigma_{city = 'dhk'} SUPPLIER$
$SUPPLIER_2 : \sigma_{city = 'ctg'} SUPPLIER$
$SUPPLY_1 : SUPPLY \bowtie_{snum=snum} SUPPLIER_1$
$SUPPLY_2 : SUPPLY \bowtie_{snum=snum} SUPPLIER_2$

Example (contd.)

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

- Apply Canonical Expression

$SUPPLIER_1 : SL_{city = 'dhk'}\ SUPPLIER$
$SUPPLIER_2 : SL_{city = 'ctg'}\ SUPPLIER$
$SUPPLY_1 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_1$
$SUPPLY_2 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_2$

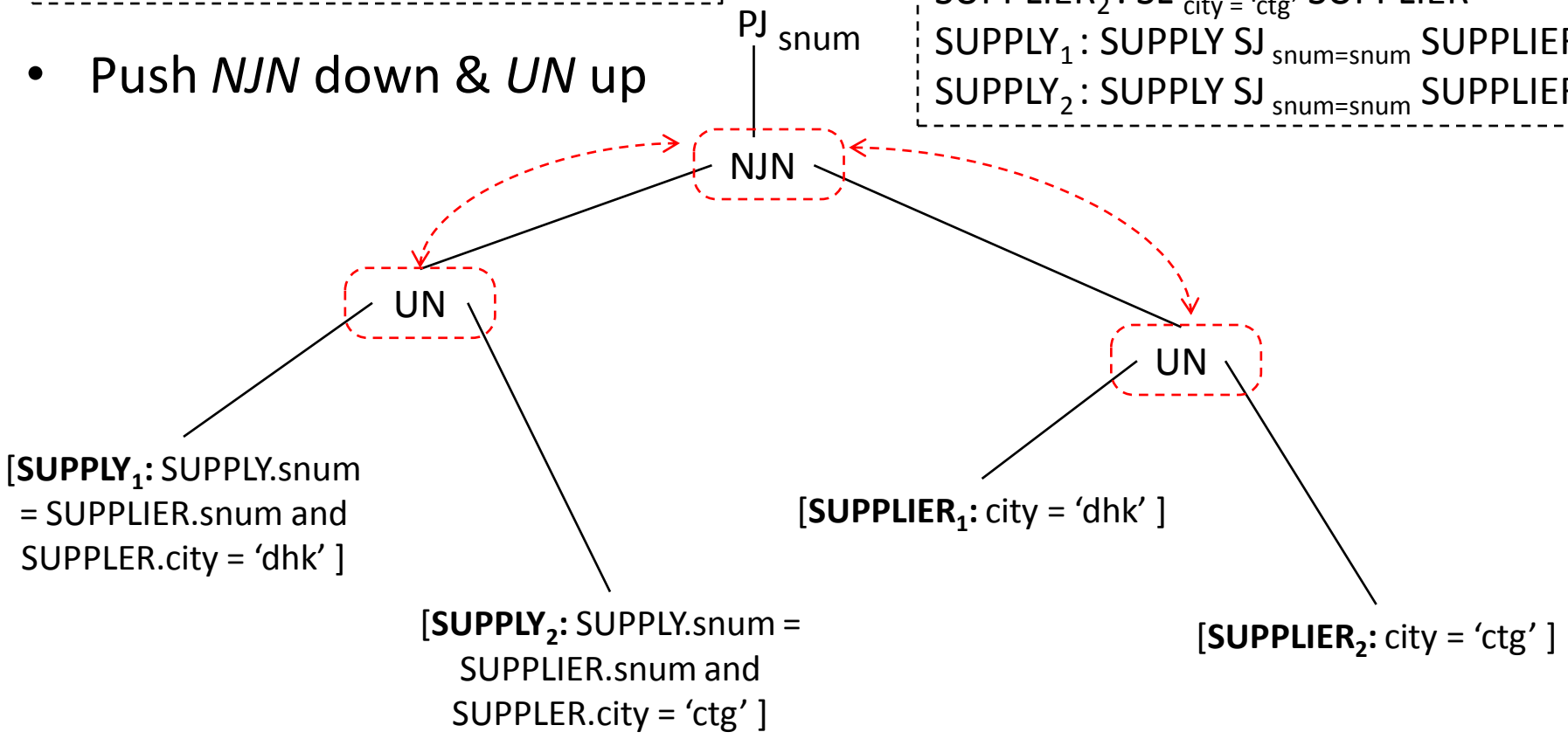


Example (contd.)

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

- Push *NJN* down & *UN* up

$SUPPLIER_1 : SL_{city = 'dhk'}\ SUPPLIER$
$SUPPLIER_2 : SL_{city = 'ctg'}\ SUPPLIER$
$SUPPLY_1 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_1$
$SUPPLY_2 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_2$

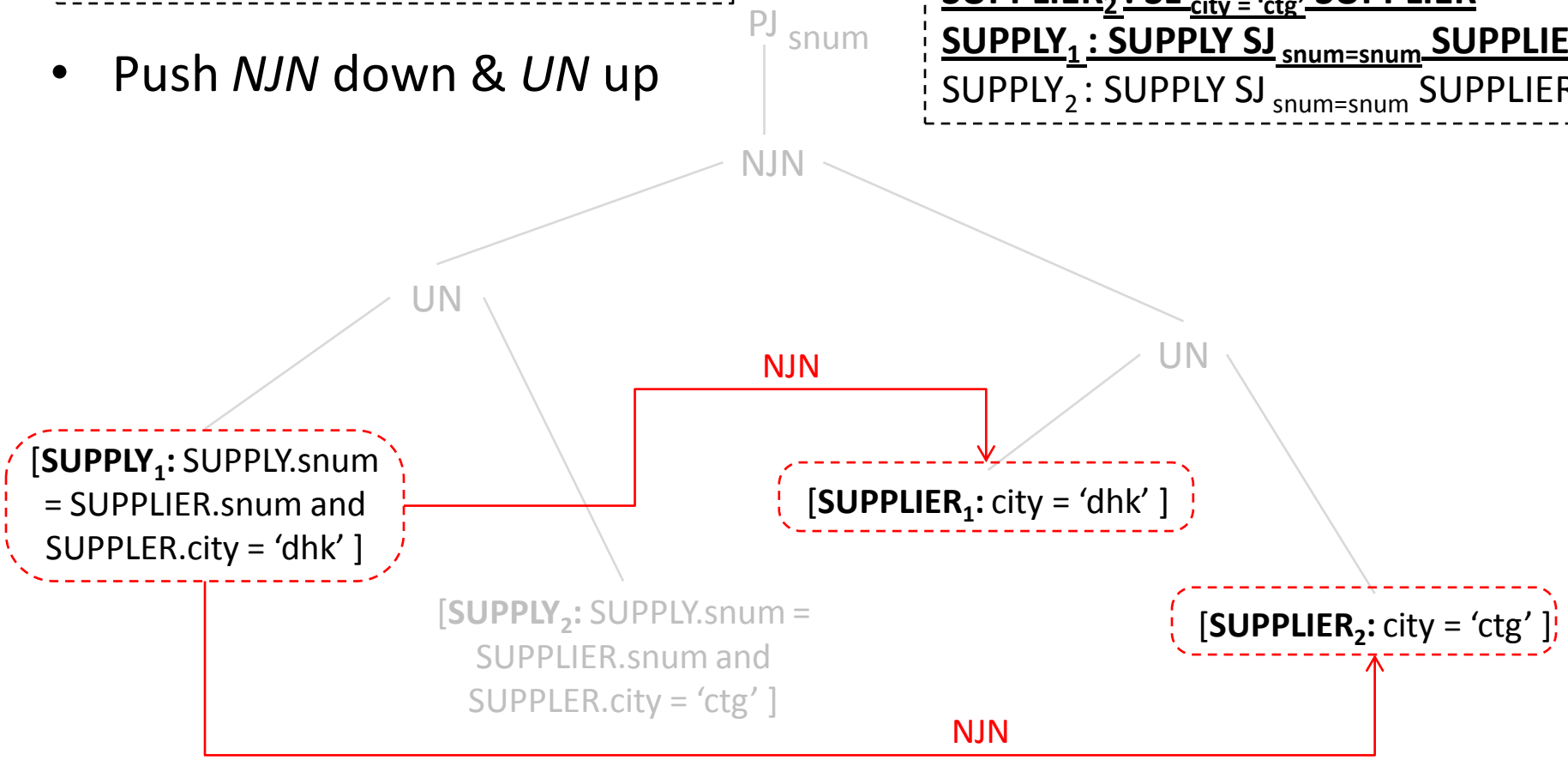


Example (contd.)

