

Relational Algebra

3. Relational Algebra

- *Relational Algebra* is a procedural DML.
- It specifies operations on relations to define new relations:
Select, Project, Union, Intersection,
Difference, Cartesian Product, Join, Divide.

3.1 SELECT

- Selects a subset of the tuples of a relation r , satisfying some condition.

$$\sigma_B(r) = \{t \in r : B(t)\}$$

- B is the selection condition, composed of selection clauses combined using AND, OR and NOT.
- A selection clause has the form

<attribute name> <op> <constant>

or

<attribute name> <op> <attribute name>

(*join*, introduce later)

where <op> is one of $=$, $<$, \leq , $>$, \geq or \neq .

STUDENT	
Person#	Name
1	Dr.C.C.Chen
3	Ms.K.Juliff
4	Ms.J.Gledill
5	Ms.B.K.Lee

RESEARCHER	
Person#	Name
1	Dr.C.C.Chen
2	Dr.R.G.Wilkinson

COURSE	
Department	Name
Psychology	Ph.D.
Comp.Sci.	Ph.D.
Comp.Sci.	M.Sc.
Psychology	M.Sc.

ENROLMENT

Enrolment#	Supervisee	Supervisor	Department	Name
1	1	2	Psychology	Ph.D.
2	3	1	Comp.Sci.	Ph.D.
3	4	1	Comp.Sci.	M.Sc.
4	5	1	Comp.Sci.	M.Sc.

- *Example:* Select the enrolment records for the students of person 1.

$\sigma_{(Supervisor=1)}(ENROLMENT)$

ENROLMENT				
Enrolment#	Supervisee	Supervisor	Department	Name
1	1	2	Psychology	Ph.D.
2	3	1	Comp.Sci.	Ph.D.
3	4	1	Comp.Sci.	M.Sc.
4	5	1	Comp.Sci.	M.Sc.

Enrolment#	Supervisee	Supervisor	Department	Name
2	3	1	Comp.Sci	Ph.D.
3	4	1	Comp.Sci	M.Sc.
4	5	1	Comp.Sci	M.Sc.

- Example: Select the enrolment records for person 1's non-Ph.D. students:

$\sigma_{(Supervisor=1) \text{ AND NOT}(Name \neq "PH.D.")}(ENROLMENT)$

ENROLMENT				
Enrolment#	Supervisee	Supervisor	Department	Name
1	1	2	Psychology	Ph.D.
2	3	1	Comp.Sci.	Ph.D.
3	4	1	Comp.Sci.	M.Sc.
4	5	1	Comp.Sci.	M.Sc.

Enrolment#	Supervisee	Supervisor	Department	Name
3	4	1	Comp.Sci	M.Sc
4	5	1	Comp.Sci	M.Sc

Properties:

- Commutative:

$$\sigma_{\langle cond1 \rangle} (\sigma_{\langle cond2 \rangle} (R)) = \\ \sigma_{\langle cond2 \rangle} (\sigma_{\langle cond1 \rangle} (R))$$

- Consecutive selects can be combined:

$$\sigma_{\langle cond1 \rangle} (\sigma_{\langle cond2 \rangle} (R)) = \\ \sigma_{\langle cond1 \rangle} \text{ AND }_{\langle cond2 \rangle} (R))$$

3.2 PROJECT

- Projects onto a subset X of the attributes of a relation.

$$\pi_X(r) = \{t[X] : t \in r\}$$

- *Remember* that a tuple, t is a mapping from attributes to elements of their domains. $t[X]$ is the restriction of that mapping to the set of attributes X .

- Example: Which courses are students enrolled in?

$$\pi_{Department, Name}(ENROLMENT) =$$

ENROLMENT				
Enrolment#	Supervisee	Supervisor	Department	Name
1	1	2	Psychology	Ph.D.
2	3	1	Comp.Sci.	Ph.D.
3	4	1	Comp.Sci.	M.Sc.
4	5	1	Comp.Sci.	M.Sc.

Department	Name
Psych.	Ph.D.
Comp.Sci	Ph.D.
Comp.Sci	M.Sc.

Properties:

- if $\langle \text{list2} \rangle$ contains all the attributes in $\langle \text{list1} \rangle$ then

$$\pi_{\langle \text{list1} \rangle} (\pi_{\langle \text{list2} \rangle} (R)) = \pi_{\langle \text{list1} \rangle} (R)$$

else

The operation is not well defined.

