SIM GLOBAL EDUCATION

UNIVERSITY OF WOLLONGONG AUSTRALIA

CSCI317 – Database Performance Tuning

Relational Algebra

sjapit@uow.edu.au

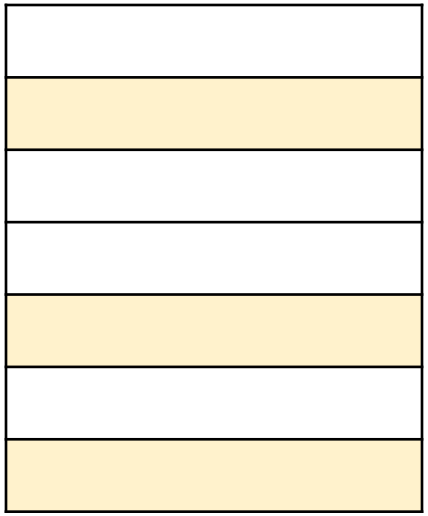
18 July 2022

Relational Algebra – An Introduction

- Relational algebra operations work on one or more relations to define another relation and keeping the original relations unchanged.
- The operands to the relational algebra operators and their output are relations. Hence, relational algebra operators can be pipelined.
- Relational algebra is useful for representing execution plans.

Relational Algebra Operations

- Five basic operations in relational algebra:
  - **Selection (σ)** - Selects a subset of rows from a relation. Works on a single relation R and defines a relation that contains only those tuples (rows) of R that satisfy the specified condition (predicate).



$$\boldsymbol{Notation} - \sigma_p(r)$$

Where $\boldsymbol{\sigma}$ stands for selection predicate and $\boldsymbol{r}$ stands for relation, and $\boldsymbol{p}$ is prepositional logic formula which may use connectors like **and**, **or**, and **not**.

Selection – An Example

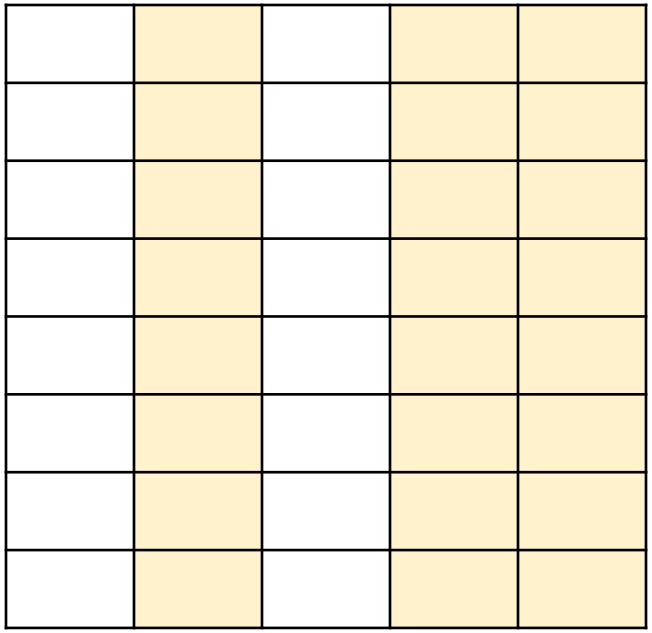
- List all staff with a salary greater than £10,000.

$$\sigma_{salary > 10000}(Staff)$$

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24- Mar-58	18000	B003
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003

Relational Algebra Operations

- **Projection** $(\boldsymbol{\pi}$ ***or*** $\sqcap)$ - Picking a subset of columns from a relation. Works on a single relation R and defines a relation that contains a vertical subset of R, extracting the values of specified attributes and eliminating duplicates.



$\boldsymbol{Notation} - \sqcap_{A_1, A_2, \cdots, A_n} (r)$

Where A_1, A_2, A_n are attribute names of a relation r.

Projection – An Example

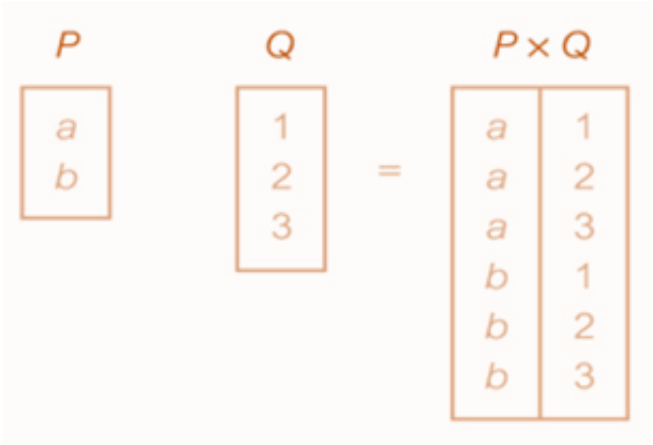
- Produce a list of salaries for all staff, showing only staffNo, fName, lName, and salary details.

$$\Pi_{staffNo,\ fName,\ lName,\ salary}(Staff)$$

staffNo	fName	lName	salary
SL21	John	White	30000
SG37	Ann	Beech	12000
SG14	David	Ford	18000
SA9	Mary	Howe	9000
SG5	Susan	Brand	24000
SL41	Julie	Lee	9000

