Chapter 4
Relational Algebra
Content

- Introduction
- Relational algebra
- Set operations on relations
- Selection
- Projection
- Cartesian production
- Join operation
- Division operation
- Other operations
- Update operations
Introduction

- Consider manipulations on the relation EMPLOYEE
  - Add a new employee
  - Move the employee whose name “Tung” to department 1
  - List names and birth dates of employees whose salary are over 20000

<table>
<thead>
<tr>
<th>FName</th>
<th>LName</th>
<th>BirthDate</th>
<th>Address</th>
<th>Sex</th>
<th>Salary</th>
<th>DNo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tung</td>
<td>Nguyen</td>
<td>12/08/1955</td>
<td>638 NVC Q5</td>
<td>Nam</td>
<td>40000</td>
<td>5</td>
</tr>
<tr>
<td>Hang</td>
<td>Bui</td>
<td>07/19/1968</td>
<td>332 NTH Q1</td>
<td>Nu</td>
<td>25000</td>
<td>4</td>
</tr>
<tr>
<td>Nhu</td>
<td>Le</td>
<td>06/20/1951</td>
<td>291 HVH QPN</td>
<td>Nu</td>
<td>43000</td>
<td>4</td>
</tr>
<tr>
<td>Hung</td>
<td>Nguyen</td>
<td>09/15/1962</td>
<td>Ba Ria VT</td>
<td>Nam</td>
<td>38000</td>
<td>5</td>
</tr>
<tr>
<td>Quang</td>
<td>Pham</td>
<td>11/10/1937</td>
<td>450 TV HN</td>
<td>Nam</td>
<td>55000</td>
<td>1</td>
</tr>
</tbody>
</table>
Introduction

- Study database programming
  - How the user can ask queries of the database
    • Select
  - How the user can modify the contents of the database
    • Insert, delete and update

- Relational model
  - Relational Algebra
    • Present a query by expressions
  - Relational Calculus
    • Present the result of a query
  - SQL (Structured Query Language)
Review

- **Algebra**
  - Operators
  - Atomic operands

- **In algebra arithmetic**
  - Operators: +, -, *, /
  - Operand – Variable: x, y, z
  - Constant
  - Expression
    - (x+7) / (y-3)
    - (x+y)*z and/or (x+7) / (y-3)
Relational algebra

- Variables – Relations
  - Set

- Operators
  - Set operations
    - Union \( \cup \)
    - Intersection \( \cap \)
    - Difference \( - \)
  - Retrieve parts of a relation
    - Selection \( \sigma \)
    - Projection \( \pi \)
  - Combine tuples of two relations
    - Cartesian product \( \times \)
    - Join \( \bowtie \)
Relational algebra

- **Constant**
  - Instance of the relation

- **Expression**
  - A query
  - A sequence of relational algebra operations

- **Operands and results of expressions**
  - Sets
Content

- Introduction
- Relational algebra
- **Set operations**
  - Selection
  - Projection
  - Cartesian product
  - Join operation
  - Division operation
  - Other operations
- Update operations
Set operation

- Relation is a set of tuples
  - The union \( R \cup S \)
  - The intersection \( R \cap S \)
  - The difference \( R - S \)

- Union Compatibility
  - Two relation schemas \( R(A_1, A_2, \ldots, A_n) \) and \( S(B_1, B_2, \ldots, B_n) \) are union compatibility if
    - The same degree \( n \)
    - And \( \text{DOM}(A_i) = \text{DOM}(B_i) \), \( 1 \leq i \leq n \)

- The result of \( \cup \), \( \cap \), and \( - \) operations
  - Relation
Union

- Given two relations R & S that are union compatible
- The union of R and S
  - Notation \( R \cup S \)
  - A relation consists of tuples that are in R or S or both (an element appears only one)

\[
R \cup S = \{ t / t \in R \lor t \in S \}
\]

- Example

<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R \cup S</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>
**Intersection**

- Given two relations R & S that are union compatible
- The intersection of R and S
  - Denotation $R \cap S$
  - A relation consists of tuples that are in R and S
    
    $$R \cap S = \{ t / t \in R \land t \in S \}$$

- Example

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>$\alpha$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>$\alpha$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R \cap S$</td>
<td>$\alpha$</td>
<td>2</td>
</tr>
</tbody>
</table>
**Difference**

