CS157b Spring 2018 Sec1 Sec Home Page/Syllabus

Database Management Systems II

Instructor: Chris Pollett
Office: MH 214
Phone Number: (408) 924 5145
Email: chris@pollett.org
Office Hours: MW 4:30-5:30pm
Class Meets:
Sec1 MW 3:00-4:15pm in MH223

Prerequisites

To take this class you must have taken: CS157A with a grade of C- or better.

Texts and Links

Required Texts:	Database Systems: The Complete Book. Hector Garcia-Molina, Jeff Ullman, and Jennifer Widom
Online References and Other Links:	Postgres. Maria DB (MySql fork). Sqlite. Oracle. DB2. Neo4j.

Description

This class continues where CS157A left off. To begin data storage will be considered. This includes a discussion of how disks works, a discussion of algorithms for disk scheduling and sorting, and a discussion of how RAID systems work. Then indexes, B+-trees, record structures, multi-dimensional indexes, etc will be discussed. The focus will then switch to how query evaluation is implemented. Next we will cover database recovery algorithms. This will be followed with some lectures on transaction processing and concurrency control techniques and then a discussion on combining data from different databases. The semester will conclude with a discussion of data mining and databases and the internet.

Course Learning Outcomes (CLOs)

By the end of this course, a student should be able to:

CLO1 -- Know common database record formats

CLO2 -- Given an index structure based on a B-tree or extensible hashing be able to figure out the effect of performing an insert or a delete

CLO3 -- Create a simple query transaction in a modern DBMS system.

CLO4 -- Tune queries and know how to perform query performance evaluations

CLO5 -- Know the ARIES recovery algorithm

CLO6 -- Be able to use isolation levels for concurrency control in a popular DBMS

CLO7 -- Be able to create or deploy a query mediator for a system with at least two data sources.

CLO8 -- Be able to determine how the A Priori algorithm would operate on a toy dataset

CLO9 -- Be able to explain Topic Sensitive Page Rank algorithm.

Course Schedule

Below is a tentative time table for when we'll do things this quarter:

Week 1: Jan 24	Star Ch 13 Secondary Storage Management
Week 2: Jan 29, Jan 31	Finish Ch 13
Week 3: Feb 5, Feb 7	Ch 14 Index Structures
Week 4: Feb 12, (HW1 due) Feb 14	Ch 15 Query Execution
Week 5: Feb 19, Feb 21	Ch 16 The Query Compiler
Week 6: Feb 26, Feb 28 (HW2 due)	Ch 17 Coping With System Failure
Week 7: Mar 5, Mar 7	Ch 18 Concurrency Control
Week 8: Mar 12, Mar 14 (Midterm)	Review
Week 9: Mar 19, Mar 21 (HW3 due)	Ch 19 Transaction Management
Week 10: Mar 26, Mar 28	Spring Recess
Week 11: Apr 2, Apr 4	Ch 20 Parallel and Distributed Databases
Week 12: Apr 9 , Apr 11	Start Ch 21 Information Integration
Week 13: Apr 16, (HW4 due) Apr 18	Finish Ch 21
Week 14: Apr 23, Apr 25	Ch 22 Data Mining
Week 15: Apr 30 , May 2	Start Ch 23 Databases and the Internet
Week 16: May 7 , May 9 (HW5 due)	Finish Ch 23
Week 17: May 14 , May 16 (No Class)	Review
	The final will be Thursday, May 17 from 12:15-2:30pm

Grading

HWs and Quizzes	50%
Midterm	20%
Final	30%
Total	100%

Grades will be calculated in the following manner: The person or persons with the highest aggregate score will receive an A+. A score of 55 will be the cut-off for a C-. The region between this high and low score will be divided into eight equal-sized regions. From the top region to the low region, a score falling within a region receives the grade: A, A-, B+, B, B-, C+, C, C-. If the boundary between an A and an A- is 85, then the score 85 counts as an A-. Scores below 55 but above 50 receive the grade D. Those below 50 receive the grade F.

If you do better than an A- in this class and want me to write you a letter of recommendation, I will generally be willing provided you ask me within two years of taking my course. Be advised that I write better letters if I know you to some degree.

Course Requirements, Homework, Quiz Info, and In-class exercises

This semester we will have five homeworks, weekly quizzes, and weekly in-class exercises.