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

- Push *NJN* down & *UN* up

$\underline{SUPPLIER_1 : SL_{city = 'dhk'} SUPPLIER}$
$\underline{SUPPLIER_2 : SL_{city = 'ctg'} SUPPLIER}$
$\underline{SUPPLY_1 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_1}$
$SUPPLY_2 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_2$

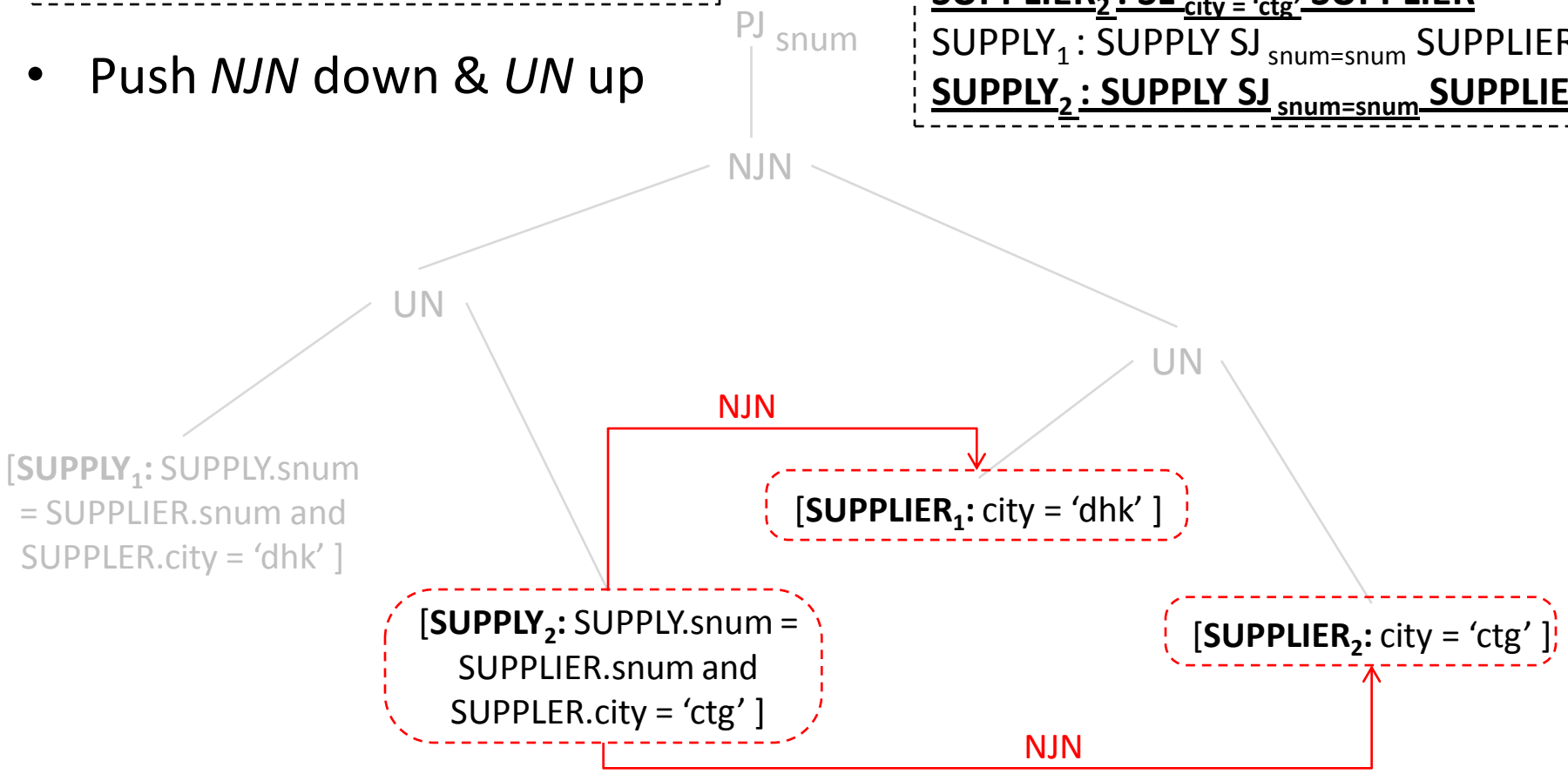


Example (contd.)

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

- Push *NJN* down & *UN* up

$SUPPLIER_1$: $SL_{city = 'dhk'}$ SUPPLIER
$SUPPLIER_2$: $SL_{city = 'ctg'}$ SUPPLIER
$SUPPLY_1$: SUPPLY $SJ_{snum=snum}$ $SUPPLIER_1$
$SUPPLY_2$: SUPPLY $SJ_{snum=snum}$ $SUPPLIER_2$

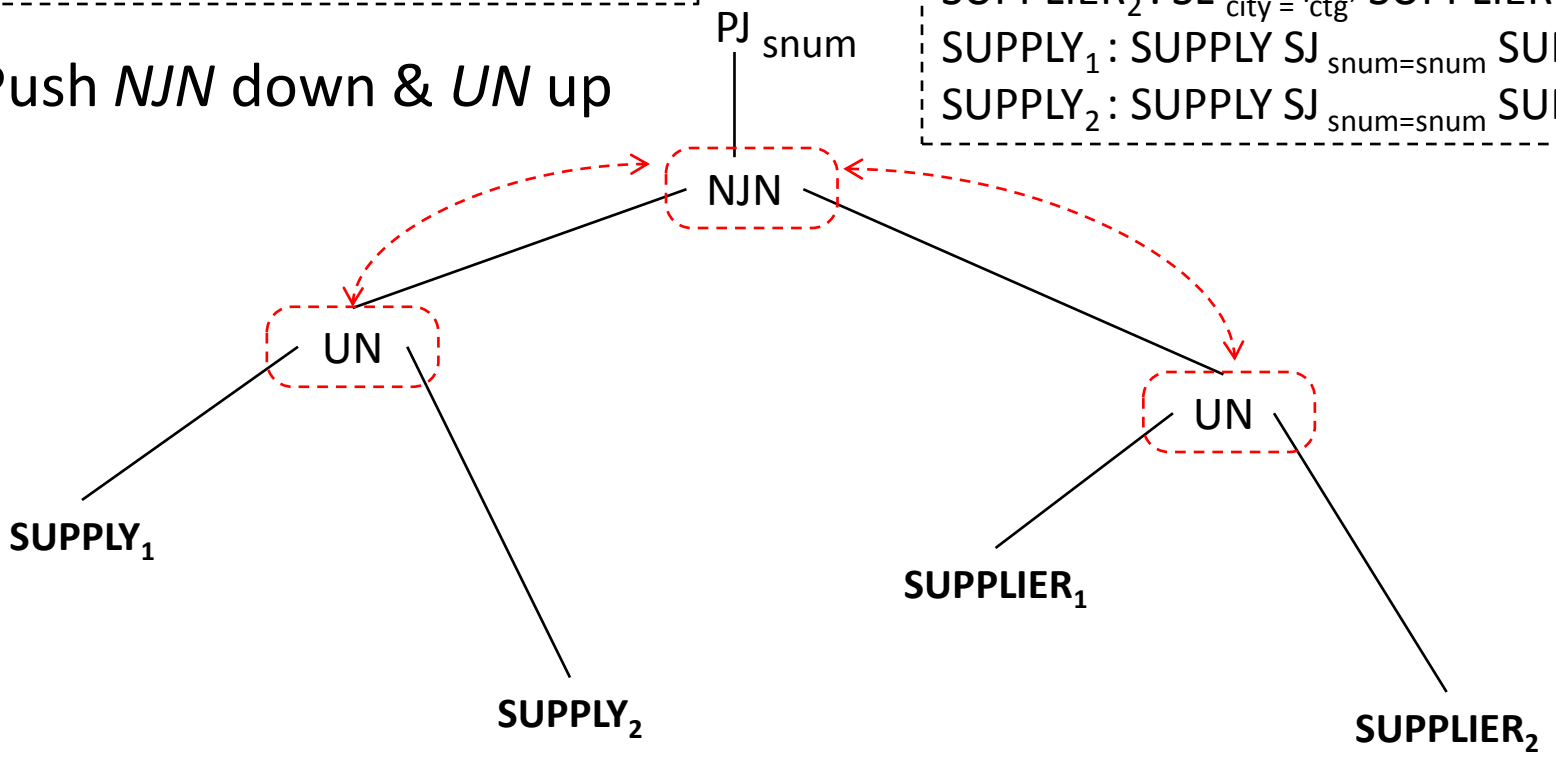


Example (contd.)

