Relational Database manipulation

1. Relational Algebra
2. Relational Calculus
3. Standard Query Language (SQL)

Relational Algebra

Relational algebra allows expressing queries as algebraic expression, where operands are relations. In example below, sigma (σ) is an operation, employee is operand relation, and logical expression is parameter to the operation.

$$\sigma_{DNO = 4 \text{ AND } SALARY > 50000} (employee)$$

This expression outputs tuples that member of employee relation, and value of salary attribute is > 50000 and value of dno attribute is equals to 4.

Relational algebra expressions have closure property that is result of an expression is a valid relation.

Relational algebra was introduced in original proposal of relational model.

There are definite set of operations defined on relations. Some operations are unary while others are binary.

Complex queries are expressed by sequencing multiple operations.

Relational Calculus

A query in relational calculus is expressed as a logic expression; above relational algebra query is expressed in relational following is a query, expressed in tuple relational calculus, as following.

$$\{t \mid EMPLOYEE(t) \text{ AND } t.salary > 50000 \text{ AND } t.dno=4) \}$$

Below is another query that outputs specific attributes of tuple tuple t, where t is such that member of employee, and there t.dno and d.dno matches where d is tuple of department relation, and d.name is 'Research'

$$\{t.\text{Fname}, t.\text{Lname}, t.\text{Address} \mid \text{EMPLOYEE}(t) \text{ AND } (\exists d)(\text{DEPARTMENT}(d) \text{ AND } d.\text{Dname}=\text{'Research'} \text{ AND } d.\text{Dnumber}=t.\text{Dno})\}$$

It may be noted that in calculus expressions are more declarative (rather than procedural) in nature; no concept of operation and their order.

Standard Query Language (SQL)

Structured Query Language (SQL) was not part of original proposal of Relational Model. Some people pronounce SQL as "SEQUEL".

Originally developed at IBM with System R. System R was first implementation of Relational Model in Early 70s.

SQL has evolved since then, and now a standard language for relational databases.

ANSI has published SQL-86, SQL-89, SQL-92, SQL-99, SQL-2003, SQL-2006, and SQL-2008, 2013 or so. SQL that discussed in most texts is SQL-99.

Despite ANSI SQL standard, RDBMS often do some kind of deviation to it. Postgres's SQL is probably most ANSI compliant.

SQL provides following categories of commands:

- Data Definition Language (DDL)
  - Used for defining database schema.
- Data Manipulation Language (DML)
  - Update commands
  - Querying commands
- Physical level commands
  - Used for improving performance of database.
- Transaction Control
  - Begin transaction, commit transaction, rollback etc
- Authorization
  - Grant permission to other users for access (read/update /delete, etc.)

Note: relational algebra and calculus do not allow expressing data definitions, and data update operations. SQL being the language that is used in practice, it provides full stack of operations.

SQL was initially influenced by tuple relation calculus, however later updates incorporated most algebraic concepts as well. We will learn most of SQL in algebraic perspective.