Every Monday this semester, except the first day of class, the Midterm Review Day, and holidays, there will be a quiz on the previous week's material. The answer to the quiz will either be multiple choice, true-false, or a simple numeric answer that does not require a calculator. Each quiz is worth a maximum of 1pt with no partial credit being given. Out of the total of twelve quizzes this semester, I will keep your ten best scores.

On Wednesday's, we will spend 15-20 minutes of class on an in-class exercise. You will be asked to post your solution to these exercises to the class discussion board. Doing so is worth 1 "pre-point" towards your grade. A "pre-point" can be used to get one missed point back on a midterm or final, up to half of that test's total score. For example, if you scored 0 on the midterm and have 10 pre-points, you can use your pre-points, so that your midterm score is a 10. On the other hand, if you score 18/20 on the midterm, you can use at most 1 pre-point since half of what you missed (2pts) on the midterm is 1pt.

Links to the current list of homeworks and quizzes can be found on the left hand frame of the class homepage. After an assignment has been returned, a link to its solution (based on the best student solutions) will be placed off the assignment page. Material from assignments may appear on midterms and finals. **For homeworks you are encouraged to work in groups of up to three people. Only one person out of this group needs to submit the homework assignment; however, the members of the group need to be clearly identified in all submitted files.**

Homeworks for this class will be submitted and returned completely electronically. To submit an assignment click on the submit homework link for your section on the left hand side of the homepage and filling out the on-line form. Hardcopies or e-mail versions of your assignments will be rejected and not

receive credit. Homeworks will always be due by the start of class on the day their due. Late homeworks will not be accepted and missed quizzes cannot be made up; however, your lowest score amongst the five homeworks and your quiz total will be dropped.

When doing the programming part of an assignment please make sure to adhere to the specification given as closely as possible. Names of files should be as given, etc. Failure to follow the specification may result in your homework not being graded and you receiving a zero for your work.

Classroom Protocol

I will start lecturing close to the official start time for this class modulo getting tangled up in any audio/visual presentation tools I am using. Once I start lecturing, please refrain from talking to each other, answering your cell phone, etc. If something I am talking about is unclear to you, feel free to ask a question about it. Typically, on practice tests days, you will get to work in groups, and in so doing, turn your desks facing each other, etc. Please return your desks back to the way they were at the end of class. This class has an online class discussion board which can be used to post questions relating to the homework and tests. Please keep discussions on this board civil. This board will be moderated. Class and discussion board participation, although not a component of your grade, will be considered if you ask me to write you a letter of recommendation.

Exams

The midterm will be during class time on: Mar 14.

The final will be: Thursday, May 17 from 12:15-2:30pm.

All exams are closed book, closed notes and in this classroom. You will be allowed only the test and your pen or pencil on your desk during these exams. The final will cover material from the whole semester although there will be an emphasis on material after the last midterm. No make ups will be given. The final exam may be scaled to replace a midterm grade if it was missed under provably legitimate circumstances. These exams will test whether or not you have mastered the material both presented in class or assigned as homework during the quarter. My exams usually consist of a series of essay style questions. I try to avoid making tricky problems. The week before each exam I will give out a list of problems representative of the level of difficulty of problems the student will be expected to answer on the exam. Any disputes concerning grades on exams should be directed to me, Professor Pollett.

Regrades

If you believe an error was made in the grading of your program or exam, you may request **in person** a regrade from me, Professor Pollett, during my office hours. **I do not accept e-mail requests for regrades.** A request for a regrade must be made no more than a week after the homework or a midterm is returned. If you cannot find me before the end of the semester and you would like to request a regrade of your final, you may see me **in person** at the start of the immediately following semester.

University Policies and Procedures

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic

integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/. Below are some brief comments on some of these policies as they pertain to this class.

Academic Integrity

For this class, you should obviously not cheat on tests. For homeworks, you should not discuss or share code or problem solutions between groups! At a minimum a 0 on the assignment or test will be given. A student caught using resources like Rent-a-coder will receive an F for the course. Faculty members are required to report all infractions to the Office of Student Conduct and Ethical Development.

Accommodations

If you need a classroom accommodation for this class, and have registered with the Accessible Education Center, please come see me earlier rather than later in the semester to give me a heads up on how to be of assistance.