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

$SUPPLIER_1 : SL_{city = 'dhk'}\ SUPPLIER$
$SUPPLIER_2 : SL_{city = 'ctg'}\ SUPPLIER$
$SUPPLY_1 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_1$
$SUPPLY_2 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_2$

- Push *NJN* down & *UN* up

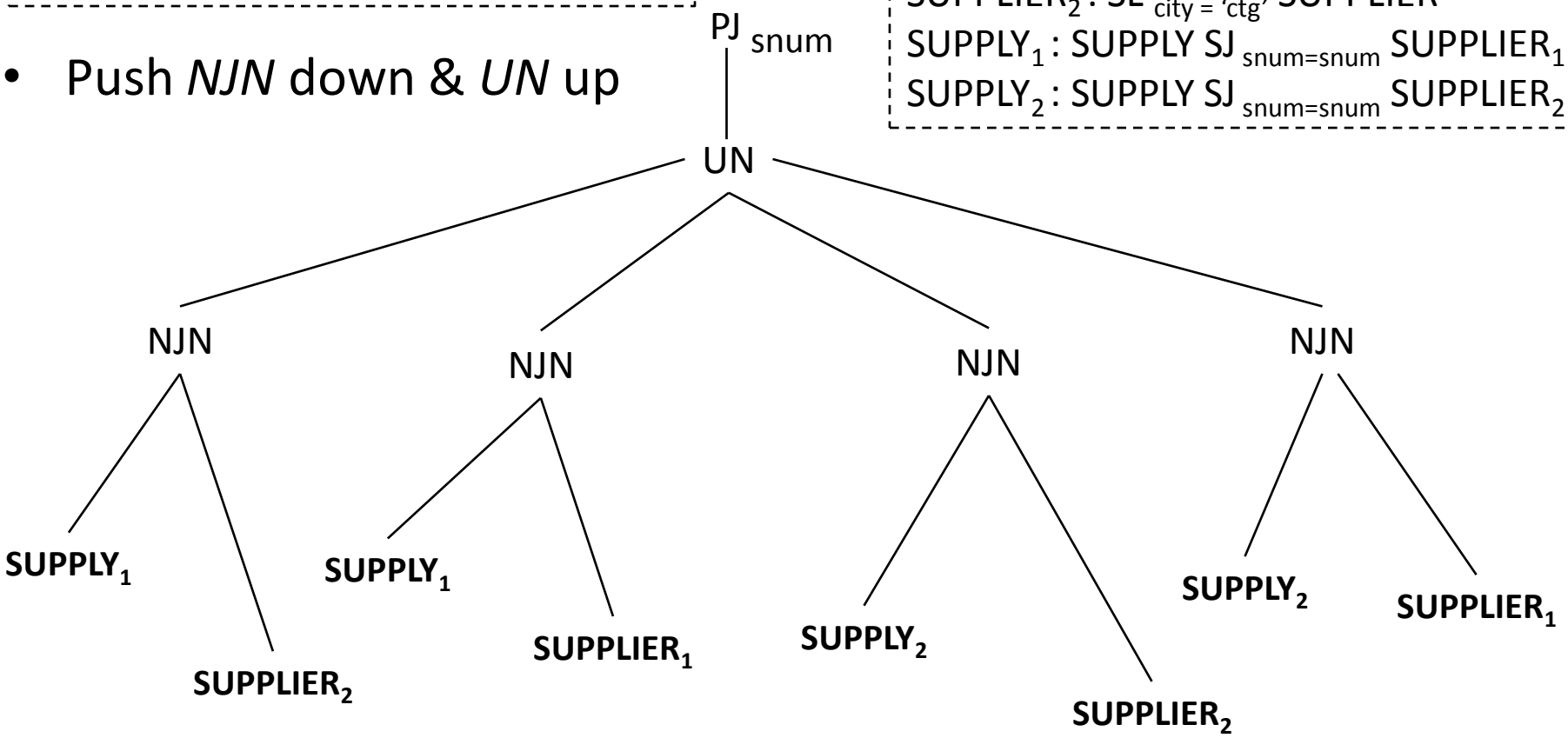


Example (contd.)

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

- Push *NJN* down & *UN* up

$SUPPLIER_1 : SL_{city = 'dhk'} SUPPLIER$
$SUPPLIER_2 : SL_{city = 'ctg'} SUPPLIER$
$SUPPLY_1 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_1$
$SUPPLY_2 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_2$

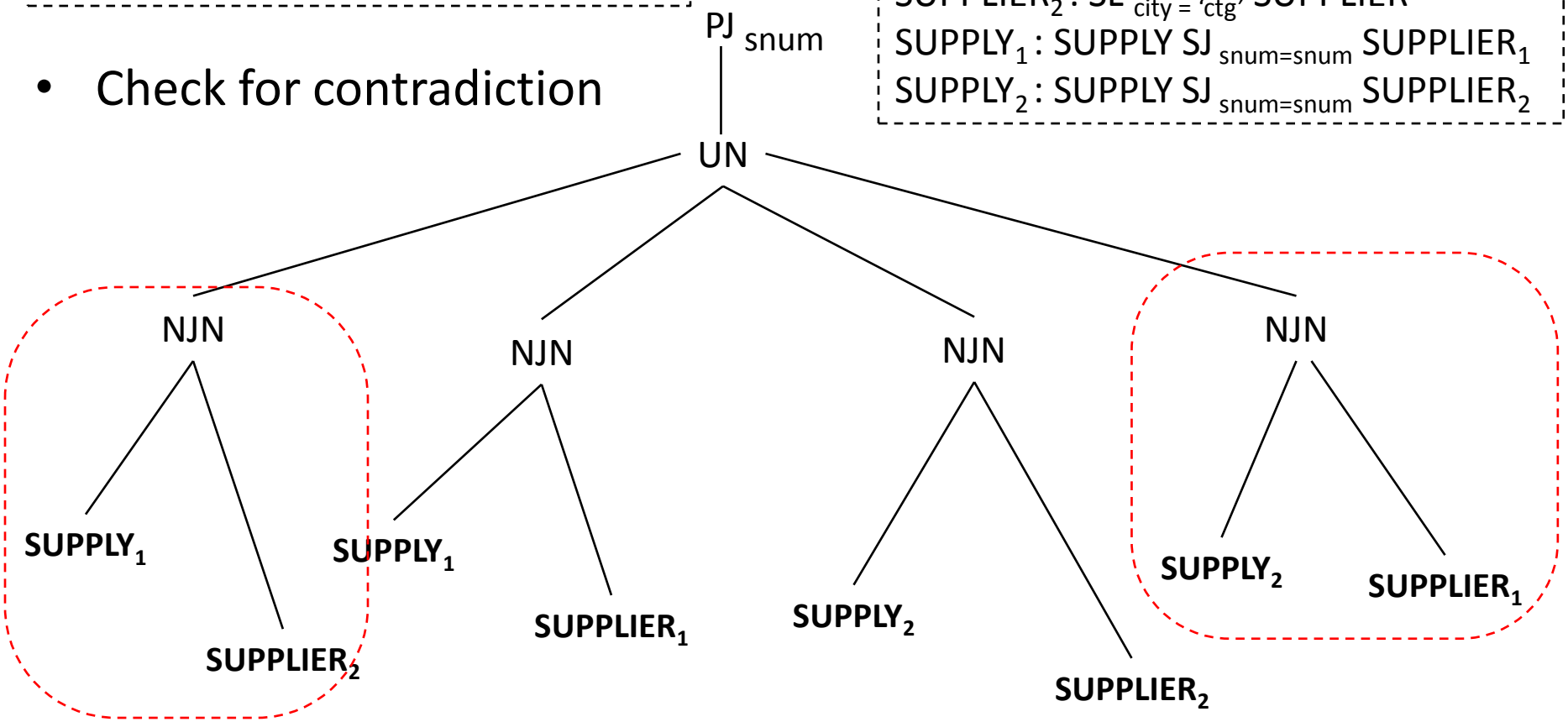


Example (contd.)