Relational Algebra Operations

- **Cross-product** $(\times)$ - Combining two relational instances. The two relational instances may or may not have the same type of tuples. In other words, the two relational instances do not need to be UNION compatible.



$$Notation - p \times q$$

Where p and q are relations and their output will be defined as - $p \times q = \{st | s \in p \; and \; t \in q\}$

Client

clientNo	fName	lName
CR76	John	Kay
CR56	Aline	Steward
CR74	Mike	Ritchie
CR62	Mary	Tregear

Viewing

clientNo	propertyNo	comment
CR56	PA14	Too small
CR76	PG4	Too remote
CR56	PG4	
CR62	PA14	No dining room
CR56	PA14	Too small

Cross-product – An Example

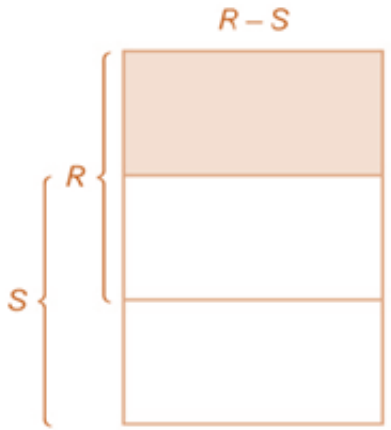
- **List the names and comments of all clients who have viewed a property for rent.**

$$(\Pi_{clientNo,\ fName,\ lName}(Client)) \times (\Pi_{clientNo,\ propertyNo,\ comment}(Viewing))$$

client.clientNo	fName	lName	Viewing.clientNo	propertyNo	comment
CR76	**John**	**Kay**	**CR56**	**PA14**	**too small**
CR76	**John**	**Kay**	**CR76**	**PG4**	**too remote**
CR76	John	Kay	CR56	PG4	
CR76	John	Kay	CR62	PA14	no dining room
CR76	John	Kay	CR56	PG36	
CR56	Aline	Stewart	CR56	PA14	too small
CR56	Aline	Stewart	CR76	PG4	too remote
CR56	Aline	Stewart	CR56	PG4	
CR56	Aline	Stewart	CR62	PA14	no dining room
CR56	Aline	Stewart	CR56	PG36	
CR74	Mike	Ritchie	CR56	PA14	too small
CR74	Mike	Ritchie	CR76	PG4	too remote
CR74	Mike	Ritchie	CR56	PG4	
CR74	Mike	Ritchie	CR62	PA14	no dining room
CR74	Mike	Ritchie	CR56	PG36	
CR62	Mary	Tregear	CR56	PA14	too small
CR62	Mary	Tregear	CR76	PG4	too remote
CR62	Mary	Tregear	CR56	PG4	
CR62	Mary	Tregear	CR62	PA14	no dining room
CR62	Mary	Tregear	CR56	PG36	

Relational Algebra Operations

- **Set-difference** $(-)$ - Tuple is in relation 1, but not in relation 2. It defines a relation consisting of the tuples that are in relation 1, but not in relation 2.



$Notation\ -r-s$

Set-different – An Example

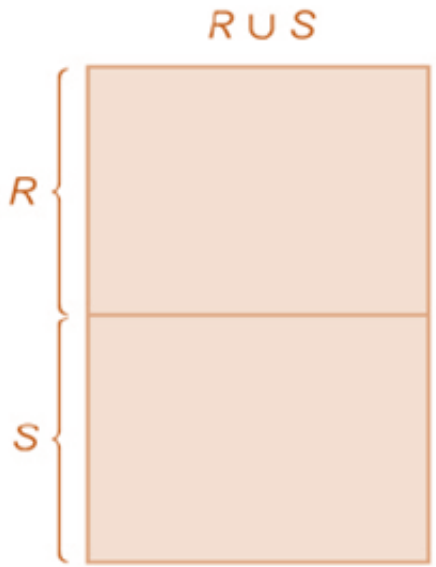
- **List all cities where there is a branch office but no properties for rent.**

$$\Pi_{city}(Branch) - \Pi_{city}(PropertyForRent)$$

city
Bristol

Relational Algebra Operations

- **Union** (∪) - Tuples are in relation 1 and in relation 2. The Union of two relations r and s defines a relation that contains all the tuple of r, or s, or both r and s, duplicate tuples being eliminated.



$Notation\ -\ r \cup s$

Where r and s are either database relations or relation result set (temporary relation).

Union – An Example

- **List all cities where there is either a branch office or a property for rent.**

$$\Pi_{city}(Branch) \cup \Pi_{city}(PropertyForRent)$$

city
London
Aberdeen
Glasgow
Bristol

Relational Algebra Operations

- Additional operations:
  - Intersection, join, division, renaming: Not essential, but is useful.

