Principles of Database Systems (CS6083)

Course Prerequisites: Good programming skills, familiarity with basics of operating systems (incl. file systems, caching, locking and concurrency), and knowledge of basic algorithms and data structures (incl. sorting, heaps, search trees, and hashing). Formally, you need to have taken a course on data structures or algorithms, and a course on operating systems, at the graduate or undergraduate level, or have permission from the instructor.

Instructor: The course is taught by Prof. Torsten Suel (email `torsten.suel@nyu.edu`). Phone: 646 997-3354. Office: 10.046 in 2MTC.

Time and Location: Tuesdays from 3:20pm to 5:50pm in 370J-202. Note that this is in 370 Jay Street.

Office Hours: Tuesdays 2-3pm in 10.046, or by appointment. The course TAs will also announce separate office hours.

Course Webpage: On NYU Classes. All announcements will also be sent via the course mailing list to your NYU student email account.

Required Text: A. Silberschatz, H. Korth, S. Sudarshan, *Database System Concepts*, 6th Edition, McGraw-Hill, 2009. (You may also use the 4th or 5th edition if those are cheaper or easier to get.)

Grading Policy: Problem sets: 20%. Project: 20%. Midterm or tests: 20%. Final exam: 40%.

General Information: The course aims to give a broad introduction to relational database systems, including the relational data model, query languages, index and file structures, query processing and optimization, concurrency and recovery, transaction management, and database design, plus optional material if time permits. The workload of the course is expected to be above average, and class attendance is strongly recommended.

Policy on Academic Dishonesty: Please see the schools policy on academic dishonesty on our schools website at `https://engineering.nyu.edu/sites/default/files/2018-06/code-conduct2-2-16.pdf`. Common examples of misconduct include cheating, fabrication, plagiarism, and/or unauthorized collaboration. Students are expected to work on their own, with the possible exception of group projects if allowed by the Professor. Students may discuss work with other individuals either in the class or outside the class, but they may not reuse code, results, or text received or retrieved from any source unless clearly disclosed in their submissions. Any student who is found to be violating this policy will be given a failing grade for the work and will be reported to the authorities, including the CSE departments student records, as described in the Universitys Student Code.

Moses Center Statement of Disability: If you are a student with a disability who is requesting accommodations, please contact New York Universitys Moses Center for Students with Disabilities (CSD) at 212-998-4980 or `mosescsd@nyu.edu`. You must be registered with CSD to receive accommodations. See `http://www.nyu.edu/students/communities-and-groups/students-with-disabilities.html`. for information about the Moses Center. The Moses Center is located at 726 Broadway on the 2nd and 3rd floors.

Programming Assignments: In addition to written problems, there will also be one or two programming assignments as part of the homework, plus a two-part project using a database system of your choice:

1. Programming Assignments: How to use a DBMS, create schemas, pose queries, and program in PHP.

2. Project Part I: Designing a complex schema and queries for a given application scenario.

3. Project Part II: implementing a web-based application, typically with PHP and Oracle, mySQL, or another database system, based on the schema and scenario from Part I.

Course Outline (tentative)

1. Introduction and Basics (Chapters 1 to 2 and Chapter 6) *(weeks 1-2)*

    (a) Overview
    (b) Relational Model
    (c) Relational Algebra
    (d) Relational Calculus

2. Relational Queries (Chapters 3 to 5) *(weeks 3-4)*

    (a) Basic SQL
    (b) Advanced SQL
    (c) Integrity Constraints

3. Database and Application Design (Chapters 7 to 9) *(weeks 5-6)*

    (a) Intro to the ER Model
    (b) Relational Design and Normal Forms
    (c) Application Development

Midterm Exam: (around late October, covering material up to Chapter 7 or 8)

4. Storage and Indexing (Chapters 10 and 11) *(weeks 8-9)*

    (a) Disk Models
    (b) External Sorting
    (c) Disk and File Organization
    (d) Indexing

5. Query Processing (Chapters 12 and 13) *(week 10-11)*

    (a) Query Evaluation
    (b) Query Optimization

6. Transaction Management (Chapters 14 to 16) *(weeks 12-13)*

    (a) Transactions
    (b) Concurrency Control
    (c) Recovery