RELATIONAL DATA LANGUAGES

Part 1

Relational Querying

- Relational model helps in simple and powerful data retrieval
- Output of query modelled as a relation.
- Based on formal mathematical model.
  - *First order predicate Logic*
  - *Eg : Book('B101')*
- Allows for much optimization

Relational Data Languages

- Manipulation and Retrieval of data
- Two Types of Query Languages
  - *Relational Algebra*
    - Procedural
    - Set of operators operating on relations
  - *Relational Calculus*
    - NonProcedural
    - Users describe what they want rather than how to compute

Formal Relational Query Languages

- **Procedural** – (Relational Algebra)
  - *User specifies what data is required and how to get those data*
  - *Operational*
  - *Execution plans can be represented*
- **Nonprocedural** – (Relational calculus)
  - *User specifies what data is required without specifying how to get those data*
  - *Declarative*
  - *Query semantics can be represented*
- SQL is the most widely used query language based on Relational Algebra

Relational Algebra

- Algebra ?
- Operands
  - Variables or values from which new values can be constructed.
- Operators
  - *Symbols denoting procedures to construct new values from Operands.*
- Relational Algebra has relations as operands and set operations as operators.
- Satisfies Closure property
  - *Output of an operation on relations is a relation itself*
  - *Operations can be composed.*

Types of Operations

- Unary Relational Operations
  - *Select (**σ**)*
  - *Project (**Π**)*
  - *Rename (ρ)*
- Binary Relational Operations
  - *Join* ⋈
    - natural,semi,Θ-join
  - *Division (÷)*

- Set theory Operations
  - *Union (**U**)*
  - *Intersection (∩)*
  - *Difference (**-**)*
  - *Cartesian Product (**X**)*
- Additional Relational Operations
  - *Outer Joins*
  - *Outer Union*
  - *Aggregate Functions*
    - **Eg.** Sum,Count, Avg..

Select (denoted by σ (sigma))

- Retrieval of subset of the tuples from a relation based on a selection condition
- Selection condition acts as filter

$\sigma_{<selection\_condition>}(R)$; R is a relation

Selection condition is a boolean formula.

Tuples satisfying the condition are retained.

Ex : $\sigma_{ISBN='B110'}(Book)$

Properties of Selection operation

- $\sigma_{<selection\_condition>}(R) = S$; R and S have same schema
- Number of tuples in S <= Number of tuples in R
- Is commutative

$$\sigma_{<cn1>}(\sigma_{<cn2>}(R)) = \sigma_{<cn2>}(\sigma_{<cn1>}(R))$$

- Cascade sequence of SELECT operations may be applied in any order:

$$\sigma_{<cn1>}(\sigma_{<cn2>}(\sigma_{<cn3>}(R))) = \sigma_{<cn2>}(\sigma_{<cn3>}(\sigma_{<cn1>}(R)))$$

- Cascade equivalent to conjunction of all the conditions

$$\sigma_{<cn1>}(\sigma_{<cn2>}(\sigma_{<cn3>}(R))) = \sigma_{<cn2>AND<cn3>AND<cn1>}(R)))$$

Project (denoted by Π (pi))

- Retrieval of the subset of columns from a relation based on a specified list of attributes
- Specified lists forms a projection of attributes

$$\Pi_{<attr\_list>}(R)$$; R is a relation

All the tuples of R with only the specified attribute values are retrieved.

Ex : $\Pi_{ISBN, Title}(Book)$

Is $\Pi_{ISBN}(\sigma_{Publ\_code='P010'}(Book))$ valid ?

Which property of relational algebra ?

BS <- $\sigma_{Publ\_code='P010'}(Book)$; BS2 <- $\Pi_{ISBN}(BS)$

Properties of Projection operation

- $\Pi_{<attr\_list>}(R) = S$;
- Removes duplicate tuples. True ?
- Number of tuples in S <= Number of tuples in R

ISBN	Title	Category	Publ_code
B111	FISH	ARTICLE	P010
B112	GLOW	ARTICLE	P212
B110	FERT	NEWS	P010

$\Pi_{ISBN, Title}(Book)$

Attribute list contains Key.

$\Pi_{Publ\_code}(Book)$

Removes duplicate tuples.

Not commutative

$\Pi_{ISBN, Title}(\Pi_{category,publ\_code}(Book))$

$\Pi_{ISBN, Title}(\Pi_{category,ISBN}(Book))$

Results

ISBN	Title	Category	Publ_code
B111	FISH	ARTICLE	P010
B112	GLOW	ARTICLE	P212
B110	FERT	NEWS	P010

- $\Pi_{category,ISBN}(Book)$ $\Pi_{category}(Book)$

Category	ISBN
ARTICLE	B111
ARTICLE	B112
NEWS	B110

Category
ARTICLE
NEWS

- $\Pi_{ISBN,\ category}(\sigma_{Publ\_code='P010'}(Book))$

ISBN	Category	Publ_code
B111	ARTICLE	P010
B110	NEWS	P010

Rename (denoted by ρ (rho))

- The general RENAME operation ρ can be
  - $\rho_{S(B1,B2,\ldots,Bn)}(R)$ *changes*

  the relation name to S, *and* the column (attribute) names to B1, B2, ...Bn
- $\rho_S(R)$ changes:

  the *relation name only to S*

  $\rho_{(B1,B2,\ldots,Bn)}(R)$ changes:

  the *column (attribute) names only to B1, B2, ...Bn*

THANK YOU

References

- Silberschatz A Korth H F and SudharshanS , “Database System Concepts”, 6th Edition, TMH publishing company limited, 2011.