Book Chapters

(4th) Chapter 4.1-6, 4.8-10, 3.3.4
(5th) Chapter 3.1-8, 3.10-11, 4.7
(6th) Chapter 3.1-9, 4.1, 4.3, 5.4, 5.5
(7th) Chapter 3.1, 3.3-5, 3.8

Things to Learn

- Basic SELECT query
- SQL set operator

SQL

- Structured Query Language
- The standard language for all commercial RDBMS
- SQL has many aspects
  - DDL: schema definition, constraints, index, ...
  - DML: query, update, ...
  - triggers, transaction, authorization, ...
- In this lecture, we cover the DML aspect of SQL
  - How to query and modify existing databases
- SQL and DBMS
  - SQL is high-level description of user’s query
    * No concrete procedure for query execution is given
  - The beauty and success of DBMS
    * The system understands the query and find the best way possible to execute it automatically
Example to Use in the Class

- School information
  - Student(sid, name, age, GPA, address, ...)
  - Class(dept, cnum, sec, unit, title, instructor, ...)
  - Enroll(sid, dept, cnum, sec)
Basic SELECT statement

- **Query 1:** Find the titles and instructors of all CS courses

- **Semantics**
  - Interpret and write FROM → WHERE → SELECT
    - FROM: the list of tables to look up
    - WHERE: conditions to meet
    - SELECT: the attributes to return
  - *Conceptual* execution (table cursor diagram)

```
    Check condition       Return attr
    Check condition       Return attr
```

General SQL statement

- SELECT A_1, ..., A_n
  FROM R_1, ..., R_m
  WHERE C
  \[\equiv \pi_{A_1, \ldots, A_n}(\sigma_C(R_1 \times \cdots \times R_m))\]

- SELECT *: all attributes

- **SELECT** is “projection” not “selection”: can be confusing

- SQL does not remove duplicates: Major difference between SQL and relational algebra
  - More examples will follow

SQL join

- **Query 2:** Find the names and GPAs of all students taking CS classes

```
    Check condition       Return attr
    Check condition       Return attr
```

- Conceptually **WHERE** R, S
 (Table join diagram)
For every pair of tuples from R and S, we check condition and produce output

Notes:

- S, E: tuple variable
  * renaming operator
  * We can consider that S and E are variables that bind to every pair of tuples
- Attributes can also be renamed
  * GPA (AS) grade
- DISTINCT: remove duplicates in the results

**WHERE conditions**

- **Query 3:** All student names and GPAs who live on Wilshire

- %: any length (0–∞) string
  \_: one character
  'Wilshire%': Any string containing Wilshire

Q: What does '___%’ mean?

- Other useful string functions: UPPER(), LOWER(), CONCAT(), ...
Set operators

- $\cap$: INTERSECT, $\cup$: UNION, $-$: EXCEPT
- Can be applied to the result of SELECT statements or to relations
- **Query 4:** All names of students and instructors

**Important points to note**

- Set operators should have the same schema for operands
  - In practice, it is okay to have just compatible types
- Set operators follow set semantics and remove duplicates
  - Set semantics is well understood for set operations. Not many people know bag semantics.
  - Efficiency
- To keep duplicates, use UNION ALL, INTERSECT ALL, EXCEPT ALL
- **Query 5:** Find ids of all students who are not taking any CS courses.

**MySQL support:**

- Standard MySQL does not support INTERSECT or EXCEPT.
- MariaDB v10.3 introduced supports for INTERSECT and EXCEPT.