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

- Check for contradiction

$SUPPLIER_1 : SL_{city = 'dhk'}\ SUPPLIER$
$SUPPLIER_2 : SL_{city = 'ctg'}\ SUPPLIER$
$SUPPLY_1 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_1$
$SUPPLY_2 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_2$

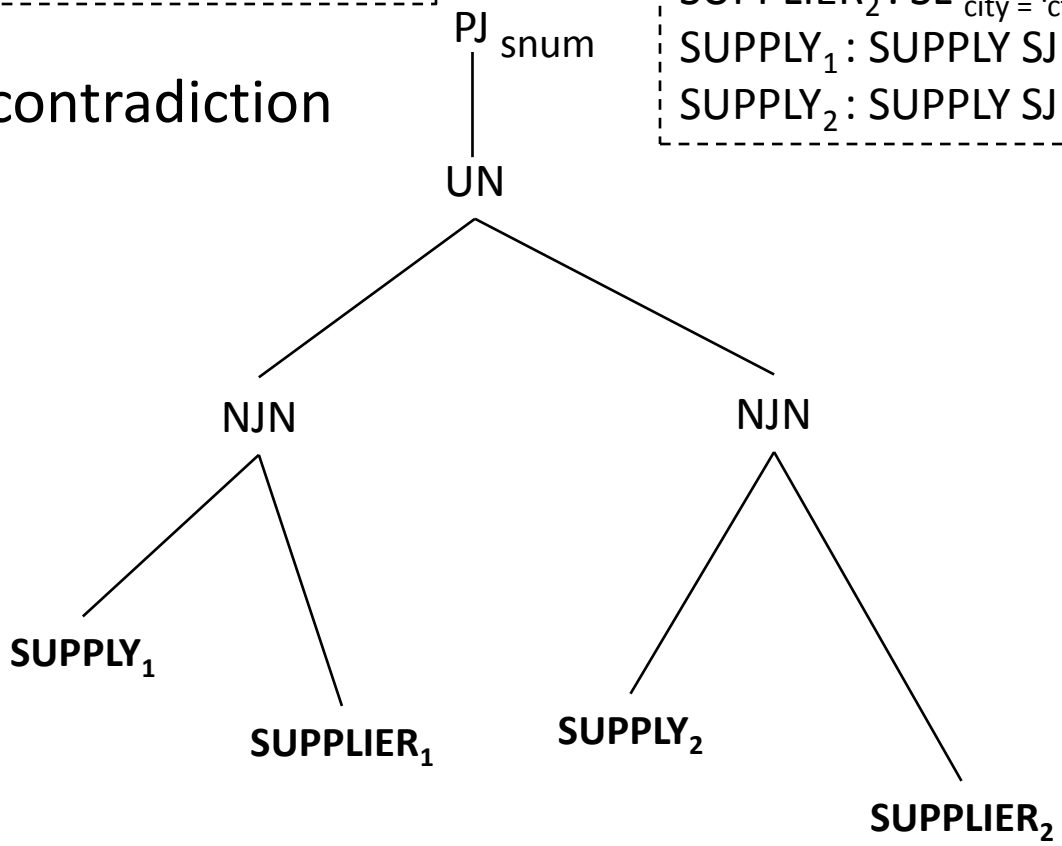


Example (contd.)

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

- Check for contradiction

$SUPPLIER_1 : SL_{city = 'dhk'} SUPPLIER$
$SUPPLIER_2 : SL_{city = 'ctg'} SUPPLIER$
$SUPPLY_1 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_1$
$SUPPLY_2 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_2$

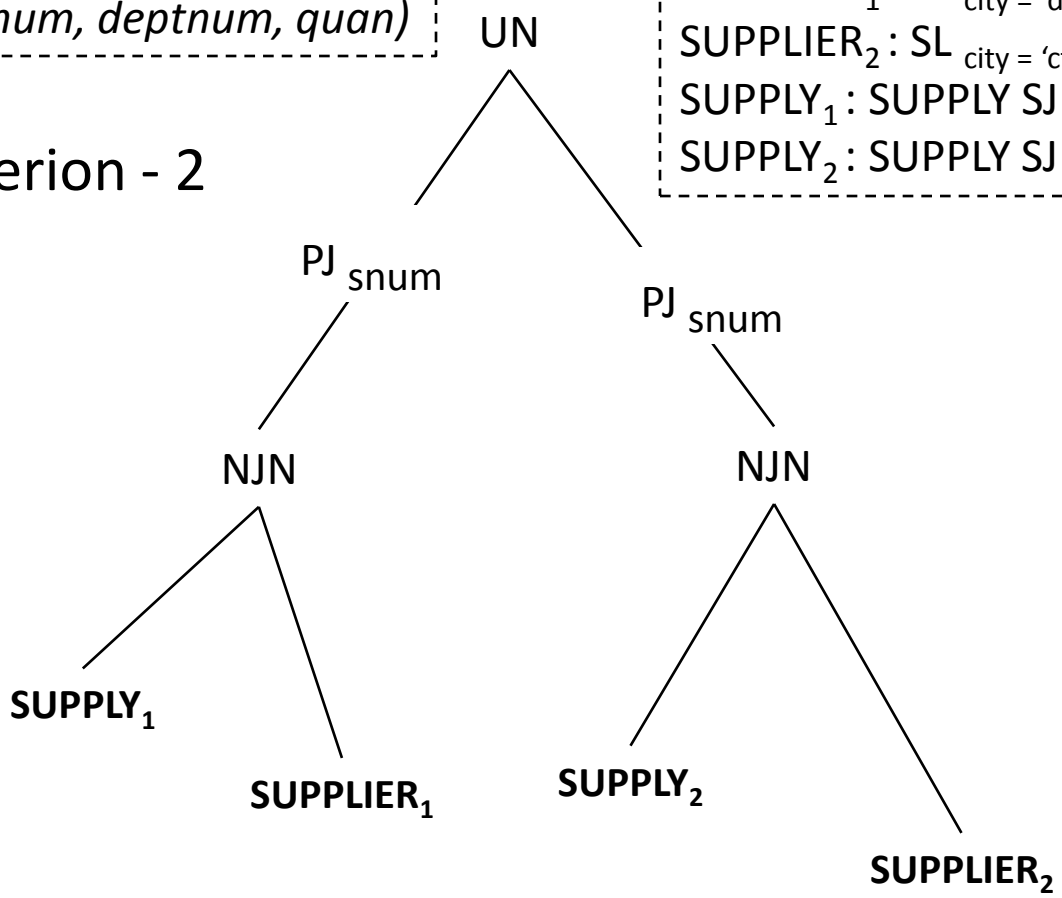


Example (contd.)

SUPPLIER (<u>snum</u>, name, city)
SUPPLY (<u>snum</u>, pnum, deptnum, quan)

- Apply criterion - 2

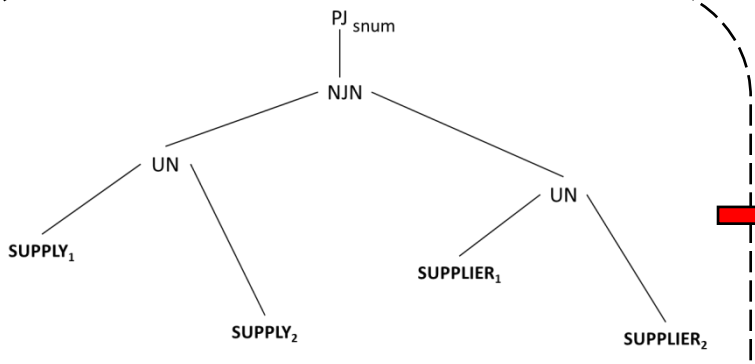
$SUPPLIER_1 : SL_{city = 'dhk'} SUPPLIER$
$SUPPLIER_2 : SL_{city = 'ctg'} SUPPLIER$
$SUPPLY_1 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_1$
$SUPPLY_2 : SUPPLY\ SJ_{snum=snum}\ SUPPLIER_2$



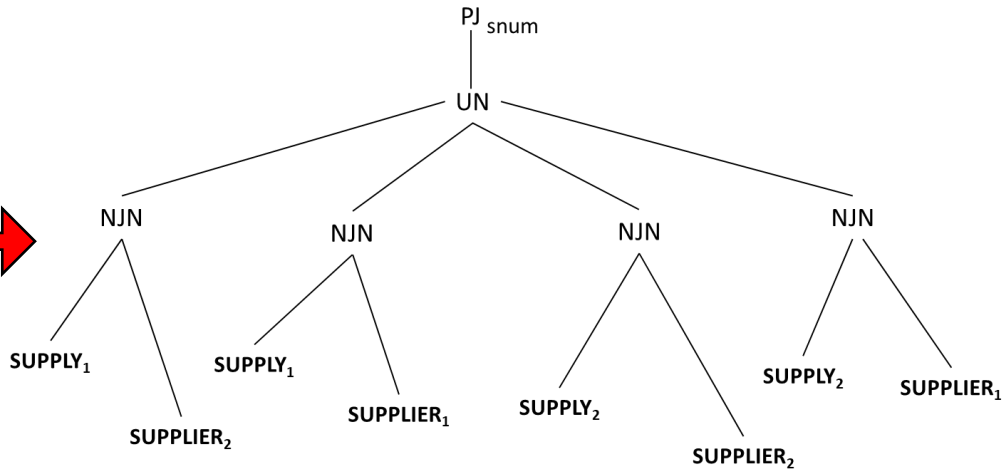
Criterion – 5 and 4

Criterion - 5:

Push *JN* down and UN up.



i. First, collect data (UN)
ii. Then Join (NJN)



