Spring 2014 COMP 430: Introduction to Database Systems

Course Objectives and Outcomes: This course provides an in-depth view of databases from a sophisticated user's perspective, as well as an introductory overview from an implementer's perspective.

The first one half of the class will cover the right way to design and implement software that uses a database as a back end to store and manage data. Even if you think that you know the right way to do this because you have played with MySQL previously, the simple fact is that you probably do not know the *right* way. There will be a lot of emphasis on SQL programming in this part of the class. There is no assumption that you know SQL before you show up.

The second half of the class will be an introduction to how databases work. We will cover how a database compiles, optimizes, and executes a computation that is described in SQL. You will actually implement your own little database execution engine. This part of the class will require some significant Java programming.

Even though we are only looking at database technology from a user's perspective, there are still a lot of topics to cover:

- The ER model
- Conceptual database design and modeling
- The relational model
- Relational query languages (relational algebra, relational calculus, and SQL)
- Logical database design (functional dependencies and normal forms)
- Query optimization
- Query execution

One thing to note is that the class is targeted at someone who (even though they might have used databases extensively and even professionally in the past) has never had a formal, semester- or quarter-long class in database systems. If you have had such as class before (for example, at a prior undergraduate institution), then the first half of the class might be a bit of a review.

Instructor: Professor Christopher Jermaine (I prefer to be called "Chris"). My office hours will be MWF immediately after class until people stop coming, or by appointment. My mobile number is 321-231-0769, my email is cmj4@cs.rice.edu, and my office is 3028.

TAs: We will have three or four TAs, though their identity is unknown at this point.

Meeting times and locations: 11AM to 11:50, MWF Duncan Hall 1070, though hopefully this will change.

Class Attendance: I will not use PowerPoint. I also don't publish my lecture notes. So if you miss lecture, you will have to get the notes from someone else in the class. But more importantly, I tend to teach via a questions and answers. As you will see, class will be highly interactive. As a group, we will have good ideas and bad ideas and we will iterative explore each of these, until we finally reach a high quality solution. If you miss class, you will never learn this process, and you will be at a distinct disadvantage when it comes to . So lecture attendance is very important.

Prerequisites: COMP 280 or 182, plus COMP 211 or 215. Basically, you need to have programming skills, you should know Java (or at least chat with me if you do not) plus you probably should have taken a good, high quality class on discrete mathematics or logic.

Grading: Your grade in 430 is determined as a combination of four quizzes and five assignments. I will also periodically assign homeworks that are not turned in, but will help you to learn the course material.

The last four assignments are programming assignments. The fourth assignment has to parts; the second part is due during finals time (it is the class' final project). Be aware of this as you schedule your travel from Houston after classes end. The fifth assignment is also due during finals time, but it is optional (if you don't complete it, all of your other scores will be multiplied by 1.0/0.95).

The grading breakdown and rough schedule for the assignments and quizzes are as follows:

Assignment 1: 5% of grade; due 3rd week: ER modeling assignment

Quiz 1: 12% of grade; 4th week: ER modeling, relational model and translation from ER to relational

Assignment 2: 15% of grade; due 6th week: SQL in MS SQL Server (programming assignment)

Quiz 2: 12% of grade; 7th week: Relational query languages

Assignment 3: 8% of grade; due 10th week: SQL Server PSMs and triggers (programming assignment)

Quiz 3: 12% of grade; 11th week: SQL DDL and Logical database design

Assignment 4.1: 6% of grade; due 12th week: SQL Parsing and Semantic Checking

Quiz 4: 12% of grade; last week of classes. Covers optimization and execution of queries.

Assignment 4.2: (final project) 13% of grade; due during finals: database execution engine.

Assignment 5: (optional) 5% of grade; due during finals: web programming.

There is no curve in this class, and the grading scale is pre-defined: 90% is an A, 80% is a B, 70% is a C, 60% is a D, and anything lower than that is a failing grade. Also note that in general, I do not give "plus" grades, except in exceptional circumstances. In general, I will give out more B's than C's, more C's than A's, more A's than D's, and the only people who fail the class are those who don't put in much effort. I'm not sure if that makes me a tough grader or not, but the one thing that I can tell you is that I strive to be a fair grader. If you master the material taught in the class, you will earn an A. It's that simple!

Regrade Requests: Much of the grading in the class is subjective. As such, no grade is viewed as final when you first receive it. It is your responsibility to look over every paper that is returned and to carefully check to make sure that it was graded correctly. You are free to discuss grading orally with me or one of the TAs. However, any request for an actual regrade must be made in writing within one week of the time that the paper is returned, with no exceptions. **All regrade requests must be prepared using a word processing program; a hard copy should then be submitted to me, along with the original graded work.** On your regrade request, carefully describe why you feel that you were scored unfairly and/or incorrectly. Even if you discussed the grading issue orally with someone, the written discussion must be self-contained and be able to evaluated based only on what is written on the paper.

Late Assignments/Quizzes. I am extremely sympathetic (some people may even call me a pushover) if you have some conflict that will make it difficult for you to attend a quiz or get an assignment completed on time. However, I don't like people to take advantage of my sympathy. So I have a very strict and explicit set of rules governing missed quizzes and late assignments. To be fair to everyone in the class, these rules are always followed to the letter and without exception, so don't even ask!

1. If you have some conflict and feel like you may need an extra day on the assignment or need to take the quiz a day or two later, it must be cleared with me no fewer than one week (seven days) before the assignment is due or the quiz will be held. I am generally sympathetic to the standard excuses, if you sound credible ("I have three exams that day", "My brother's bar mitzvah is the day before the assignment is due", are the usual type of excuses that I hear). However, if you do not clear it a week in advance, you will receive a zero, with only a two exceptions, given below.
2. If you are ill at the time of a quiz or right before an assignment is due and so you miss the one week window, or if there is a death in your immediate family, I will allow a late assignment or a make-up provided (a) you can give me proof of the circumstances, and (b) you let me know before the quiz is held or the assignment is due.

3. If you simply don't turn in the assignment or don't show up for the quiz, the only valid excuse is a note from a doctor given as proof that you were injured or ill at the due data to such an extent that it would be unreasonable for you to send me an email or leave a message letting me know of your illness or injury. In any other case, the result is a zero on the assignment.

Textbook: This book is not strictly required, and I will not refer to it much until we are several weeks into the class. However, I think it is a really good book: *Database Systems: the Complete Book* Second Edition, by Hector Garcia-Molina, Jeffrey Ullman, and Jennifer Widom (Prentice Hall). This book is not in the bookstore; you should simply order this from your favorite online retailer (if this is a problem for you, let me know).

Electronic Communication: We will associate a Google group with this class (instructions on how to sign up for the group are TBD). It is expected that you will check the set of posts to our group regularly during the semester, and at least once a day in the couple of days before each quiz. Whenever I need to send a message and reach everyone quickly, I will post a message to this group. We will use OwlSpace to handle assignments and turnins.

Students with Disabilities: Please let me know ASAP if you have special needs that require alternative arrangements (such as more time to take each quiz). Students with disabilities will need to also contact Disability Support Services in the Ley Student Center.