- Given two relations R & S that are union compatible
- The difference of R and S
  - Denotation $R - S$
  - A relation consists of tuples that are in R but not in S
    $$R - S = \{ t \mid t \in R \land t \notin S \}$$

**Example**

<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R - S</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Properties

- Commutative

\[ R \cup S = S \cup R \]
\[ R \cap S = S \cap R \]

- Associative

\[ R \cup (S \cup T) = (R \cup S) \cup T \]
\[ R \cap (S \cap T) = (R \cap S) \cap T \]
Content

- Introduction
- Relational algebra
- Set operations
- **Selection**
- Projection
- Cartesian product
- Join operation
- Division operation
- Other operations
- Queries in relational algebra
Selection

- Is applied to relation R to produce a new relation with a subset of R’s tuples
- Tuples in the resulting relation satisfy some condition C
- Denotation $\sigma_c(R)$

- C is a Boolean expression made up of clauses
  - $<\text{attribute}> <\text{comparison operator}> <\text{constant}>$
  - $<\text{attribute}> <\text{comparison operator}> <\text{attribute}>$

  - $<\text{comparison op}> : <, >, \le, \ge, \ne, =$
  - Clauses are connected by Boolean operator : $\land, \lor, \neg$
Selection

- The result is a relation
  - The same list of attributes as $R$
  - The number of tuples is less than or equal to the number of tuples of $R$

Example

$\sigma_{(A=B) \land (D>5)}(R)$

<table>
<thead>
<tr>
<th>$R$</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>$\alpha$</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>5</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>$\beta$</td>
<td>12</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>$\beta$</td>
<td>23</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>
Selection

- Selection operator is commutative

\[ \sigma_{c_1}(\sigma_{c_2}(R)) = \sigma_{c_2}(\sigma_{c_1}(R)) = \sigma_{c_1 \land c_2}(R) \]
Example 1

- List all employees who work in department 4
  - Relation: EMPLOYEE
  - Attribute: DNo
  - Condition: DNo=4

\[ \sigma_{DNo=4}(EMPLOYEE) \]
Example 2

- Select tuples for all employees who either work in department 4 and make over $25,000 per year or work in department 5 and make over $30,000
  - Relation: EMPLOYEE
  - Attributes: SALARY, DNO
  - Condition:
    • (SALARY>25000 and DNO=4) or
    • (SALARY>30000 and DNO=5)

\[ \sigma \,(\text{SALARY}>25000 \, \land \, \text{DNO}=4 \, \lor \, (\text{SALARY}>30000 \, \land \, \text{DNO}=5)) \,(\text{EMPLOYEE}) \]
Content

- Introduction
- Relational algebra
- Set operations
- Selection
- **Projection**
- Cartesian product
- Join operation
- Division operation
- Other operations
- Update operations
Projection

- Is used to produce from a relation R a new relation that has only some of R’s columns
- Denotation $\pi_{A_1, A_2, \ldots, A_k}(R)$

- The result is a relation
  - Has k attributes
  - The number of tuples is less than or equal to the number of tuples of R

Example

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>10</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>20</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>30</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>40</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

$\pi_{A, C}(R)$
Projection

- Projection operator is not commutative

\[ \pi_{A_1, A_2, \ldots, A_n}(R) = \pi_{A_1}(\pi_{A_2}(\ldots(\pi_{A_n}(R)\ldots))) \]

\[ \pi_{A_1, A_2, \ldots, A_n}(R) = \pi_{A_1, A_2, \ldots, A_n}(R), \ n \leq m \]
Example 3

- List out the name and salary of employees

\[ \pi_{\text{LNAME, FNAME, SALARY}}(\text{EMPLOYEE}) \]
Example 4

- Find the SSN of employees who either work on projects or have dependents
Example 5

- Find the SSN of employees who work on projects and have dependents
Example 6

- Find the SSN of employees who do not have any dependents
Extended projection

- Extending the projection operator to allow it to compute with components of tuples

- Denotation $\pi_{F_1, F_2, \ldots, F_n}(E)$
  - $E$ is a relation algebra expression
  - $F_1, F_2, \ldots, F_n$ are arithmetic expressions involving
    - Attributes in $E$
    - Constants
    - Arithmetic operators ($a + b : \text{sum}$)
    - String operators ($c || d : \text{concatenate}$)
Example 7

