Database Management System – 18 (Relational Algebra – Unary Relational operations)

Ajay James
Asst. Prof in CSE
Government Engineering College Thrissur

Outline

- Relational Algebra
- Why Relational Algebra
- SELECT
- PROJECT
- RENAME

Relational Algebra

- Basic set of operations for the relational model is known as the ***relational algebra***
- Operations enable a user to specify basic retrieval requests
- Result of a retrieval is a new relation
- **Algebra operations** produce new relations
- Sequence of relational algebra operations forms a **relational algebra expression**
- *Unary Relational Operations*
- *Operations from set theory*

Why relational algebra?

- Provides a ***formal foundation*** for relational model operations
- Used as a basis for implementing and ***optimizing queries***
- Some of its concepts are incorporated into the ***SQL*** standard query language for RDBMSs
  - Core operations and functions in the internal modules of most relational systems are based on relational algebra operations

Unary operation - SELECT

- SELECT operation is used to select a *subset* of the tuples from a relation that satisfy a **selection condition**
- A filter that keeps only those tuples that satisfy a qualifying condition
- Select operation is denoted by
  - $\sigma_{\text{<selection condition>}}(R)$
  - σ (sigma) is used to denote the select operator
  - selection condition is a Boolean expression specified on the attributes of relation R
- To select the EMPLOYEE tuples whose department number is four

  $\sigma_{DNO = 4}(EMPLOYEE)$
- To select those whose salary is greater than $30,000

  $\sigma_{SALARY > 30,000}(EMPLOYEE)$

Example

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

$\sigma_{(Dno=4\ \textbf{AND}\ Salary>25000)\ \textbf{OR}\ (Dno=5\ \textbf{AND}\ Salary>30000)}(EMPLOYEE)$

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5

SELECT Operation Properties

- SELECT operation $\sigma_{<\text{selection condition}>}(R)$ produces a relation S that has the ***same schema*** as R
- SELECT operation σ is ***commutative***

  $\sigma_{<\text{condition1}>}(\sigma_{<\text{condition2}>}(R)) = \sigma_{<\text{condition2}>}(\sigma_{<\text{condition1}>}(R))$
- Cascaded SELECT operation **may be applied in any order**

  $\sigma_{<\text{condition1}>}(\sigma_{<\text{condition2}>}(\sigma_{<\text{condition3}>}(R)) = \sigma_{<\text{condition2}>}(\sigma_{<\text{condition3}>}(\sigma_{<\text{condition1}>}(R)))$
- Cascaded SELECT operation may be replaced by a single selection with a conjunction of all the conditions

  $\sigma_{<\text{condition1}>}(\sigma_{<\text{condition2}>}(\sigma_{<\text{condition3}>}(R)) = \sigma_{<\text{condition1}> \text{ AND } <\text{condition2}> \text{ AND } <\text{condition3}>}(R))$

Unary operation - PROJECT

- Selects certain *columns* from the table and discards the other columns
- Creates a vertical partitioning
- To list each employee's first and last name and salary, the following is used:

  $\pi_{\text{LNAME, FNAME,SALARY}}(\text{EMPLOYEE})$
- General form of the project operation is

  $\pi_{<\text{attribute list}>}(R)$

  π (pi) is the symbol used to represent the project operation

  <attribute list> is the desired list of attributes from the attributes of relation R
- Project operation *removes any duplicate tuples*

Example

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

$\pi_{Lname, Fname, Salary}(EMPLOYEE)$.

Lname	Fname	Salary
Smith	John	30000
Wong	Franklin	40000
Zelaya	Alicia	25000
Wallace	Jennifer	43000
Narayan	Ramesh	38000
English	Joyce	25000
Jabbar	Ahmad	25000
Borg	James	55000

$\pi_{Sex, Salary}(EMPLOYEE)$.

Sex	Salary
M	30000
M	40000
F	25000
F	43000
M	38000
M	25000
M	55000

PROJECT Operation Properties

- Number of tuples in the result of projection $\pi_{\text{<list>}}(R)$ is always less or equal to the number of tuples in R
- If the list of attributes includes a key of R, then the number of tuples is equal to the number of tuples in R
- $\pi_{\text{<list1>}}(\pi_{\text{<list2>}}(R)) = \pi_{\text{<list1>}}(R)$ as long as <list2> contains the attributes in <list2>

Unary operation - RENAME

- Apply several relational algebra operations one after the other
  - Write the operations as a single **relational algebra expression** by nesting the operations
  - Apply one operation at a time and create **intermediate result relations**
- To retrieve the first name, last name, and salary of all employees who work in department number 5
- $\pi_{\text{FNAME, LNAME, SALARY}}(\sigma_{\text{DNO=5}}(\text{EMPLOYEE}))$
- OR
- $\text{DEP5\_EMPS} \leftarrow \sigma_{\text{DNO=5}}(\text{EMPLOYEE})$
- $\text{RESULT} \leftarrow \pi_{\text{FNAME, LNAME, SALARY}}(\text{DEP5\_EMPS})$

Example

$\pi_{\text{Fname, Lname, Salary}}(\sigma_{\text{Dno=5}}(\text{EMPLOYEE}))$

Fname	Lname	Salary
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000
Joyce	English	25000

Example

$TEMP \leftarrow \sigma_{Dno=5}(EMPLOYEE)$

$R(First\_name, Last\_name, Salary) \leftarrow \pi_{Fname, Lname, Salary}(TEMP)$

TEMP

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston,TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston,TX	M	40000	888665555	5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble,TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5

R

First_name	Last_name	Salary
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000
Joyce	English	25000

Rename Operator

- Rename operator is ρ
- Can be expressed by any of the following forms:
  - $\rho_{S(B_1, B_2, \ldots, B_n)}(R)$ is a renamed relation S based on R with column names $B_1, B_1, \ldots B_n$.
  - $\rho_S(R)$ is a renamed relation S based on R (which does not specify column names)
  - $\rho_{(B_1, B_2, \ldots, B_n)}(R)$ is a renamed relation with column names $B_1, B_1, \ldots B_n$ which does not specify a new relation name.

Reference

- Elmasri R. and S. Navathe, Database Systems: Models, Languages, Design and Application Programming, Pearson Education 6th edition and 7th edition

Thank you