

IDB Lecture 4: Relational Algebra (RA)

Relational Algebra

Relational Algebra Expression takes an input of relation(s) (R), applies a sequence of operations and returns a relation as an output.

Operations

Projection (π) vertical operation which chooses columns. Of general form

$$\pi_{A_1, \dots, A_n}(R)$$

taking only the values of attributes A_1 to A_n for each tuple in R .

Selection (σ) horizontal operation on rows. Of general form

$$\sigma_{condition}(R)$$

taking only the tuples in R for which the condition is satisfied.

- for $\sigma_{\theta_1}(\sigma_{\theta_2}(R)) = \sigma_{\theta_1 \wedge \theta_2}(R)$, the RHS generally has faster runtime.

Product (\times) cartesian product *concatenates* each tuple of R with each tuples of S . Of general form

$$R \times S.$$

- relations must have a disjoint set of attributes
- $cardinality(R \times S) = cardinality(R) \times cardinality(S)$
 - where **Cardinality** is the number of attributes.
- $arity(R \times S) = arity(R) + arity(S)$
 - where **Arity** is the number of *attributes*.

Renaming (ρ) gives a new name to some attribute of a relation with syntax

$$\rho_{replacements}(R)$$

where a replacement has the form $A \rightarrow B$.

Union Intersection & Difference

Note: Relations must have the same attributes.

Union (\cup) set of all rows in R and S

Intersection (\cap) all rows that belong to both R and S

Difference ($-$) all rows in A that are not in B

Joining relations

Joins can be created by combining Cartesian product (\times) with selection (σ).

Natural Join (\bowtie) joins two tables on their *common attributes*

Theta-join $R \bowtie_{\theta} S = \sigma_{\theta}(R \times S)$

Equijoin \bowtie_{θ} where θ is a *conjunction of equalities*

Semijoin $R \ltimes_{\theta} S = \pi_X(R \bowtie_{\theta} S)$ where X is the set of attributes of R

Antijoin $R \bar{\ltimes}_{\sigma} S = R - (R \ltimes_{\sigma} S)$

Translating SQL to/from Relational Algebra

SELECT \iff projection (π)

FROM \iff Product (\times)

WHERE \iff selection (σ)

SELECT A_1, \dots, A_n

FROM T_1, \dots, T_m

WHERE

\updownarrow

$\pi_{A_1, \dots, A_n}(\sigma_{\langle \text{condition} \rangle}(T_1 \times \dots \times T_m))$

where common attributes in T_1, \dots, T_m must be renamed.