Relational Algebra Operations

- **Join** (⋈) - Join is a derivative of cross-product. That is, it is equivalent to performing a Selection, using join predicate as selection formula, over cross-product of the two operand relations.
- Various forms of join operations
  - Theta join
  - Equijoin (a particular type of Theta join)
  - Natural join
  - Outer join
  - Semijoin

Relational Algebra Operations

- **Theta join** $(\bowtie)$ **-** Defines a relation that contains tuples satisfying the predicate F from the cross product of R and S.
- The predicate F is of the form $R.a_i \Theta S.b_i$ where Θ may be one of the comparison operators $(<, \leq, >, \geq, =, \neq)$.
- Can rewrite Theta join using basic Selection and Cross product operations.

$$R \bowtie_F S = \sigma_F(R \times S)$$

Client

clientNo	fName	lName
CR76	John	Kay
CR56	Aline	Steward
CR74	Mike	Ritchie
CR62	Mary	Tregear

Viewing

clientNo	propertyNo	comment
CR56	PA14	Too small
CR76	PG4	Too remote
CR56	PG4	
CR62	PA14	No dining room
CR56	PA14	Too small

Theta join – An Example

$$R \bowtie_F S = \sigma_F(R \times S)$$

$$\sigma_{Client.clientNo = Viewing.clientNo}((\prod_{clientNo,\, fName, lName}(Client)) \times (\prod_{clientNo, propertyNo, comment}(Viewing)))$$

Client.clientNo	fName	lName	Viewing.clientNo	propertyNo	comment
CR76	John	Kay	CR76	PG4	Too remote
CR56	Aline	Steward	CR56	PA14	Too small
CR56	Aline	Steward	CR56	PG4	
CR56	Aline	Steward	CR56	PG36	
CR62	Mary	Tregear	CR62	PA14	No dining room

Relational Algebra Operations

- **Equijoin** $(\bowtie)$ **-** Defines a relation that contains tuples satisfying the predicate F from the cross product of R and S.
- The predicate F is of the form $R.a_i \Theta S.b_i$ where Θ may be one of the comparison operators $(<, \leq, >, \geq, =, \neq)$.
- If predicate F contains only equality (=), the term Thetabjoin is the same as Equijoin.

$$R \bowtie_F S = \sigma_F(R \times S)$$

Equijoin – An Example

- List the names and comments of all clients who have viewed a property for rent.

$$\left(\Pi_{clientNo,fname,lName}\ (Client)\right)_{Client.clientNo=Viewing.clientNo}$$
$$\left(\Pi_{clientNo,propertyNo,comment}\ (Viewing)\right)$$

Client.clientNo	fName	lName	Viewing.clientNo	propertyNo	comment
CR76	John	Kay	CR76	PG4	Too remote
CR56	Aline	Steward	CR56	PA14	Too small
CR56	Aline	Steward	CR56	PG4	
CR56	Aline	Steward	CR56	PG36	
CR62	Mary	Tregear	CR62	PA14	No dining room

Relational Algebra Operations

- **Natural join** $(\bowtie)$ **-** An Equijoin of the two relations R and S over all common attributes x. One occurrence of each common attribute is eliminated from the result.

Natural Join – An Example

- List the names and comments of all clients who have viewed a property for rent.

$$\left(\Pi_{clientNo,fName,lName}\ (Client)\right) \bowtie \left(\Pi_{clientNo,propertyNo,comment}(Viewing)\right)$$

clientNo	fName	lName	propertyNo	comment
CR76	John	Kay	PG4	Too remote
CR56	Aline	Steward	PA14	Too small
CR56	Aline	Steward	PG4	
CR56	Aline	Steward	PG36	
CR62	Mary	Tregear	PA14	No dining room

Relational Algebra Operations

- Outer join (⟕) - (left) outer join is join in which tuples from R that do not have matching values in common columns of S are also included in result relation.
- Outer join is used to display rows in the result that do not have matching values in the join column.

Outer join – An Example

- Produce a status report on property viewings.

$$\Pi_{propertyNo,street,city}\ (PropertyForRent) \bowtie Viewing$$

propertyNo	Street	City	clientNo	viewDate	comment
PA14	16 Holhead	Aberdeen	CR56	24-May-22	Too small
PA14	16 Holhead	Aberdeen	CR62	14-May-22	No dining room
PL94	6 Argyll St	London	Null	Null	Null
PG4	6 Lawrence St	Glasgow	CR76	20-Apr-22	Too remote
PG4	6 Lawrence St	Glasgow	CR56	26-May-22	
PG36	2 Manor Rd	Glasgow	CR56	28-May-22	
PG21	18 Dale Rd	Glasgow	Null	Null	Null
PG16	5 Novar Dr	Glasgow	Null	Null	Null

Semijoin

- **Semijoin** $(\triangleright_F)$ - Defines a relation that contains the tuples of R that participate in the join of R with S.
- Can rewrite Semijoin using ProjectFon and Join.

$$R \triangleright_F S = \Pi_A (R \bowtie_F S)$$

Semijoin – An Example

- List complete details of all staff who work at the branch in Glasgow.

$$Staff \triangleright_{Staff.branchNo=Branch.branchNo} \left(\sigma_{city='Glasgow'}(Branch)\right)$$

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003