- commutes with selection:

$$\pi_X (\sigma_B (R)) = \sigma_B (\pi_X (R))$$

Exercise: Verify the above with:

$$\pi_{\{\text{Department}\}} (\sigma_{(\text{Department} = \text{"Psychology"})} (ENROLMENT)).$$

Properties:

- if $\langle \text{list2} \rangle$ contains all the attributes in $\langle \text{list1} \rangle$ then

$$\pi_{\langle \text{list1} \rangle} (\pi_{\langle \text{list2} \rangle} (R)) = \pi_{\langle \text{list1} \rangle} (R)$$

else

The operation is not well defined.

- commutes with selection: B cannot be specified outside of X

$$\pi_X(\sigma_B(R)) = \sigma_B(\pi_X(R))$$

Exercise: Verify the above with:

$$\pi_{\{\text{Department}\}} (\sigma_{(\text{Department} = \text{"Psychology"})}(\text{ENROLMENT})).$$

Questions

$$1) \pi (R \cup S) = \pi (R) \cup \pi (S)?$$

$$2) \pi (R \cap S) = \pi (R) \cap \pi (S)?$$

Answer:

$$2) \pi (R \cap S) \neq \pi (R) \cap \pi (S)$$

Example:

$$R = (Animal, Cat), S = (Animal, Dog)$$

π : project on the first column

$$\pi (R \cap S) = \{\}$$

$$\pi (R) \cap \pi (S) = \{Animal\}$$

3.3 UNION

- Is the set theoretic union of the tuples of two relations.

$$r \cup s = \{t: t \in r \text{ or } t \in s\}$$

- Note: Requires R and S to be union compatible: that there is a 1-1 correspondence between their attributes, in which corresponding attributes are over the same domain.

- Example:

$$R1 \leftarrow \sigma_{(Supervisor=2)}(ENROLMENT)$$

$$R2 \leftarrow \sigma_{(Name="M.Sc")}(ENROLMENT)$$

$$R1 \cup R2 =$$

Enrolment#	Supervisee	Supervisor	Department	Name
1	1	2	Psych.	Ph.D.
3	4	1	Comp.Sci	M.Sc
4	5	1	Comp.Sci	M.Sc

- Example: $STUDENT \cup RESEARCHER =$

Person#	Name
1	Dr C.C.Chen
3	Ms K.Juliff
4	Ms J.Gledhill
5	Ms B.K.Lee
2	Dr R.G.Wilkinson

3.4 INTERSECTION

- Is the set theoretic intersection of the tuples of two relations.

$$r \cap s = \{t: t \in r \text{ and } t \in s\}.$$

- Example:

$$R_1 \leftarrow \sigma_{(\text{Supervisor}=1)}(ENROLMENT)$$

$$R_2 \leftarrow \sigma_{(\text{Name}=\text{"Ph.D."})}(ENROLMENT)$$

$$R_1 \cap R_2 =$$

Enrolment#	Supervisee	Supervisor	Department	Name
2	3	1	Comp.Sci.	Ph.D.

- Example: $\text{STUDENT} \cap \text{RESEARCHER} =$

STUDENT	
Person#	Name
1	Dr.C.C.Chen
3	Ms.K.Juliff
4	Ms.J.Gledill
5	Ms.B.K.Lee

RESEARCHER	
Person#	Name
1	Dr.C.C.Chen
2	Dr.R.G.Wilkinson

Person#	Name
1	Dr C.C. Chen

3.5 DIFFERENCE

- Is the set difference of the tuples of two relations.

$$r - s = \{t: t \in r \text{ and } t \notin s\}$$

- Example: STUDENT – RESEARCHER =

Person#	Name
3	Ms K. Juliff
4	Ms J. Gledhill
5	Ms B.K. Lee

3.6 CARTESIAN PRODUCT

$$r \times s = \{t_1 || t_2 : t_1 \in r \text{ and } t_2 \in s\}$$

Where $t_1 || t_2$ indicates the concatenation of tuples.

