Module 1

Database Systems

Connor Baker

Compiled on February 1, 2020 at 2:00pm

What are Database and DBMS?

A database

- An integrated collection of data
- Data integrity is a concern
- Models real-world enterprise
  - Entities (e.g., students, courses)
  - Relationships (e.g., Frodois taking INFS614)

A database management system (DBMS) is a software package designed to store, provide access to, and manage databases

Why use a DBMS?

In general, because they're *easier and more efficient* than directly managing a database.

They provide

- Data independence and efficient access
- Reduced application development time
- Data integrity and security
- Uniform data administration
- Concurrent access
- Recovery from crashes

Data Models

A *data model* is a collection of concepts for describing data.

A *schema* is a description of a particular collection of data, using the given data model.

The *relational model of data* is the most widely used model today.

- Main concept: a relation is a table with rows and columns
- Every relation has a schema which describes columns or fields

Levels of Abstraction

Inside a database we have multiple *views* and a single *conceptual (logical) schema* and *physical schema*.

- Views describe how users see the data
- Conceptual schema defines logical structure
- Physical schema describes the files and indices used

NB: Schemas are defined using a Data Definition Layer (DDL); data is modified/queried using Data Manipulation Layer (DML).

Data Independence

- Applications insulated from how data is structured and stored
- *Logical data independence:* Protection from changes in the logical structure of data.
- *Physical data independence:* Protection from changes in the physical structure of data.

NB: This is one of the most important benefits of using a DBMS.

Transaction

A *transaction* is the execution of a DB program – an *atomic* sequence of database actions (reads/writes).

ACID properties

- A: Atomicity
- C: Consistency
- I: Isolation
- D: Durability

ACID properties are achieved by logging and concurrency control in the subsystems of the DBMS.

Database Users

We have several different categories of users

- End users
  - DB application users

- DB application programmers
  - More precisely, they are DBMS users (like webmasters)
- Database administrator (DBA)
  - Designs logical/physical schemas
  - Handles security and authorization
  - Data availability, crash recovery
  - Database tuning as needs evolve

Summary

- DBMS used to maintain and query large datasets
- Benefits include recovery from system crashes, concurrent access, quick application development, data integrity, and security
- Levels of abstraction give data independence
- We'll learn how to
  - Set up a database
    - *Design* (ERD and Relational Model) and *refine* (Relational Normalization Theory)
  - Query the database
    - Relational Algebra and SQL