CS525: Advanced Database Organization

Notes 0: Course Organization

Yousef M. Elmehdwi

Department of Computer Science

Illinois Institute of Technology

yelmehdwi@iit.edu

January 8, 2018

Welcome to CS525

Credits

- The slides are a modified version of the slides used by Hector Garcia-Molina for the CS 245 course at Stanford.
- Responsibility of any errors due to modifications belongs to *me*.

This Course

- Graduate-level introduction to the design and implementation of database management systems

Faculty

- **Instructor:** Yousef Elmehdwi, Senior Lecturer, CS, 2nd time teach this class!
  - Office Hours: F, 11:50am-1:50pm or by appointment at Stuart Building, Room 237D
- **TA**: TBD.

Prerequisite(s)

- Courses: CS425
- Programming experience in C, C++
- Unix OS and file system knowledge is helpful
- Data structures (.e.g., CS401)

Workload and Grading

- Schedule and Important Dates
  - On blackboard
- Programming Assignments 45%(10%10%10%15%)
  - 4 Assignments
  - Groups of 4 students
  - Plagiarism $\rightarrow$ 0 points and administrative action
- Homeworks (5%)
- Class participation and Quizzes (5%)
- Exams: Close book, close notes
  - Midterm Exam (20%): **03/05/2018**
  - Final (25%): TBA

Letter Grade Distribution

Points	Grade
90 - 100	A
80 - 89	B
70 - 79	C
60 - 69	D
0 - 59	E

Programming Assignments

- 4 assignments one on-top of the other
- Optional 5th assignment for extra credit
- Source code managed in **git** repository on Bitbucket.org
- Handing in assignments = submit (push) to repository
- One repository per student
- You will get an invitation from the TA soon (wait one week, the if not, contact TA)
- Git tutorials:
  - http://www-cs-students.stanford.edu/~blynn/gitmagic/book.pdf
  - https://git-scm.com/documentation.

Course Policy

1. Make-up Exams: Only for officially proven health reasons.
2. Late Work:
   - 1-3 days late: -10% points
   - 4-7 days late: -20% points
   - $>$ 7 days late: 0 points

Attendance Policy

- Students are expected to attend all classes and are responsible for all material covered in class, even when absent.
- Students should understand that some material discussed in class is not covered in the textbook.
- Attendance is required.
- I realize that some absences are unavoidable, and you should inform your instructor prior to missing any classes.
- Missing more than 6 classes will decrease your overall grade by a letter grade.
- You will be advised to withdraw from the course if you miss more than 10 classes.

Course Overview

- File organization and access, buffer management, performance analysis, and storage management
- Database system architecture, query optimization, transaction management, recovery, concurrency control
- Reliability, protection, and integrity
- And more when time permits

Course Objectives

After attending the course students should be able to:

- Understand the design decisions behind textbook DBMS architectures
- Know the trade-offs of various storage organization techniques
- Be able to build parts of a small-sized data processing system from scratch
- Understand the basics of query optimization
- Know standard implementations of relational operators such as join, aggregation, and set operations
- Be able to estimate the cost of executing an operator/query based on DB statistics
- Know standard database indexing techniques
- Understand concurrency control and recovery mechanisms

Tentative Course Outline

The weekly coverage might change as it depends on the progress of the class.

Week	**Content**
Week 1	Introduction/ Hardware
Week 2	File and System Structure
Weeks 3-4	Indexing and Hashing
Weeks 5-7	Query Processing
Week 8-9	Crash Recovery
Weeks 10-11	Concurrency Control
Weeks 12-13	Concurrency Control
Week 14	Transaction Processing
Week 15	Advanced topics

Suggested Texts, Readings & Materials

- Garcia-Molina, Ullman, and Widom, Database Systems: The Complete Book, $2^{nd}/3^{rd}$ Edition, Prentice Hall, 2008
- Elmasri and Navathe , Fundamentals of Database Systems , 6th Edition , Addison-Wesley , 2003
- Ramakrishnan and Gehrke , Database Management Systems , 3nd Edition , McGraw-Hill , 2002
- Silberschatz, Korth, and Sudarshan , Database System Concepts , 6th Edition , McGraw Hill , 2010

Next

Notes 1: Introduction