Relational Algebra:

* Translate from SQL to Algebra:

$\sigma_{condition}$ ~ Where CONDITION

π_{column} ~ SELECT column

$A \times B$ ~ FROM A, B

$A \cap B$ ~ Sel. From A intersect select .. from B

$A \cup B$ ~ " " A UNION " " B

$A - B$ ~ " " A EXCEPT " " B

PS: Algebraic operations returns sets while SQL Returns Duplicates.

$A \times B$: each row of A is paired with all rows of B

$\rho(C(\text{position} \geq 1 \rightarrow \text{new name}), \text{input})$

$R \bowtie_c S = \sigma_c(R \times S)$
Conditional join

A

i	ii
0	a
1	b
2	c

B

i1	ii1
0	u
1	v
2	w

A x B

i	ii	i1	ii1
0	a	0	u
0	a	1	v
0	a	2	w
1	b	0	u
1	b	1	v
1	b	2	w
2	c	0	u
2	c	1	v
2	c	2	w

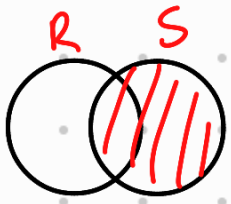
$\sigma_{i=i1}(A \times B) = A \bowtie_{i=i1}^{B}$

i	ii	i1	ii1
0	a	0	u
1	b	1	v
2	c	2	w

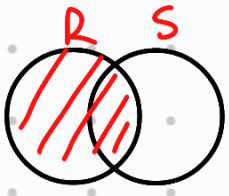
$R \bowtie S$

Natural join : equijoin on all matching column names

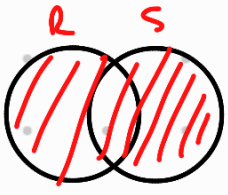
types of joins
- inner : discards different columns
- outer : includes columns with same and diff. columns.

$R \overset{\leftarrow}{\bowtie}_c S$: 

: replaces missing colums in R that satisfy "c" with NULL

$R \overset{\rightarrow}{\bowtie}_c S$: 

: replaces missing colums in S that satisfy "c" with NULL

$R \overset{\leftrightarrow}{\bowtie}_c S$: 

: Combination of R and S that satisfy "c" fill difference with NULL.

→ what is A/B ??

$A/B : \pi_x(A) - \pi_x(\pi_x(A) \times B - A)$

used for queries that involve the "all" or "every".

A/B = tuples of A that are associated with all tuples of B

⟹ Extended Relational Algebra (not very useful).

γ : Grouping operator → used to group tuples according to one or more attributes using the "Aggregate function" (COUNT, SUM, AVG, MIN, MAX)

$$\gamma_{G;\ A_1 \leftarrow f_1(E_1)\ ;\ A_2 \leftarrow f_2(E_2) \ldots}(R)$$

f_i

- G: Grouping Attributes
- A_1: name of result Aggregate Attr.
- R: input relation.

Example

Student (sid, ... dept);
Course (cid, ... dept);
Enroll (sid, cid, ...);

if I want to count all students enrolled in each dept ie CC, FMS, GTI, ...

$$\gamma_{dept}(COUNT(sid) \rightarrow n)\ (Student)$$

Output : (dept, n)

If I want to calculate the avg grade in each course

$$\gamma_{cid}(AVG(grade) \rightarrow a)(Enroll)$$

output (cid, a)

If I want to calculate Average Grade per department.

$$\gamma_{dept}(AVG(grade) \rightarrow a)(Enroll \bowtie Course)$$

output (dept, a)

If I want to select all departements with more than 20 students

$$\sigma_{count>20}(\gamma_{dept}(COUNT(sid) \rightarrow count)(Student))$$

Output (dept, count)