- List out the employees’ name and salary increased by 10%

\[ \pi_{\text{LNAME, FNAME, SALARY*1.1}}(\text{EMPLOYEE}) \]
Sequences of operations

- Apply several relational algebra operations one after one
  - A single relational algebra expression
  \[
  \pi_{A_1, A_2, \ldots, A_k} (\sigma_C (R)) \quad \sigma_C (\pi_{A_1, A_2, \ldots, A_k} (R))
  \]
  
  - Break down a complex expression into simpler steps
    - Step 1 \( \sigma_C (R) \)
    - Step 2 \( \pi_{A_1, A_2, \ldots, A_k} (\text{the result of step 1}) \)

Giving a name

CuuDuongThanCong.com
https://fb.com/tailieudientucntt
Assignment operator

- Is often used to receive the result of an operation
  - The intermediate result in a sequence of operations
- Denotation ←

- Example
  - Step 1: \( S \leftarrow \sigma_C(R) \)
  - Step 2: \( \text{RESULT} \leftarrow \pi_{A_1, A_2, ..., A_k}(S) \)
Rename operator

- Is used to rename either the relation name or attribute name
  - Relation
    Examine $R(B, C, D)$
    $\rho_S(R)$ : Rename the name of relation $R$ to $S$
  - Attribute
    $\rho_{X, C, D}(R)$ : Rename the name of attribute $B$ to $X$
    Rename the name of relation $R$ to $S$ and the name of attribute $B$ to $X$
    $\rho_{S(X,C,D)}(R)$.
Example 8

- List out the name of employees who work in department 4

- **Case 1:** \( \pi_{\text{LNAME, FNAME}}(\sigma_{DNO=4}(\text{EMPLOYEE})) \)

- **Case 2:**  
  \[ \text{EMP\_DEP4} \leftarrow \sigma_{DNO=4}(\text{EMPLOYEE}) \]  
  \[ \text{RESULT} \leftarrow \pi_{\text{LNAME, FNAME}}(\text{EMP\_DEP4}) \]  
  \[ \rho_{\text{RESULT(LASTNAME, FIRSTNAME)}}(\pi_{\text{LNAME, FNAME}}(\text{EMP\_DEP4})) \]  
  \[ \text{RESULT(LASTNAME, FIRSTNAME)} \leftarrow \pi_{\text{LNAME, FNAME}}(\text{EMP\_DEP4}) \]
Content

- Introduction
- Relational algebra
- Set operations
- Selection
- Projection
- **Cartesian product**
- Join operation
- Division operation
- Other operations
- Update operations
Cartesian product

- Cross-product
  - Is used to combine tuples from two relations in a combinatorial fashion
- Denotation \( R \times S \)
- The result is a relation \( Q \)
  - \( Q \) has one tuple for each combination of tuples, one from \( R \) and one from \( S \)
  - If \( R \) has \( u \) tuples and \( S \) has \( v \) tuples,
    - Then \( Q \) will have \( (u \times v) \) tuples
  - If \( R \) has \( n \) attributes and \( S \) has \( m \) attributes,
    - Then \( Q \) will have \( (n + m) \) attributes \((R^+ \cap S^+ = \emptyset)\)
Cartesian product

- Example

<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>β</td>
<td>20</td>
<td>-</td>
</tr>
</tbody>
</table>

\[ R \times S \]

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>X</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1</td>
<td>α</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>α</td>
<td>1</td>
<td>β</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>α</td>
<td>1</td>
<td>β</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>α</td>
<td>1</td>
<td>γ</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>α</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>β</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>β</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>γ</td>
<td>10</td>
<td>-</td>
</tr>
</tbody>
</table>

\[ \rho_{(X,C,D)}(S) \]
Cartesian product

- **Example**

<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>10</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>10</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>20</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>γ</td>
<td>10</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R × S</th>
<th>A</th>
<th>R.B</th>
<th>S.B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1</td>
<td>α</td>
<td>10</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>1</td>
<td>β</td>
<td>10</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>1</td>
<td>β</td>
<td>20</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>1</td>
<td>γ</td>
<td>10</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>α</td>
<td>10</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>β</td>
<td>10</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>β</td>
<td>20</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>γ</td>
<td>10</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

unambiguous
Cartesian product

- Cartesian product is often followed by a selection operation

\[ R \times S \]

\[ \sigma_{A=S.B} (R \times S) \]

