

Relational Algebra - extended

Relational Algebra

- Relational Algebra/language used by a DBMS
 - When coming up with query Plans and
 - Optimizing query plans
- Relational Algebra based on sets
 - Duplicates eliminated
 - $\Pi_{\text{lastname}}(\text{STUDENTS})$ will remove duplicate last names
- SQL based on multisets
 - Duplicates allowed
 - `SELECT lastname FROM STUDENTS` will allow duplicate last names

Multisets (Bags)

- Set is collection of distinct objects
 - $\{2,4,6\}$
- Bag is like a set but
 - Allows repeated elements
 - $\{2,4,4,6,6\} \neq \{2,4,6\}$
 - Order of elements does not matter
 - $\{2,4,6,2\} = \{6,2,4,2\}$

Bags (Union, Project, Aggregation)

- ❑ Union can combine tuples of the two relations without worrying about duplication elimination
 - $\Pi_{c1} (R1) \cup \Pi_{c1} (R2)$
- ❑ Projection will output the desired columns of the relation without elimination of duplications
- ❑ Selection will output tuple if it matches the condition and don't need to worry about duplicate
- ❑ Some aggregation functions like SUM, AVG will make sense with bags interpretation of relations

$\Pi_{\text{SUM}(\text{salary})} (\sigma_{\text{dept_name} = \text{"Comp. Sci."}} (\text{INSTRUCTOR}))$

Duplicates δ

- Delta δ to remove duplicate tuples
 - $\delta_{\text{lastname}}(R)$ eliminates duplicate tuples from relation R
 - `SELECT DISTINCT lastname FROM STUDENT;`

Projection – expression/rename

□ $\Pi_{\text{Exp} \rightarrow A} (R)$

- Exp is relational algebra expression involving attributes of relation R
- A is new name of the output/resultant attribute

□ $\Pi_{A1+A2 \rightarrow A} (R)$

□ SELECT A1+A2 as A FROM R

Grouping γ

□ $\gamma_L (R)$

- Group tuples in a relation R into distinct group
- L is list consisting of grouping attributes and aggregation attributes

□ $\gamma_{\text{dept_name}, \text{SUM}(\text{salary}) \rightarrow \text{sum_sal}} (\text{INSTRUCTOR})$

```
SELECT dept_name, SUM(salary) as sum_sal  
FROM INSTRUCTOR GROUP BY dept_name
```

Sorting

- Tau τ

- $\tau_{A1,A2}(R)$

- $A1,A2$ is list of attributes to be sorted by

- $\tau_{\text{last,birthdate}}(\text{STUDENT})$

Outer Joins

□ Full Outer Natural Join

■ R1 $\overset{O}{\bowtie}$ R2

□ Left Outer Natural Join

■ R1 $\overset{O}{\bowtie}_L$ R2