

Principles of Database Systems (CS 6083)

Course Prerequisites: Good programming skills, familiarity with basics of operating systems (file systems, locks, caching), and knowledge of basic algorithms and data structures (such as sorting, search trees, and hashing).

Time and Location: T 6:00–8:25 pm, room RH 721.

Instructor: Prof. Torsten Suel, suel@poly.edu, phone 260–3354.

Office Hours: T 4:45–5:45pm in Room 10.046 (2 MetroTech Center), or by email appointment.

Course Webpage: <http://cis.poly.edu/cs6083/> (class announcements will also be sent via course mailing lists to your Poly student email accounts)

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 that is easier or cheaper to get.)

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

General Information: The course aims to give a broad introduction to 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 at the end if time permits. The workload of the course is expected to be above average, and class attendance is strongly recommended.

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 use PHP.
2. Project Part I: Designing a complex schema and queries for a given application scenario.
3. Project Part II: implement 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: (Probably right after spring break)

4. Storage and Indexing (Chapters 10 and 11) (*weeks 7-8*)
 - (a) Disk Models
 - (b) External Sorting
 - (c) Disk and File Organization
 - (d) Indexing
5. Query Processing (Chapters 12 and 13) (*week 10*)
 - (a) Query Evaluation
 - (b) Query Optimization
6. Transaction Management (Chapters 14 to 16) (*weeks 11-12*)
 - (a) Transactions
 - (b) Concurrency Control
 - (c) Recovery
7. Databases and the Web (*week 13, if time permits – unlikely*)
 - (a) Architecture of the Web
 - (b) Web Search Engines
 - (c) Text Indexing and Information Retrieval
 - (d) Semistructured Data and XML