<table>
<thead>
<tr>
<th>A</th>
<th>R.B</th>
<th>S.B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1</td>
<td>α</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>α</td>
<td>1</td>
<td>β</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>α</td>
<td>1</td>
<td>β</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>1</td>
<td>γ</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>α</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>β</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>β</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>γ</td>
<td>10</td>
<td>-</td>
</tr>
</tbody>
</table>
Example 9

- For each department, list out the information of the manager

<table>
<thead>
<tr>
<th>DNAME</th>
<th>DNUMBER</th>
<th>MGRSSN</th>
<th>MGRSTARTDAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nghien cuu</td>
<td>5</td>
<td>333445555</td>
<td>05/22/1988</td>
</tr>
<tr>
<td>Dieu hanh</td>
<td>4</td>
<td>987987987</td>
<td>01/01/1995</td>
</tr>
<tr>
<td>Quan ly</td>
<td>1</td>
<td>888665555</td>
<td>06/19/1981</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DNAME</th>
<th>DNUMBER</th>
<th>MGRSSN</th>
<th>SSN</th>
<th>FNAME</th>
<th>LNAME</th>
<th>BIRTHDATE</th>
<th>ADDRESS</th>
<th>SEX</th>
<th>SALARY</th>
<th>DNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tung Nguyen</td>
<td>333445555</td>
<td>987987987</td>
<td>888665555</td>
<td>Vinh</td>
<td>Pham</td>
<td>12/08/1955</td>
<td>638 NVC Q5</td>
<td>Nam</td>
<td>40000</td>
<td>5</td>
</tr>
<tr>
<td>Hang Bui</td>
<td>999887777</td>
<td>888665555</td>
<td>888665555</td>
<td>332 NTH Q1</td>
<td>Nu</td>
<td>07/19/1968</td>
<td>638 NVC Q5</td>
<td>Nu</td>
<td>25000</td>
<td>4</td>
</tr>
<tr>
<td>Nhu Le</td>
<td>987654321</td>
<td>987987987</td>
<td>987987987</td>
<td>291 HVH QPN</td>
<td>Nu</td>
<td>06/20/1951</td>
<td>638 NVC Q5</td>
<td>Nu</td>
<td>43000</td>
<td>4</td>
</tr>
<tr>
<td>Hung Nguyen</td>
<td>987987987</td>
<td>888665555</td>
<td>987987987</td>
<td>Ba Ria VT</td>
<td>Nam</td>
<td>09/15/1962</td>
<td>638 NVC Q5</td>
<td>Nam</td>
<td>38000</td>
<td>5</td>
</tr>
</tbody>
</table>
Example 9

- **Step 1:**
  - Cartesian product DEPARTMENT & EMPLOYEE
    
    \[
    \text{EMP\_DEP} \leftarrow (\text{DEPARTMENT} \times \text{EMPLOYEE})
    \]

- **Step 2:**
  - Select tuples that satisfy the condition MgrSSN=SSN
    
    \[
    \text{RESULT} \leftarrow \sigma_{\text{MgrSSN}=\text{SSN}}(\text{EMP\_DEP})
    \]
Example 10

- Find the highest salary in company

<table>
<thead>
<tr>
<th>FNAME</th>
<th>LNAME</th>
<th>SALARY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tung</td>
<td>Nguyen</td>
<td>40000</td>
</tr>
<tr>
<td>Hang</td>
<td>Bui</td>
<td>25000</td>
</tr>
<tr>
<td>Nhu</td>
<td>Le</td>
<td>43000</td>
</tr>
<tr>
<td>Hung</td>
<td>Nguyen</td>
<td>38000</td>
</tr>
</tbody>
</table>
Example 10

**Step 1:**
- Select salaries which are not the highest one

\[
R1 \leftarrow (\pi_{\text{SALARY}}(\text{EMPLOYEE}))
\]

\[
R2 \leftarrow \sigma_{\text{EMPLOYEE.SALARY} < R1.\text{SALARY}}(\text{EMPLOYEE} \times R1)
\]

\[
R3 \leftarrow \pi_{\text{EMPLOYEE.SALARY}}(R2)
\]

**Step 2:**
- Let do the difference of the set of salary and salary in R3

\[
\text{RESULT} \leftarrow \pi_{\text{SALARY}}(\text{EMPLOYEE}) - R3
\]
**Example 11**

- Find the departments that have the same locations as the department 5

Which locations does the department 5 have?

