CMPT 354
Database Systems 1

Lecture Handout

Instructor: Eugenia Ternovska

Simon Fraser University

About Me

Office: TASC 8025

Evgenia (Eugenia) Ternovska

Associate Professor
School of Computing Science
Simon Fraser University

Education

PhD in Computer Science
University of Toronto

Research Interests

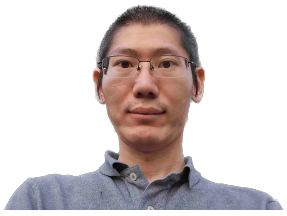
Theory of Computation
Logic in Computer Science
and in Database theory

CMPT 409:/701 Computability and Logic

Teaching Assistants

Heng Liu

MSc, Currently PhD Student,
SFU



Two more TAs

Will be introduced to you later

Website and Contacts

Course Website: SFU Canvas

- Lecture notes, homework, assignments announcements,
- Class discussions: Piazza
- Get help easily from classmates, the TAs and myself
- **Important:** if contacting us by email,
  use subject line that contains "CMPT 354"

Check Syllabus frequently for what is due each week

Make sure you get **email notifications from Canvas**

Why Do We Study Database Systems?

- Many important applications need to manage, update, and query databases (banks, fleet control, search, school system)
- Data quality, quantity, and timeliness becoming more important with AI & Machine Leaning(algorithms that generalize from data)
- Growth of user and machine generated data

Why Do We Study Database Systems?

Challenges of data-intensive systems

- Reliability
- Concurrency
- Performance
- Accessibility
- Security

What is a database?

A collection of data items, related to a specific enterprise, which is structured and organized so as to be more easily accessed, managed, and updated

What is a Database Management System (DBMS)?

- system software for creating and managing databases

Examples:

- Commercial: Oracle, IBM DB2, Microsoft SQL Server
- Open source: MySQL (Sun/Oracle), PostgreSQL, SQLite

Course Goals

Students will

- Learn to query a database
- Learn to design a database
- Learn how databases work with applications using them
- Learn working basics and important issues of Database Management Systems

Our Teaching Goals

- Teach you to be **good end-users** (SQL is not enough)
- Provide you with **solid foundations** of how a DBMS works

Two-way learning

- We are open to any feedback you might have!

Course Overview

- Introduction: Relational Database Definition, History, Applications, and Trends
- Introduction to SQL: Create, Modify, and Query Relational Databases
- Relational Data Model, and Relational Algebra (and Calculus)
- Database Design: Steps and Procedures
- Entity Relationship Model: E/R Diagram Design and Translation to Relational Databases
- Design Theory: Functional Dependencies, Normalization (1NF, 2NF, 3NF, BCNF)
- Constraints and Triggers, Views and Indexes

Course Overview (Continued)

- Database access from applications
- Query Processing and Optimization: join strategies, query plans
- Incomplete data: null values and certain answers
- Storage and indexing: B+ trees, hashing
- Scheduling and concurrency control: transaction management, serializability, locking
- Transaction Processing: Anomalies, ACID properties, Concurrency Control
- Conclusion and Further Topics to Study

Prerequisites

CMPT 225, (and MACM 101 or (ENSC 251 and ENSC 252)).

Some background in discrete mathematics:

- Basics of set theory (sets, set operations, relations, orders)
- Basics of graph theory (directed/undirected graphs, trees)
- Basics of algorithm complexity analysis (Big-O notation)
- Basics of logic (predicate logic, inference, satisfiability)
  $\Longrightarrow$ helps to understand and write correct SQL queries

Assessment

The course assessment will be based on:

- Homework Assignments (incl. mini-project) **25 %**
- Quizzes: **10 %** (in-person)
- Interactive Sessions: **5 %** (unlimited attempt quizzes)
- Midterm **20 %** (in-person)
- Final Exam: **40 %** (in-person)

Overall passing grade on the weighted average of exams (midterm & final) to obtain a clear pass (C- or better)

Bonus Marks will be given to five students who are most active on Piazza (i.e., most helpful to other students and the whole class)

Practice Exercises

They will be provided throughout the course

Goal: help you to understand the material

- Not part of the Assessment – for your own practice
- Solve, do not submit

Software: SQLite as our DBMS (database engine)

In contrast to many other database management systems (e.g., Oracle, DB2, and SQL Server),
SQLite is not a client–server database engine

SQLite is used to develop embedded software for devices like televisions, cell phones, cameras, etc.

Used worldwide for testing and development

Less functionality, e.g., does not validate datatypes

SQLite: Most Widely Used Database Engine

It is likely used more than all other database engines combined

SQLite is found in:

- Every Android device
- Every iPhone and iOS device
- Every Mac
- Every Windows10 machine
- Every Firefox, Chrome, and Safari web browser
- Every instance of Skype
- Every instance of iTunes
- Every Dropbox client
- Every TurboTax and QuickBooks
- PHP and Python
- Most television sets and set-top cable boxes
- Most automotive multimedia systems
- Countless millions of other applications

SQLite (Continued)

In SQLite, a database is stored in a single file, a feature that distinguishes it from other database engines

Small database size 2GB

Not scalable, e.g. for when multiple users need to access the database at the same time

Very simple to set up and operate

See Modules in Canvas for instructions and initial setup

Laptops

You will need to bring your laptops to quizzes and the sessions that are marked interactive in the schedule (see Syllabus).

Please check SFU libabry laptop borrowing service if necessary: https://www.lib.sfu.ca/borrow/borrow-materials/laptops-equipment/borrow-laptop

(see course web page for details)

Note: you cannot install any software on library laptops

Textbooks

[1] **Database Systems: The Complete Book**, 2nd EditionHector Garcia-Molina, Jeffrey D. Ullman, Jennifer WidomPrentice Hall

[2] **Database System Concepts**, Seventh EditionAvi Silberschatz, Henry F. Korth, S. SudarshanMcGraw-Hill, March 2019 www.db-book.com

[3] **A First Course in Database Systems** Ullman, Widom, Pearson Prentice Hall, 3d edition

[4] **Database Management Systems** Ramakrishnan, Gehrke, McGraw-Hill, 3rd edition, A broader (longer) exposition

Further reading: Abiteboul, Vianu, Hull Foundations of Databases Addison-Wesley, 1995, Mostly theoretical topics

- Out of print but freely available (for personal use only)

Acknowledgements

Additional references and resources used in preparation of this course are listed on https://canvas.sfu.ca/courses/77505/pages/references-and-resources

or mentioned in slides.