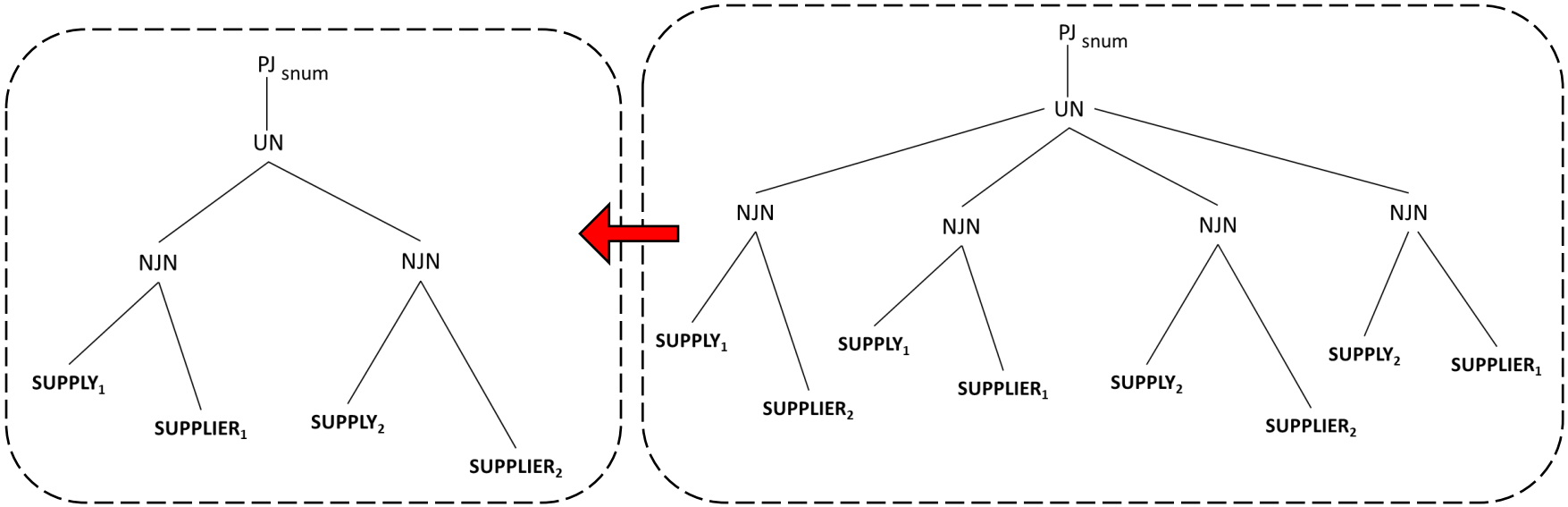
i. First, do join (NJN)
ii. Then collect data (UN)

(this is called *distributed join*)

Criterion – 5 and 4

Criterion - 4:

Eliminate *JN* between fragments that do not contribute to results.



Simplification of Vertically Fragmented Relations

Simplification of Vertically Fragmented Relations

- To determine a proper subset of the fragments which is sufficient for generating result for a query.

Example

Global schema:

EMP (empno, name, sal, tax, mgrnum, deptnum)

Fragmentation schema:

$EMP_1 = SL_{deptnum <= 10} PJ_{empnum, name, magnum, deptname} (EMP)$
$EMP_2 = SL_{10 < deptnum <= 20} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_3 = SL_{deptnum > 20} PJ_{empnum, name, mgrnum, deptname} (EMP)$
$EMP_4 = PJ_{empnum, name, sal, tax} (EMP)$

Example (contd.)

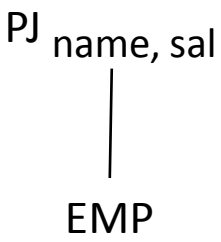
$EMP_1 = SL_{deptnum <= 10}\ PJ_{empnum, name, mgrnum, deptnum}\ (EMP)$

$EMP_2 = SL_{10 < deptnum <= 20}\ PJ_{empnum, name, mgrnum, deptnum}\ (EMP)$

$EMP_3 = SL_{deptnum > 20}\ PJ_{empnum, name, mgrnum, deptnum}\ (EMP)$

$EMP_4 = PJ_{empnum, name, sal, tax}\ (EMP)$

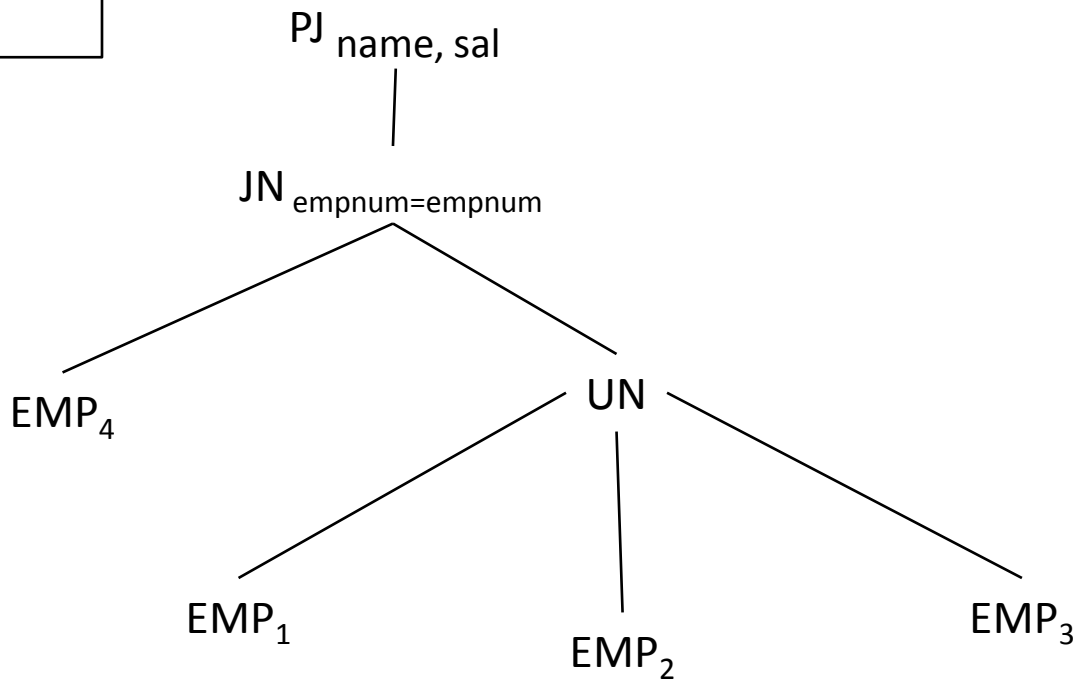
Q: $PJ_{name, sal}$ **EMP**



Example (contd.)

$EMP_1 = SL_{\text{deptnum <=10}}\ PJ_{\text{empnum, name, mgrnum, deptnum}}\ (EMP)$
$EMP_2 = SL_{\text{10 < deptnum <= 20}}\ PJ_{\text{empnum, name, mgrnum, deptnum}}\ (EMP)$
$EMP_3 = SL_{\text{deptnum > 20}}\ PJ_{\text{empnum, name, mgrnum, deptnum}}\ (EMP)$
$EMP_4 = PJ_{\text{empnum, name, sal, tax}}\ (EMP)$