Example:

ENROLMENT \times *RESEARCHER*

E'ment#	S'ee	S'or	D'ment	E'ment. Name	Person#	R'cher. Name
1	1	2	Psych.	Ph.D.	1	Dr C.C. Chen
1	1	2	Psych.	Ph.D.	2	Dr R.G.Wilkinson
2	3	1	Comp.Sci	Ph.D.	1	Dr C.C. Chen
2	3	1	Comp.Sci	Ph.D.	2	Dr R.G.Wilkinson
3	4	1	Comp.Sci	M.Sc.	1	Dr C.C. Chen
3	4	1	Comp.Sci	M.Sc.	2	Dr R.G.Wilkinson
4	5	1	Comp.Sci	M.Sc.	1	Dr C.C. Chen
4	5	1	Comp.Sci	M.Sc.	2	Dr R.G.Wilkinson

More useful is:

$R1 \leftarrow ENROLMENT \times RESEARCHER$

$\sigma_{(Supervisor=Person\#)}(R1) =$

E'ment#	S'ee	S'or	D'ment	E'ment. Name	Person#	R'cher. Name
1	1	2	Psych.	Ph.D.	2	Dr R.G.Wilkinson
2	3	1	Comp.Sci.	Ph.D.	1	Dr C.C. Chen
3	4	1	Comp.Sci.	M.Sc.	1	Dr C.C. Chen
4	5	1	Comp.Sci.	M.Sc.	1	Dr C.C. Chen

- or even better:

$R1 \leftarrow ENROLMENT \times RESEARCHER$

$R2 \leftarrow \sigma(Supervisor=Person\#)(R1) =$

$\pi\{E'ment\#, S'ee, S'or, R'cher.Name, D'ment, E'ment.Name\}(R2) =$

E'ment#	S'ee	S'or	R'cher. Name	D'ment	E'ment. Name
1	1	2	Dr R.G.Wilkinson	Psych.	Ph.D.
2	3	1	Dr C.C. Chen	Comp.Sci.	Ph.D.
3	4	1	Dr C.C. Chen	Comp.Sci.	M.Sc.
4	5	1	Dr C.C. Chen	Comp.Sci.	M.Sc.

- The last of these is also known as natural join, the next to last is equi-join.

3.7 JOIN

Is used to combine related tuples from two relations.

- 3.7.1 Theta-join

$$r \bowtie_B s = \{t_1 || t_2 : t_1 \in r \text{ and } t_2 \in s \text{ and } B\}$$

B is composed of conditions (combined with AND) of the form $A_i \theta B_j$ where A_i is an attribute of R, B_j is an attribute of S, and θ is a comparison operator.

- 3.7.2 Equi-join

Is a theta-join where each comparison operator is “=”.

Example:

ENROLMENT ⋈ *RESEARCHER* (*Supervisor=Person#*)

- 3.7.3 Natural join

Is an equi-join where only one attribute from each comparison is retained.

Example: $ENROLMENT \bowtie RESEARCHER (Supervisor), (Person\#)$

- Question: If two relations have no join attributes,

how do you define the join result? Why?

- 3.7.3 Natural join

Is an equi-join where only one attribute from each comparison is retained.

Example: $ENROLMENT \bowtie_{RESEARCHER} (Supervisor), (Person\#)$

- Question: If two relations have no join attributes,

how do you define the join result? Why?

$$R(A, B) \bowtie S(B, C) \bowtie T(C, D)$$

- *Notes:*

1. In a natural join, there may be several pairs of join attributes.

Example:

COURSE		
Department	Name	By
Comp.Sci	Ph.D.	Research
Comp.Sci.	M.Sc.	Research
Psychology	M.Sc.	Coursework

Calculate

$ENROLEMENT \bowtie_{COURSE} (Department, Name), (Department, Name)$

- 2. If the pairs of joining attributes are exactly those that are identically named, we can write

$ENROLMENT \bowtie COURSE$

3.8 DIVIDE

Suppose R is a relation over Z , S over X with $X \subseteq Z$. Let $Y = Z - X$. Then $R \div S$ is a relation over Y ,

$$R \div S = \{t : t \times S \subseteq R\}$$

Example:

P	
A	B
a ₁	b ₁
a ₁	b ₂
a ₂	b ₁
a ₃	b ₂
a ₄	b ₁
a ₅	b ₁
a ₅	b ₂

Q
B
b ₁
b ₂

$$P \div Q =$$

A
a ₁
a ₅

Typical use: Which courses are offered by all departments?

$$COURSE \div (\pi_{\{Department\}} COURSE)$$

Note: $\{\sigma, \pi, \cup, -, \times\}$ are sufficient to define all these operations: this is a relationally complete set of operators.