**MySQL:**

DISTINCT (to remove duplicates)

% = any string

\_ = one character

Set operators:

INTERSECT, UNION, EXCEPT

- same schema for operands

- based on set semantics

- keep duplicates with ‘ALL’

Subqueries:

- can rewrite them as long as we don’t have negation

- with negation, we need EXCEPT

Set membership (IN, NOT IN)

- a IN R is TRUE is a is in R

Set comparision (> ALL, < SOME, )

- ALL is , SOME is

- <> ALL NOT IN, = SOME IN

EXISTS and Correlated subqueries

- correlated- outer query looks 1 tuple at a time and binds the tuple to S

- for each S, we execute the inner query and check condition

- this is just interpretation

- Subqueries in FROM clause

- considered as a regular relation

- must be renamed inside

Aggregates

- SUM, AVG, COUNT, MIN, MAX

- COUNT (\*) counts tuples (NULLs)

GROUPBY- duplicates are removed

HAVING - used with aggregates

ORDER BY … [ASC/DESC]

- default ASC, used for looks

Order (by computer): FROM -> WHERE -> GROUP BY -> HAVING -> ORDER BY -> SELECT

NULL

- aggregates are computed ignoring NULL, except COUNT (\*)

- if input to an aggregate is empty, COUNT returns 0, all others return NULL

Set operators (

- NULL is treated like other values here

- check NULL: IS NULL, IS NOT NULL

Arithmetic op’s and comp’s

- for NULL’s, UNKNOWN, only true values returned

Three-valued logic:

Truth table:

- AND: U & T = U, U & F = F, U & U = U

- OR: U | T = T, U | F = U, U | U = U

SQL and bag semantics

- everything uses bag except set oper’s

rules for bag:

- under bag, R

**Relational Model:**

- relations = tables

- attributes = columns

- tuples = rows

- domain = type

SQL Types:

- char(n) – fixed length

- varchar(n) – var lenght

- integer – 32 bit

- decimal (5,2) – 999.99

- real,double – 32bit,64bit

- date – ‘2002-01-15’

- time – ’13:50:00’

- timestamp – ‘2002-01-15 13:50:00’

Tables:

- one primary key per table

- UNIQUE for other keys

- SQL92: no NULL in primary

- DEFAULT for default

Loading Data:

LOAD DATA INFILE <datafile> INTO TABLE <table>

Terms:

Data model = graph/tree model

Schema = table design/structure

Instance = data

Database construction steps:

1. domain analysis

2. database design: E/R model,

database design theory

3. table creation: DDL

4. load

5. query and update: DML

Three-valued logic

**Relational Algebra:**

Set: no duplicate tuples (eliminated)

Bag: allowed duplicates

SELECT

PROJECT

CROSS PRODUCT

NATURAL JOIN

THETA JOIN

RENAME

UNION

DIFFERENCE

INTERSECT

DIVISION

Core operators:

Remember:

Data manipulation language

- relation > algebra > relation

SET vs BAG semantics