Lecture 1A (Monday, Jan 8 2018)

CS 410/510 - Databases

Welcome!

- Introduce self
- Introduce databases

Class logistics:

- **Grad students cannot re-take**

Topics

What is a database?

- Ask class
- What do you think this class is about?
- What is it?
    - A collection of data
    - About things (‘entities’)
    - To answer questions

How can we efficiently and effectively

- Store
- Retrieve
- Update
- Analyze

Information?

To do this, we need to **model** it.

- Describe the way the information is organized
- Develop a **schema** (the *shape* of the information)

The **store it**, **query it**, and **write a program to use it**

Database vs. DBMS (database management system):

- A database is the collection of data
- A DBMS is a software tool for storing and managing databases

Learning outcomes:

- **Design a data model** to organize data for storage and computation

- **Describe the data model** using standard paradigms and notation
- **Implement the model** in a relational DBMS
- **Write queries** to get data out of the database
- **Build applications** that use a database
- **Understand tradeoffs** in selecting database tech

This course is not comprehensive

- Focused on *using*, not *building*
- Only covering relational DBMS in any detail
- There's a lot we won't get to!

Database systems

- Store and process data
- Interpret queries to answer questions from data
- Why not just use a file?
    - How is the data organized?
    - Can we process it quickly?
    - Can multiple users access it?
    - Lots of hard work, someone else has done it!
- Well-understood problem

Data modeling

Let's start modeling data!

We look at three kinds of things:

- Entities
- Attributes
- Relationships

Our data models will be built up from these.

Let's start looking for entities!

- Digg
- Baseball (http://m.mlb.com/scoreboard)
- Twitter (https://twitter.com/LensKitRS)

Activity! Get with your neighbor, identify entities in a site you like

Relating entities

- Show E-R syntax
- Start drawing Twitter out

The Three-Schema Model

- User schema(s)
- Logical or conceptual schema
- Physical schema

There are some extra layers, particularly in the physical schema.

Course Logistics

If we have time.

Structure

- Lectures
- Homework assignments
- Project
- 2 midterms
- Final exam

Learning by doing

Lectures

- Will publish
    - Notes
    - Slides (if any)
    - Example code
    - Doc cam drawings
- Need volunteers to capture whiteboard