<table>
<thead>
<tr>
<th>DNUMBER</th>
<th>DLOCATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TP HCM</td>
</tr>
<tr>
<td>4</td>
<td>HA NOI</td>
</tr>
<tr>
<td>5</td>
<td>VUNGTAU</td>
</tr>
<tr>
<td>5</td>
<td>NHATRANG</td>
</tr>
<tr>
<td>5</td>
<td>TP HCM</td>
</tr>
</tbody>
</table>

Which departments will have locations which are in that set

<table>
<thead>
<tr>
<th>DNUMBER</th>
<th>DLOCATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>TP HCM</td>
</tr>
<tr>
<td>4</td>
<td>HA NOI</td>
</tr>
<tr>
<td>5</td>
<td>VUNGTAU</td>
</tr>
<tr>
<td>5</td>
<td>NHATRANG</td>
</tr>
<tr>
<td>5</td>
<td>TP HCM</td>
</tr>
</tbody>
</table>
Example 11

- **Step 1:**
  - Find the locations of the department 5
  
  \[
  \text{LOC\_DEP5}(\text{Loc}) \leftarrow \pi_{\text{DLOCATION}}(\sigma_{\text{DNUMBER}=5}(\text{DEPT\_LOCATIONS}))
  \]

- **Step 2:**
  - Select the departments that have the same locations as LOC\_DEP5

\[
\begin{align*}
R1 & \leftarrow \sigma_{\text{DNUMBER} \neq 5}(\text{DEPT\_LOCATIONS}) \\
R2 & \leftarrow \sigma_{\text{DLOCATION}=\text{Loc}}(R1 \times \text{LOC\_DEP5}) \\
\text{RESULT} & \leftarrow \pi_{\text{DNUMBER}}(R2)
\end{align*}
\]
Introduction

Relational algebra

Set operations

Selection

Projection

Cartesian product

Join operation
  - Natural join
  - Theta join
  - Equi join

Divide operation

Other operations

Update operations
Join operation

- Is used to combine related tuples from 2 relations into single tuples

- Denotation $R \bowtie S$
  - $R(A_1, A_2, \ldots, A_n)$ and $S(B_1, B_2, \ldots, B_m)$

- Result is a relation $Q$
  - Has $(n + m)$ attributes $Q(A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_m)$
  - A tuple of $Q$ is a combination of tuples from $R$ and $S$ satisfying some join condition
    - The form : $A_i \theta B_j$
    - $A_i$ : the attribute from $R$, $B_j$ : the attribute from $S$
    - $A_i$ and $B_j$ have the same domain
    - $\theta$ : comparison operators $\neq, =, <, >, \leq, \geq$
Join operation

- **Categories**
  - *Theta join* pairs tuples using one specific condition
    - Denotation $R \bowtie_C S$
    - $C$ refers to an arbitrary condition for attributes
  - *Equijoin* when $C$ involves equality comparisons only
  - *Natural join*
    - Denote $R \bowtie S$ or $R \ast S$
    - $R^+ \cap S^+ \neq \emptyset$
    - Only one join attribute is kept
Join operation

- Example of theta join

\[ R \bowtie_{B < D} S \]

\[ R \bowtie_{C} S = \sigma_{C}(R \times S) \]
Join operation

- **Example of equijoin**

<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

  $R \bowtie_{C=D} S$

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>SCC</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

  $\rho_{(S,C,D)} S$
Join operation

- Example of natural join

<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R * S</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>7</td>
</tr>
</tbody>
</table>
Example 12

- Find the employees whose salary are greater than the salary of the employee ‘Tùng’

\[
\text{EMPLOYEE(LNAME, FNAME, SSN, ..., SALARY, DNO)}
\]

\[
\text{R1(SAL) } \leftarrow \pi_{\text{SALARY}} (\sigma_{\text{FNAME}=\text{‘Tung’}} (\text{EMPLOYEE}))
\]

\[
\text{RESULT } \leftarrow \text{EMPLOYEE } \bowtie_{\text{SALARY}>\text{SAL}} \text{ R1}
\]

\[
\text{RESULT(LNAME, FNAME, SSN, ..., SALARY, DNO, SAL)}
\]
Example 13

- For each employee, find the information of the department that he/she is working for

