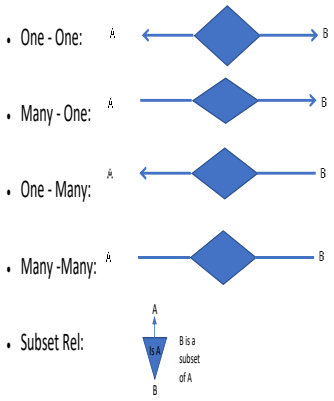
DBMS Concepts

Entity Relation Model

- Entity: Distinguishable from other objects
- Entity Set: Group of Entities
- Record: Each Row of Dataset
- Attributes: Column-ID



Models

- **Network:** ER model restricted to Binary, Many-One Relationships, To break a many many relation insert temp between them and insert 2 many-one relations
- **Hierarchical:** Kind of Decision Tree

Armstrong's Axioms

1. If $Y \subset X \subset U$ then $X \to Y$, reflexive
2. If $X \to Y$ and $Z \subset U$ then $XZ \to YZ$, augmented
3. If $X \to Y$ and $Y \to Z$ then $X \to Z$, transitive

Inference Rules

1. $\{X \to Y, X \to Z\}$ then $X \to YZ$, Union Rule
2. $\{X \to Y, WY \to Z\}$ then $XW \to Z$
3. $\{X \to Y, Z \subset Y\}$ then $X \to Z$
4. $\{X \to YZ\}$ then $X \to Y, X \to Z$

Keys

- **Super:** A combination of fields by which a row is uniquely identified. If u add any column to primary key it becomes a super key.
- **Primary key:** Any key which can uniquely identify rows of the current table
- **Candidate key:** The minimal super key. *Individual* columns in a table which qualifies for uniqueness of all rows
- **Foreign Key:** Field or collection of fields which uniquely identifies a row of another table.

3NF or Third Normal Form

$X \to A$ is in 3NF if $A \notin X$ and either X is a key of relation or A is a prime attribute.

Transformation removes Partial And Transitive Dependency

X+ Algorithm of Bernstein

Set N=0 and $X(0) = X(given)$

1. $\forall A \to B$ in f such that $A \subset X$ and $B \notin X$.then $X = X \cup B$
2. Repeat step two until no $A \to B$ such exists

Elementary and Redundant FD

$X \to Y$ is Elementary if $\nexists (X' \subset X \; and \; X' \to Y)$

F is a Redundant FD if $(F - f)^+ = F^+$

In such cases partial dependency exists

Lossless Join Property

If $r = \pi_{R1}(r) \bowtie \pi_{R2}(r)$

Lossless if $R_1 \cap R_2 \to R_1 - R_2 \; or \; R_2 - R_1 \; exists \; in \; f$

1st Normal Form(1NF)

As per the rule of first normal form, an attribute (column) of a table cannot hold multiple values. It should hold only atomic values.

Boyce Codd Normal Form(BCNF)

A relation is in BCNF *if $X \to Y$ exists & $Y \notin X$* then X is a key(candidate or super)

Prime Attribute

A attribute which is a member or a subset of the key

2nd Normal Form(2NF)

- Table is in 1NF (First normal form)
- No non-prime attribute is dependent on the proper subset of any candidate key of table.

Partial Dependency is Removed

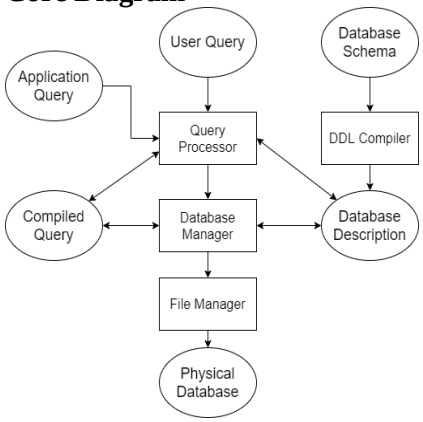
Basic Relational Algebra Operators

- **Union:** $R \cup S$ Set of Tuples either in R or S both. R,S must have same number of attributes
- **Set Difference:** $R - S$ Set of Tuples present in R but not in S. Same attributes
- **Cartesian Product:** $R \times S$, Let R,S have *arity(no of attributes)* k1,k2. Then RxS is set of tuples with arity k1+k2 with 1st k1 components form tuple of R and last k2 components tuples of S
- **Projection:** $\pi_{i_1,i_2,\ldots,i_N}(R)$ Set of tuples with attributes $i_1, i_2, \ldots, i_N$ with same content with that of R
- **Selection:** $\sigma_F(R)$ set of tuples satisfying criteria F {Select C1,C2 from T1 *where* C3>2}

Additional Relational Algebra Operators

- **Intersection:** $R \cap S = R - (R - S)$ Set of Tuples both in R or S both. R,S must have same number of attributes
- **Quotient:** $R \div S$ let arity of R is r and S is s and $r > s$, set of tuples 't' with arity $r - s$ st $\forall$ tuples 'U' in S, the tuple tu(concat) is in R
- **θ Join:** $R \bowtie_{i\theta j} S$, Set of Tuples that is in the subset of $R \times S$ satisfying $i\theta j$
- **Natural Join:** $R \bowtie S$, Subset of $R \times S$ in which only those tuples with same values for same attributes in R,S are represented in same column

Core Diagram



Terminologies

- **Cardinality:** No of Tuples
- **Arity:** No of Attributes
- **Tuples:** Represents Rows

Type of Joins

- **Natural Join:** (or Inner Join) when tuples with same values for same attributes in R,S are represented in the same column
- **Left Join:** Takes values of all tuples of left join, wherever same values for same attributes in R,S are represented in the same column, if there doesn't exists a value in R, it will print NULL
- **Right Join:** here R and S is interchanged from