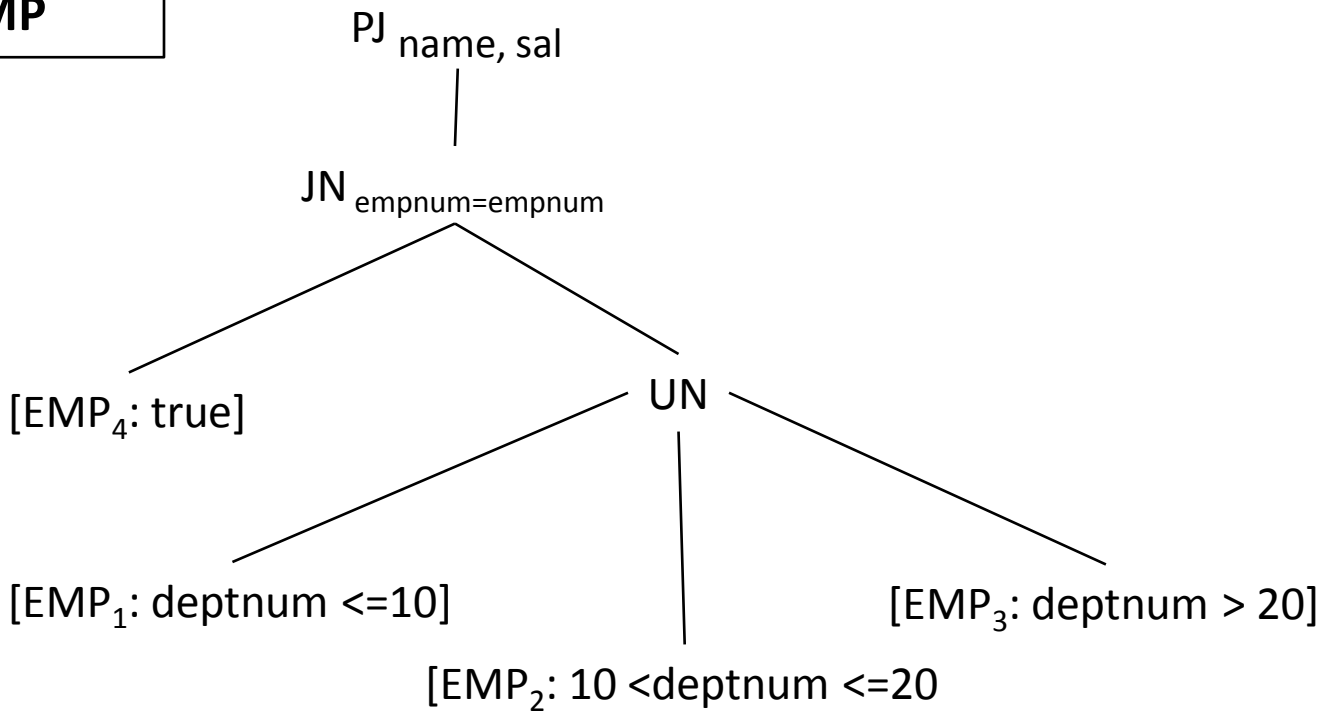
Q: $PJ_{\text{name, sal}}$ **EMP**



Example (contd.)

$EMP_1 = SL_{\text{deptnum <=10}}\ PJ_{\text{empnum, name, mgrnum, deptnum}}\ (EMP)$
$EMP_2 = SL_{\text{10 < deptnum <= 20}}\ PJ_{\text{empnum, name, mgrnum, deptnum}}\ (EMP)$
$EMP_3 = SL_{\text{deptnum > 20}}\ PJ_{\text{empnum, name, mgrnum, deptnum}}\ (EMP)$
$EMP_4 = PJ_{\text{empnum, name, sal, tax}}\ (EMP)$

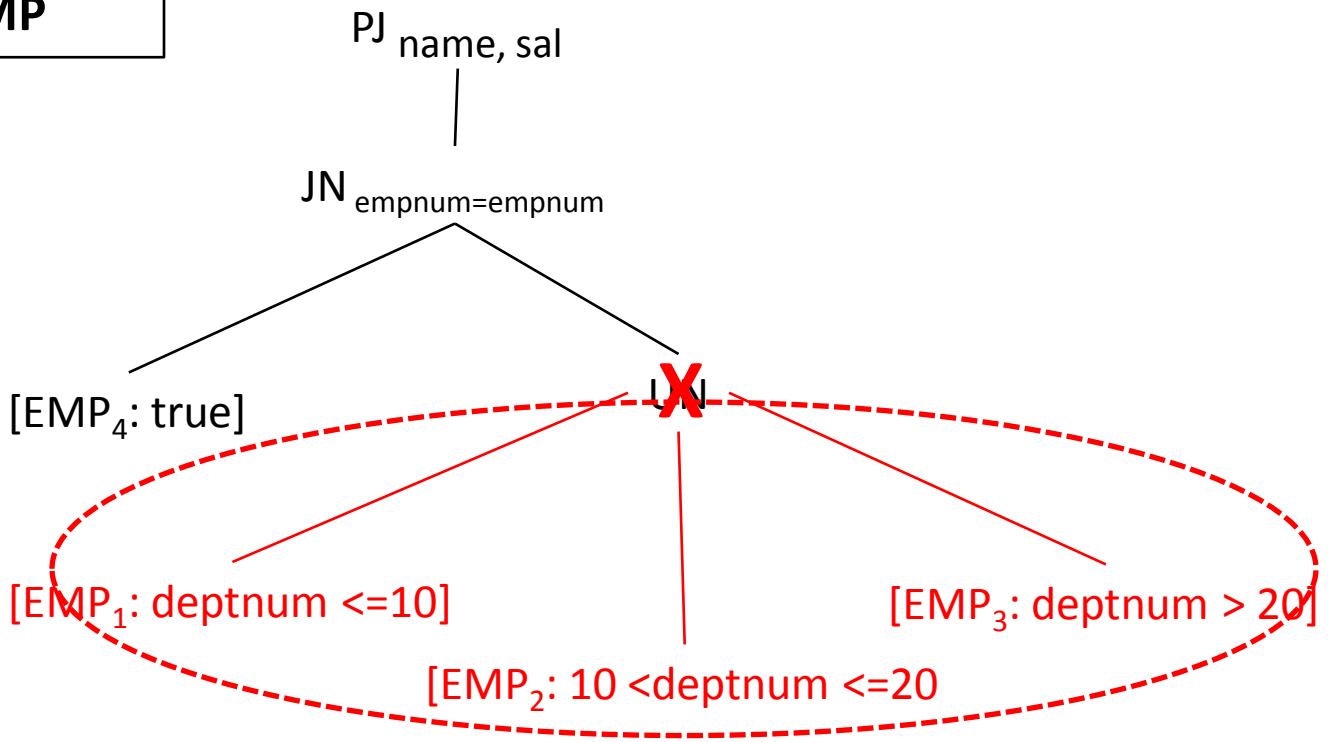
Q: $PJ_{\text{name, sal}}$ **EMP**



Example (contd.)

$EMP_1 = SL_{deptnum <= 10} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_2 = SL_{10 < deptnum <= 20} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_3 = SL_{deptnum > 20} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_4 = PJ_{empnum, name, sal, tax} (EMP)$

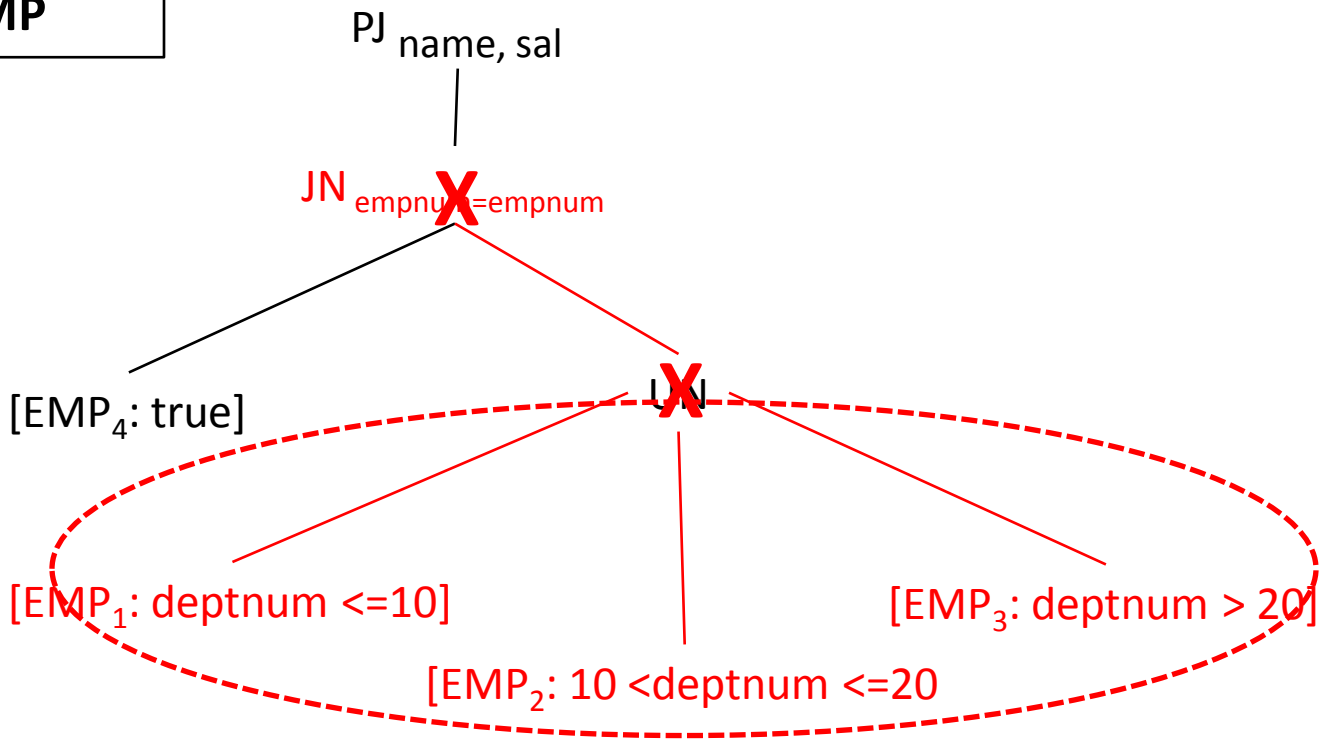
Q: $PJ_{name, sal}$ **EMP**



Example (contd.)