\[
\text{RESULT } \leftarrow \text{EMPLOYEE } \bowtie_{\text{DNO} = \text{DNUMBER}} \text{DEPARTMENT}
\]

\[
\text{RESULT(LNAME, FNAME, SSN, ..., DNO, DNAME, DNUMBER, ...)}
\]
Example 14

- Find the locations for each department

\[
\text{DEPARTMENT}(\text{DNAME}, \text{DNUMBER}, \text{MGRSSN}, \text{MGRSTARTDATE})
\]
\[
\text{DEPT\_LOCATIONS}(\text{DNUMBER}, \text{DLOCATION})
\]

\[
\text{RESULT} \leftarrow \text{DEPARTMENT} \times \text{DEPT\_LOCATIONS}
\]

\[
\text{RESULT}(\text{DNAME}, \text{DNUMBER}, \text{MGRSSN}, \text{MGRSTARTDATE}, \text{DLOCATION})
\]
Example 9

- For each department, list out the information of the manager

https://fb.com/tailieudientucntt
Example 10

- Find the highest salary in company
Example 11

- Find the departments that have the same locations as the department 5
A complete set of relational algebra operations

- The set of relational algebra operations \{\sigma, \pi, \times, -, \cup\} is called a complete set
  - Any of other relational algebra operations can be expressed as a sequence of operations from this set

- Example
  - \( R \cap S = R - (R - S) \)
  - \( R \Join_c S = \sigma_c (R \times S) \)
Content

- Introduction
- Relational algebra
- Set operations
- Selection
- Projection
- Cartesian product
- Join operation
- Division operation
- Other operations
- Update operations
Division

- Is used to retrieve tuples from $R$ that satisfy all tuples from $S$
- Denotation $R \div S$
  - $R(Z)$ and $S(X)$
    - $Z$ is attribute set of $R$, $X$ is attribute set of $S$
    - $X \subseteq Z$
- Result is a relation $T(Y)$
  - Has $Y=Z-X$
  - Includes tuples $t$, if for all $t_S \in S$, there exists a tuple $t_R \in R$ with 2 conditions
    - $t_R(Y) = t$
    - $t_R(X) = t_S(X)$
Division

- Example

\[
R \div S
\]

\[
\begin{array}{c|c|c|c|c|c}
R & A & B & C & D & E \\
\hline
\alpha & a & \alpha & a & 1 \\
\alpha & a & \gamma & a & 1 \\
\alpha & a & \gamma & b & 1 \\
\beta & a & \gamma & a & 1 \\
\beta & a & \gamma & b & 3 \\
\gamma & a & \gamma & a & 1 \\
\gamma & a & \gamma & b & 1 \\
\gamma & a & \beta & b & 1 \\
\hline
\end{array}
\]

\[
\begin{array}{c|c|c}
S & D & E \\
\hline
a & 1 \\
b & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c}
A & B & C \\
\hline
\alpha & a & \gamma \\
\gamma & a & \gamma \\
\end{array}
\]
Example 15

- Find the SSN of employees who work on all the projects
Example 16

- Find the SSN of employees who work for all projects that the department 4 controls
Division

- Express the division operation by the complete set of relational algebra operations

\[ Q_1 \leftarrow \pi_Y(R) \]
\[ Q_2 \leftarrow Q_1 \times S \]
\[ Q_3 \leftarrow \pi_Y(Q_2 - R) \]
\[ T \leftarrow Q_1 - Q_3 \]
Content

- Introduction
- Relational algebra
- Set operations
- Selection
- Projection
- Cartesian product
- Join operation
- Division operation

- Other operations
  - Aggregation operators
  - Grouping
  - Outer join

- Update operations
Aggregation operators

- Input: the collections of values from the DB
- Output: a single value
- Include
  - AVG
  - MIN
  - MAX
  - SUM
  - COUNT
Aggregation operators

- Example

<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

SUM(B) = 10
AVG(A) = 1.5
MIN(A) = 1
MAX(B) = 4
COUNT(A) = 4
Grouping

- Is used to consider a relation in groups, corresponding to the value of columns

- Denotation

\[ G_1, G_2, \ldots, G_n \mathcal{F}_{F_1(A_1), F_2(A_2), \ldots, F_n(A_n)}(E) \]