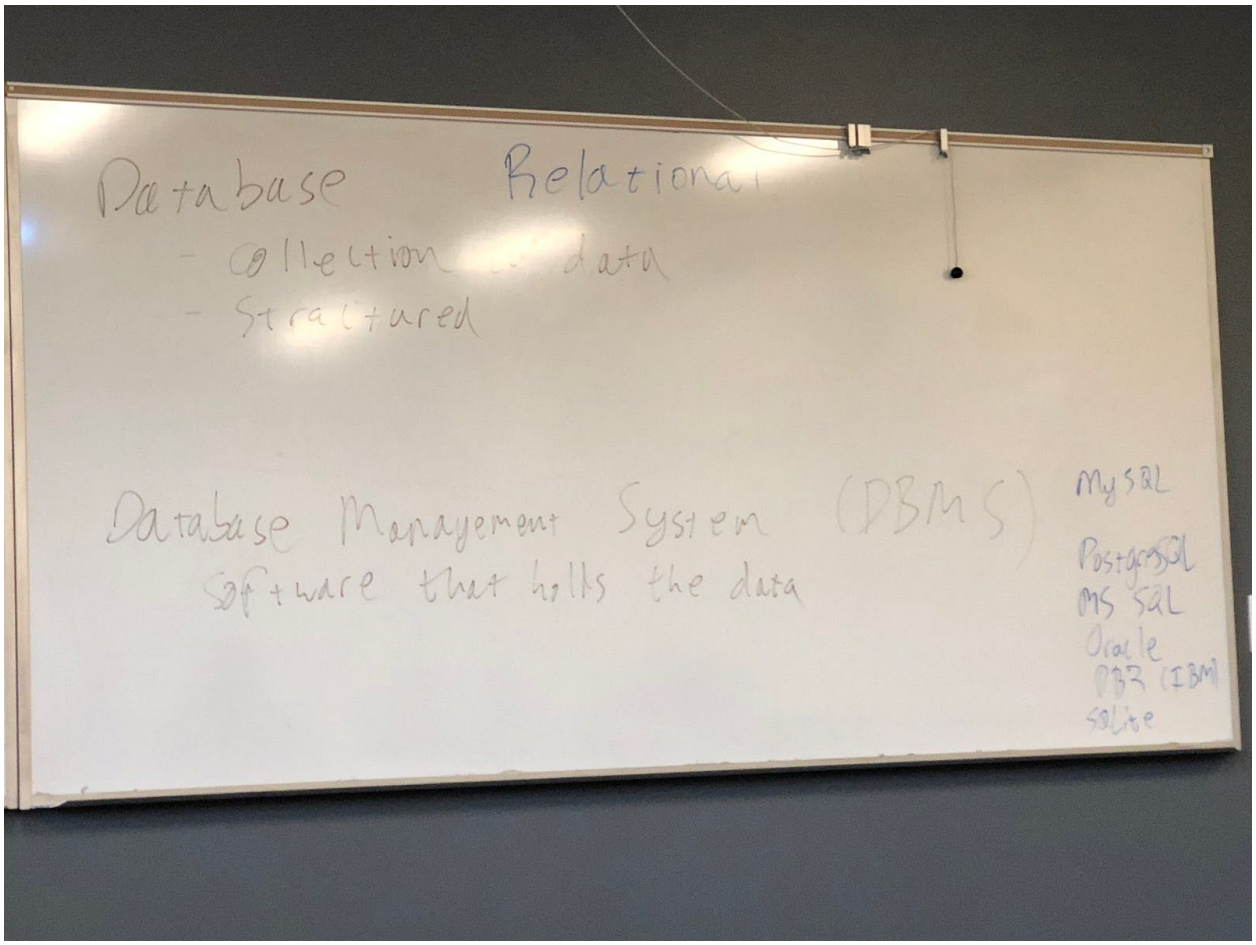
Resources

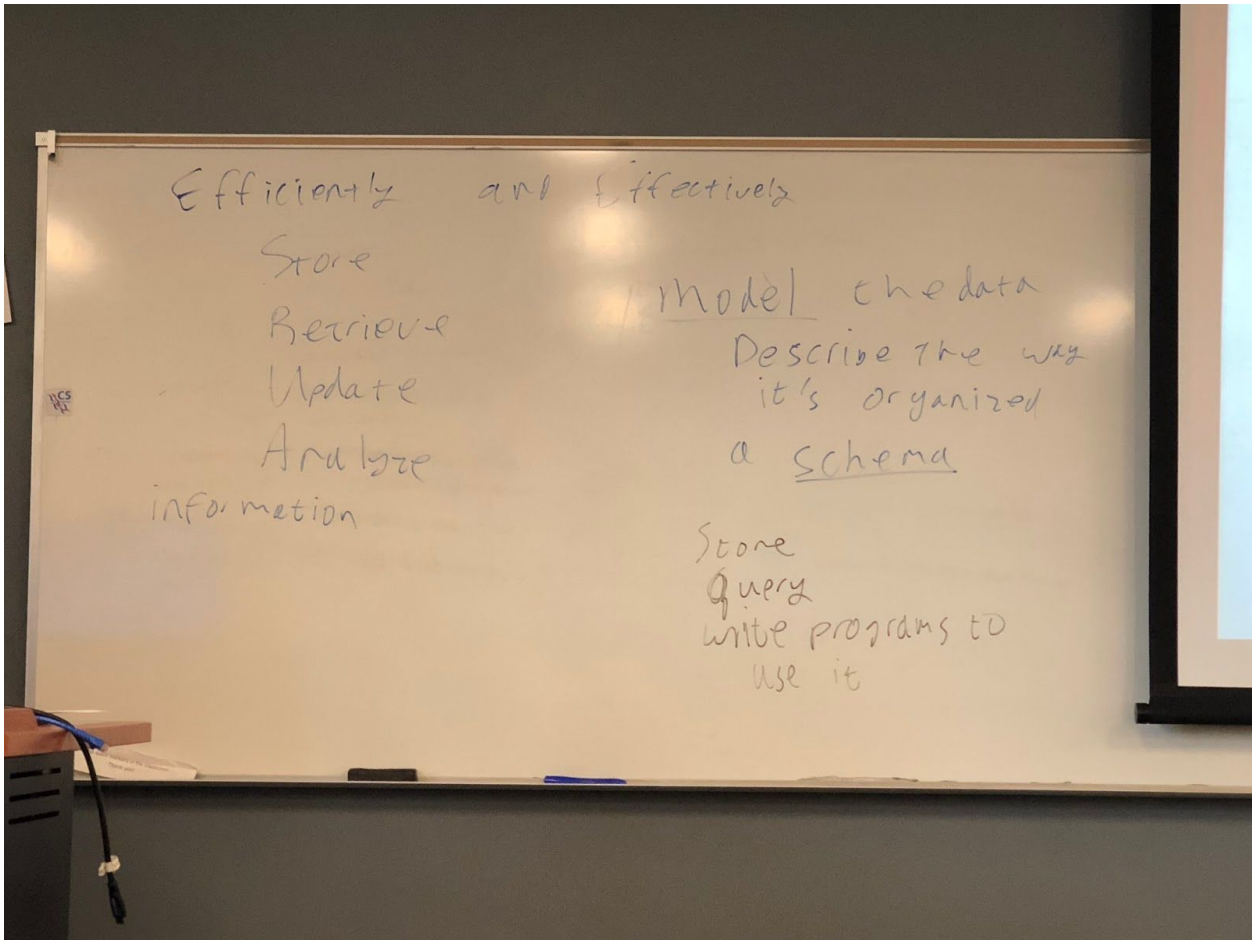
- Blackboard
- Online docs
- Textbooks

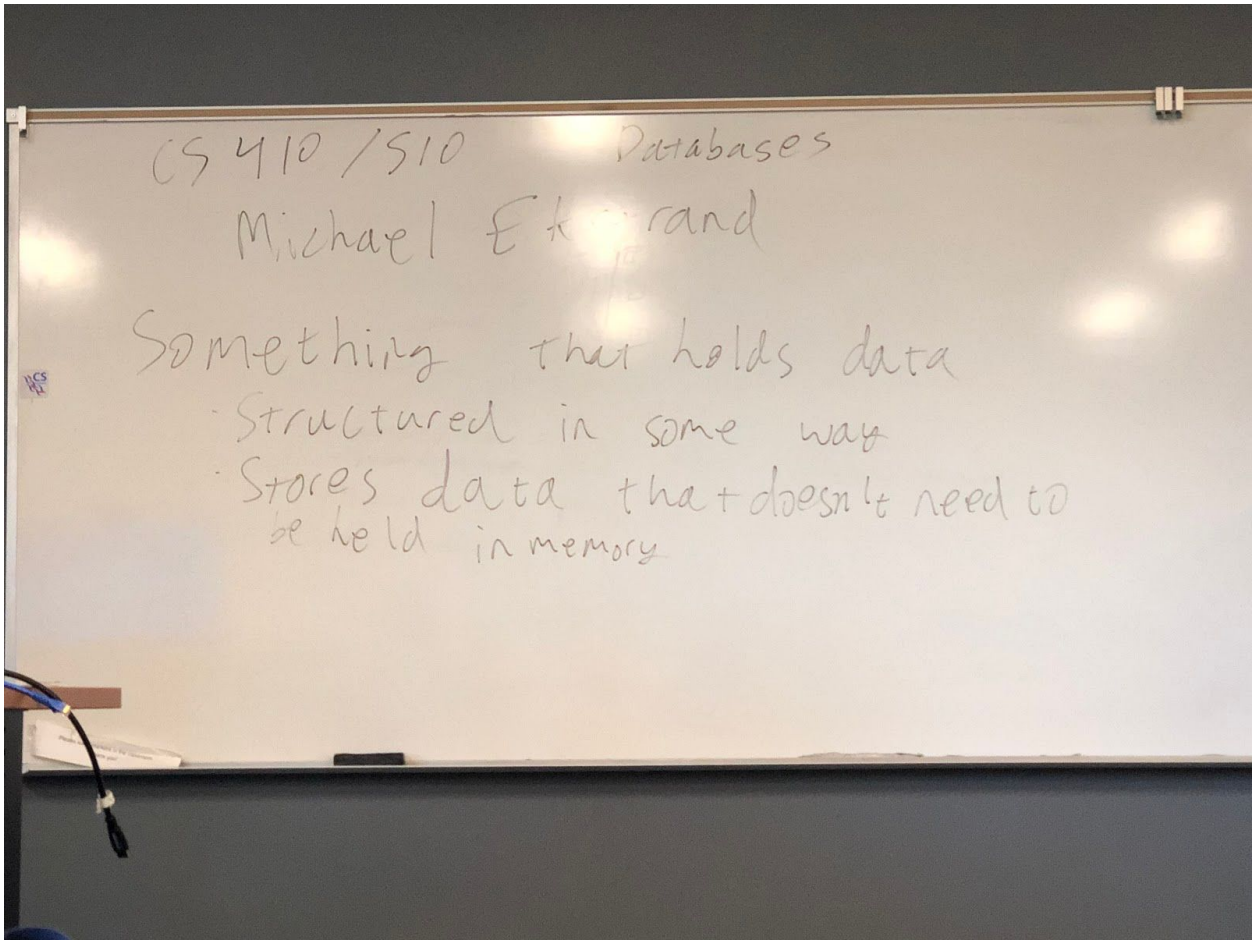
Software

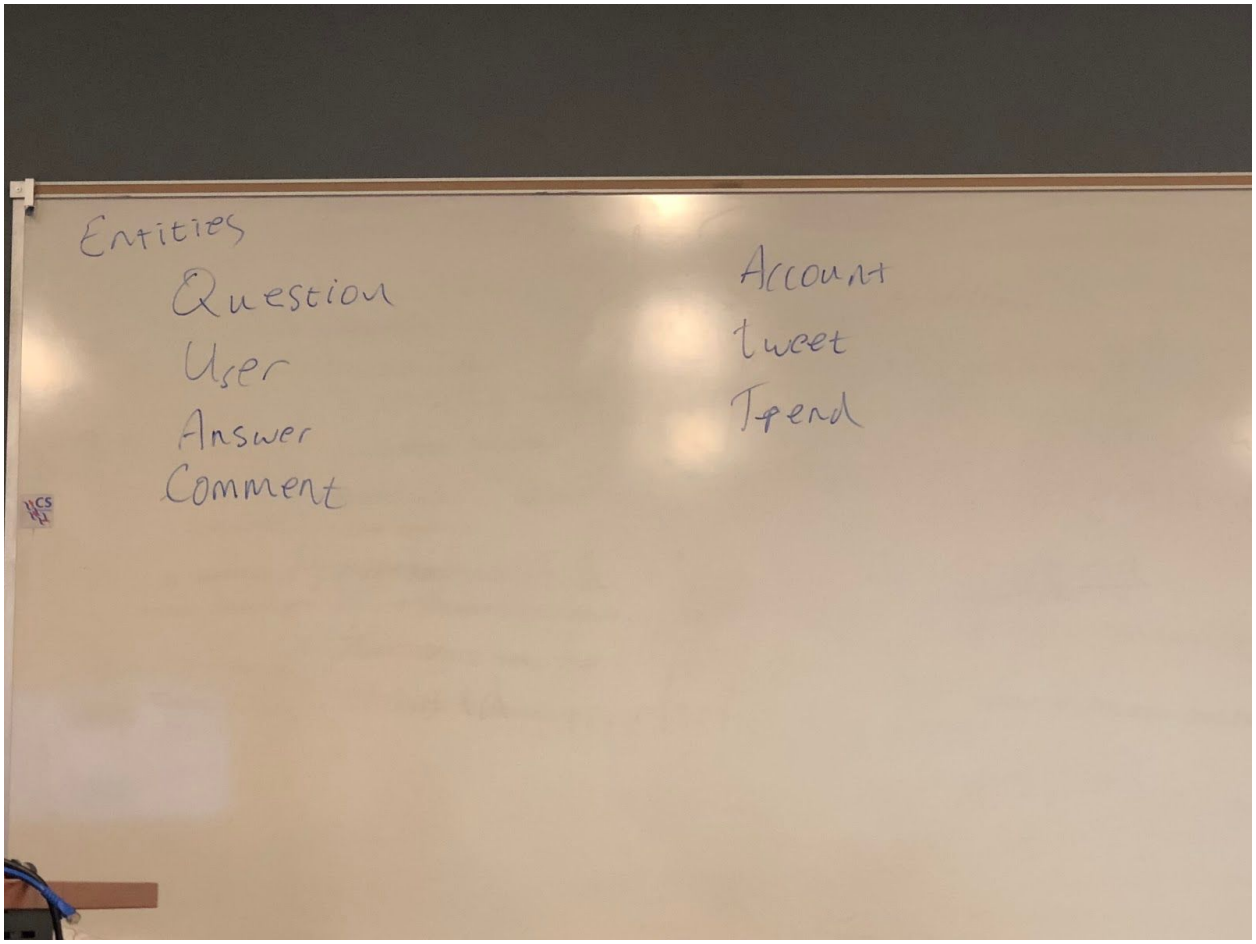
- MySQL database server
- DataGrip and IntelliJ IDEA IDEs
- Java

Pictures


Database
Relational
- Collection of data
- Structured
Database Management System (DBMS)
Software that holds the data
MySQL
PostgreSQL
MS SQL
Oracle
DB2 (IBM)
SQLite


Efficiently and Effectively
Store
Retrieve
Update
Analyze
information
model the data
Describe the way
it's organized
a schema
Store
Query
write programs to
use it


CS 410 / 510 Databases
Michael E[illegible]rand
Something that holds data
· Structured in some way
· Stores data that doesn't need to be held in memory


Entities
Question
User
Answer
Comment
Account
tweet
Trend