$EMP_1 = SL_{deptnum <= 10} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_2 = SL_{10 < deptnum <= 20} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_3 = SL_{deptnum > 20} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_4 = PJ_{empnum, name, sal, tax} (EMP)$

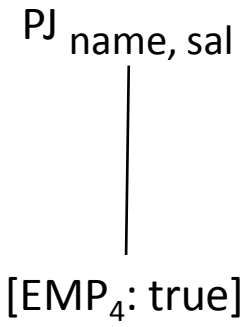
Q: $PJ_{name, sal}$ **EMP**



Example (contd.)

$EMP_1 = SL_{deptnum <= 10} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_2 = SL_{10 < deptnum <= 20} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_3 = SL_{deptnum > 20} PJ_{empnum, name, mgrnum, deptnum} (EMP)$
$EMP_4 = PJ_{empnum, name, sal, tax} (EMP)$

Q: $PJ_{name, sal}$ **EMP**



Parametric Queries

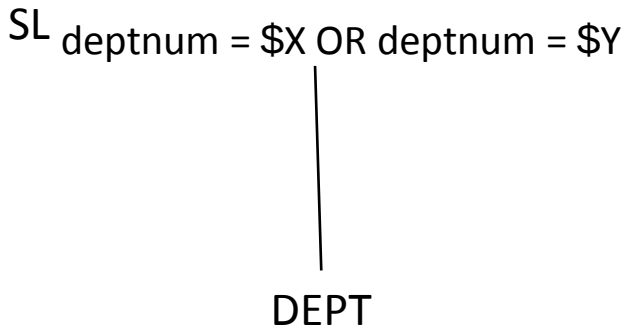
Parametric Queries

- *SL* include parameters values that are not known in *compile time*.
- When parametric queries executed, the user provides values in *runtime*, which are bound to (substituted for) parameters.
- Example: $SL_{\text{deptnum} = \$x}\ \text{DEPT}$

Simplification of Parametric Queries

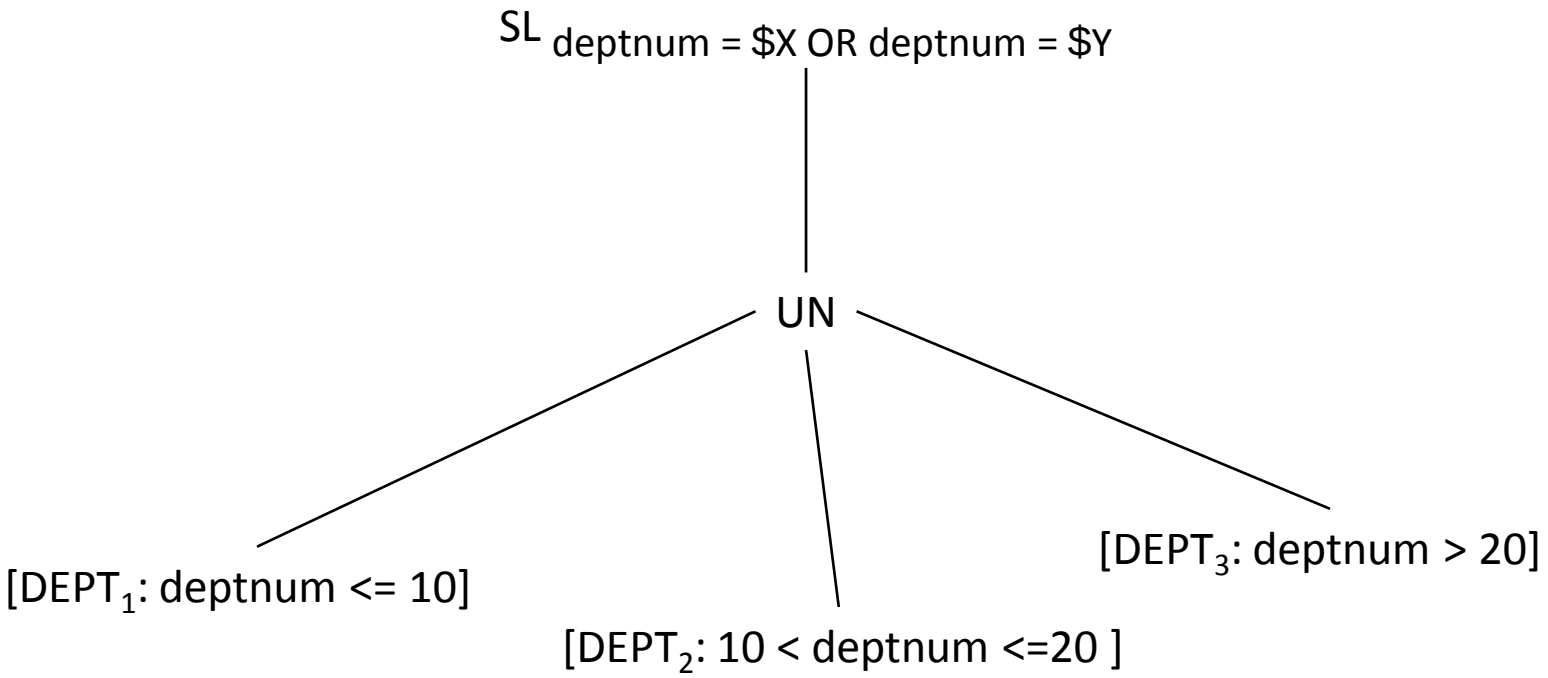
Example

Q: $\sigma_{\text{deptnum = \$X OR deptnum = \$Y}}$ **DEPT**



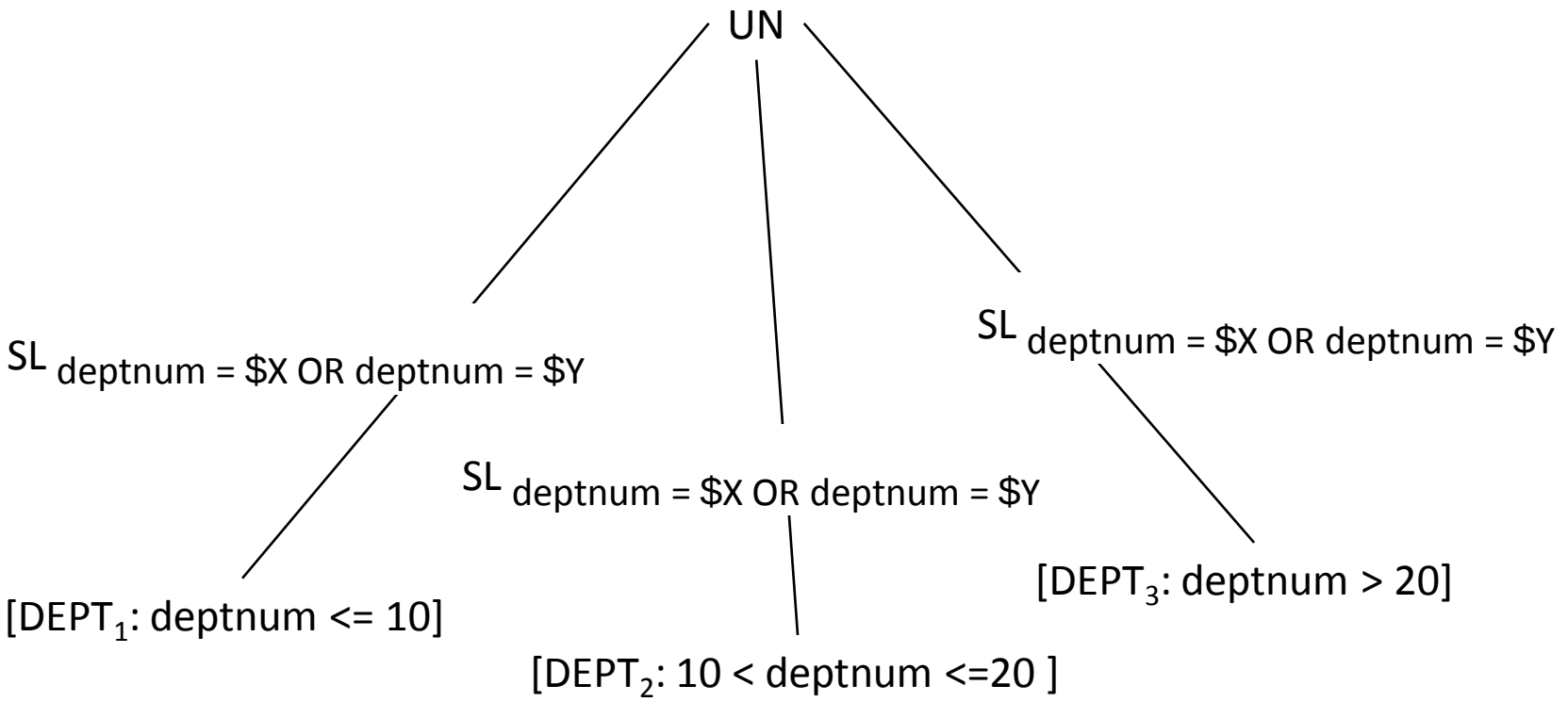
Example (contd.)