- \( E \) is relational algebra expression
- \( G_1, G_2, \ldots, G_n \): grouping attributes
- \( F_1, F_2, \ldots, F_n \): aggregation operators
- \( A_1, A_2, \ldots, A_n \): aggregated attributes
Groping

Example

<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>4</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>γ</td>
<td>2</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

\[ \mathcal{F}_{\text{SUM}(C)}(R) \]

<table>
<thead>
<tr>
<th>SUM_C</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
</tr>
</tbody>
</table>

\[ A \mathcal{F}_{\text{SUM}(C)}(R) \]

<table>
<thead>
<tr>
<th>A</th>
<th>SUM_C</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>14</td>
</tr>
<tr>
<td>β</td>
<td>3</td>
</tr>
<tr>
<td>γ</td>
<td>10</td>
</tr>
</tbody>
</table>
Example 17

- The number of employees and the average salary of the company
Example 18

- For each department, find the number of employees and the average salary
Example 19

- Find the name of departments that have the largest number of employees
Example 20

- Find the name of employees who work the largest number of projects
Outer join

- Is used to avoid the loss of information
  - A theta join is taken first
  - Then, the tuples that failed to join with any tuple of the other relation are added to the result

- Three cases
  - Left outer join
  - Right outer join
  - Left and right outer join
Example 21

- List out the name of employees and the name of department that they are the manager if any

\[
\text{R1} \leftarrow \text{EMPLOYEE} 
\quad \text{SSN=MGRSSN} 
\quad \text{DEPARTMENT} 
\]

\[
\text{RESULT} \leftarrow \pi_{\text{FNAME}, \text{LNAME}, \text{DNAME}} (\text{R1}) 
\]

<table>
<thead>
<tr>
<th>FNAME</th>
<th>LNAME</th>
<th>DNAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tung</td>
<td>Nguyen</td>
<td>Nghien cuu</td>
</tr>
<tr>
<td>Hang</td>
<td>Bui</td>
<td>null</td>
</tr>
<tr>
<td>Nhu</td>
<td>Le</td>
<td>null</td>
</tr>
<tr>
<td>Vinh</td>
<td>Pham</td>
<td>Quan ly</td>
</tr>
</tbody>
</table>
Example 22

- List out the name of departments and the number of employees of that department

*If a department has just been established and not yet been arranged the employees, then what will be the result?*
Content

- Introduction
- Relational algebra
- Set operations
- Selection
- Projection
- Cartesian product
- Join operation
- Division operation
- Other operations
- Update operations
Update operations

- The content of the database can be updated by update operations
  - Insertion
  - Deletion
  - Update

- These operations are expressed by an assignment operation

\[ R_{\text{new}} \leftarrow \text{operations on } R_{\text{old}} \]
**Insertion operation**

- **Is expressed**

\[ R_{\text{new}} \leftarrow R_{\text{old}} \cup E \]

  - \( R \) is a relation
  - \( E \) is a relational algebra expression

- **Example**

  - Assign the employee whose SSN is 123456789 the project with SSN is 20 and the number of working hours is 10

\[ \text{WORKS\_ON} \leftarrow \text{WORKS\_ON} \cup (\text{‘123456789’}, 20, 10) \]
Deletion operation

- Is expressed

\[ R_{\text{new}} \leftarrow R_{\text{old}} - E \]

- \( R \) is a relation
- \( E \) is a relational algebra expression

- Example

- Delete all work assignments of the employee 123456789

\[ \text{WORKS\_ON} \leftarrow \text{WORKS\_ON} - \sigma_{\text{SSN}=\text{‘123456789’}}(\text{WORKS\_ON}) \]
Example 23

- Remove work assignments that have locations in ‘Ha Noi’
Update operation

- Is expressed

\[ R_{\text{new}} \leftarrow \pi_{F_1, F_2, \ldots, F_n} (R_{\text{old}}) \]

- \( R \) is a relation
- \( F_i \) is an arithmetic expression that results in the new value for attributes

- Example
  - Increase working hours to 1.5 times for all employees

\[ \text{WORKS\_ON} \leftarrow \pi_{\text{SSN, PNO, HOURS*1.5}} (\text{WORKS\_ON}) \]
Example 24

- Increase working hours to 1.5 times for assignments that are over 30 hours, the remain will be increased to 2 times