Q: $\sigma_{\text{deptnum = \$X OR deptnum = \$Y}}$ **DEPT**



Example (contd.)

Q: SL deptnum = $X OR deptnum = $Y **DEPT**

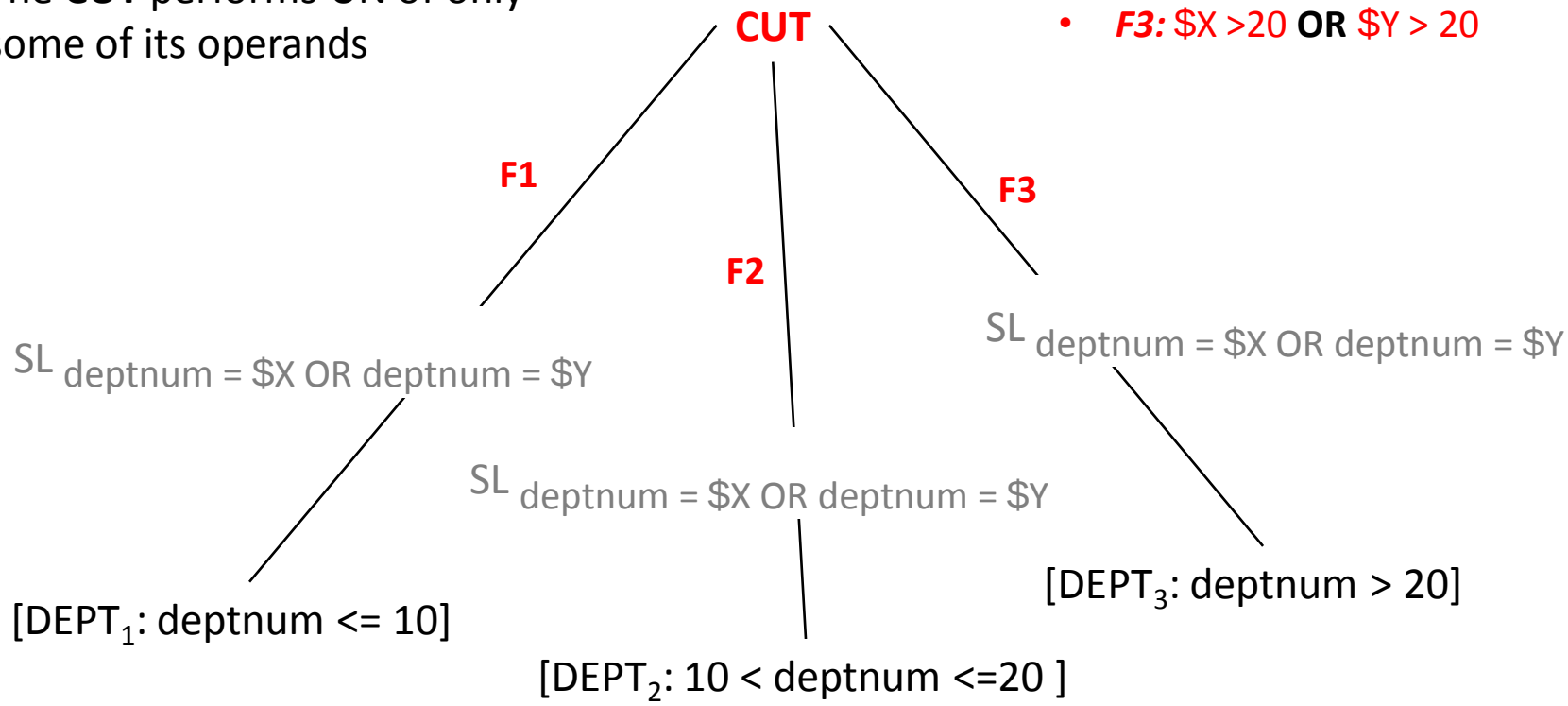


Example (contd.)

Q: $SL_{\text{deptnum = \$X OR deptnum = \$Y}}$ **DEPT**

- ***F1:*** $X <= 10 **OR** $Y <= 10
- ***F2:*** ($X > 10 AND $X <= 20) **OR** ($Y > 10 AND $Y <=20)
- ***F3:*** $X >20 **OR** $Y > 20

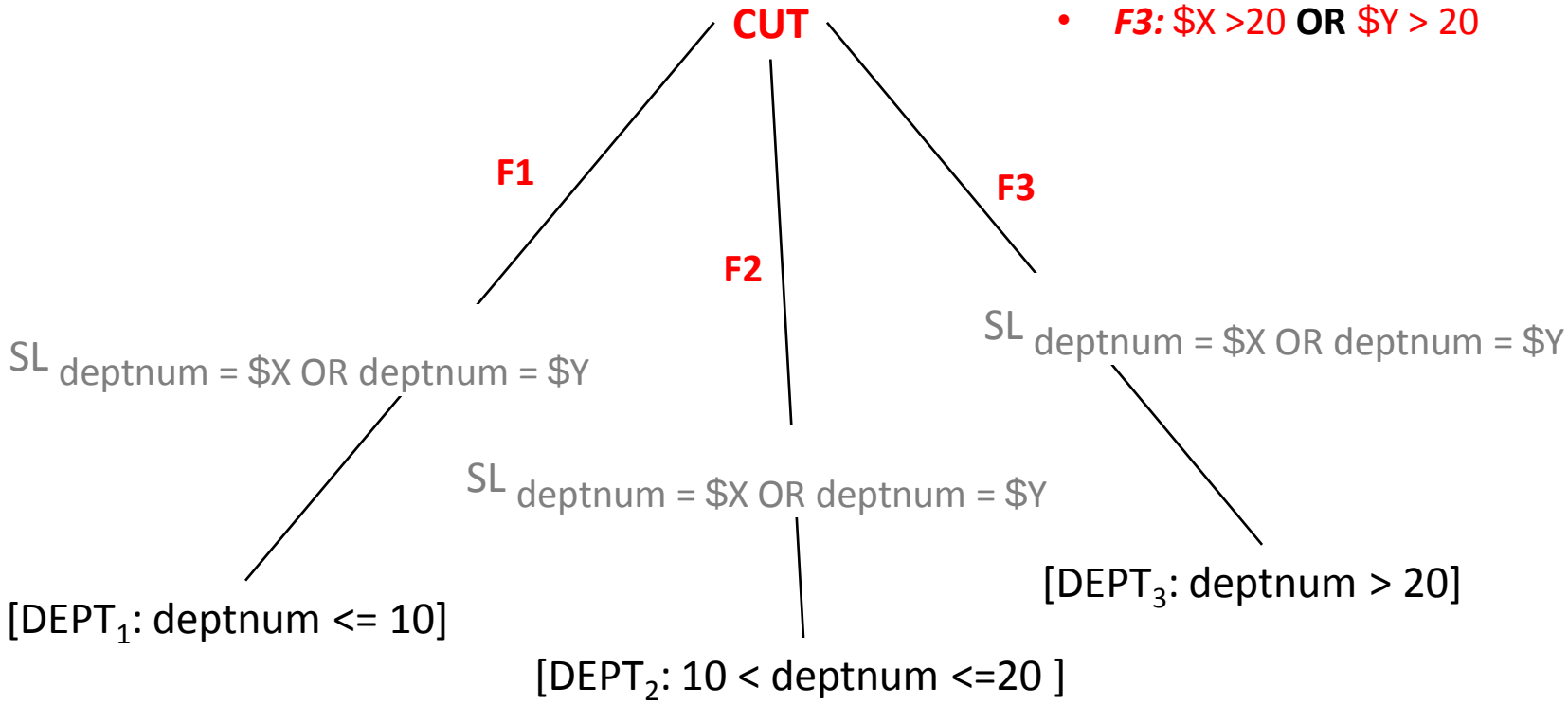
The **CUT** performs UN of only some of its operands



Example (contd.)

- ***F1:*** $X <= 10 **OR** $Y <= 10
- ***F2:*** ($X > 10 AND $X <= 20) **OR** ($Y > 10 AND $Y <=20)
- ***F3:*** $X >20 **OR** $Y > 20

QUESTION: What will happen to the tree for $X = 1, $Y = 35?



Additional Reading

- Get the intuition behind all the rules of algebra of qualified relation.
- Proof of rule – 6 and 7 of algebra of qualified relations.
  - Understand what is happening in each line of the proofs.

Practice Problems/ Questions

- Text book: exercise 5.2 (a) and (b)
  - N.B: use criteria – 1 to 5
- Text book: exercise 5.3, 5.4 and 5.7