CSE 177 – DATABASE SYSTEMS IMPLEMENTATION

- Instructor: Florin Rusu
- Office: SE 262
- Email: frusu@ucmerced.edu
- Web: http://faculty.ucmerced.edu/frusu

**Description.** This course studies the internals of a database management system, with emphasis on query execution. The final goal of the class is to build a fully-functional database execution engine consisting of all the standard components: storage manager, buffer manager, query execution engine, query optimizer, and query compiler.

**Prerequisites.**
- CSE 031 – Introduction to Computer Science and Engineering 2

**Textbook.**

**References.**

**Course format.** The course consists of 3 hour lectures and 3 hour labs per week. The fundamentals of query processing algorithms are presented during the lectures. The understanding of the students is checked through weekly quizzes. The lab is entirely dedicated to the semester-long project which requires students to implement a fully-functional database system based on the concepts presented during the lectures.

**Topics.**

1. Secondary Storage Management
   - The Memory Hierarchy
   - Disks
   - Accelerating Access to Secondary Storage
   - Disk Failures
   - Arranging Data on Disk
   - Representing Block and Record Addresses
   - Variable-Length Data and Records
   - Record Modifications

2. Index Structures
   - Index-Structure Basics
Learning goals. The goal of this course is to expose students to the internals of a database execution engine. At the end of the course, students will be able to understand how a query is executed from the moment a user launches it to getting back the result. This is achieved by a thorough analytical study of the components and algorithms involved in query processing and a hands-on project that requires the implementation of a full-fledged relational database engine.
Learning outcomes. Students will learn the fundamental data processing algorithms implemented inside all the existing relational database engines. They will get detailed exposure to the current research in data management and direct experience with research prototype systems. By the end of the course, students will understand the software architecture of a relational database engine and will know the role and functionality of each component. They will be able to design and implement each of the components in the architecture and to define the interface and the data flow between the components. In essence, the students will learn how to build a relational database engine from ground up.

The progress students make in assimilating the class material is continuously tested through weekly quizzes and a semester-long project in which the students are required to implement a full-fledged relational database system. Quizzes are meant to test the understanding of the algorithms implemented in the engine components. The project is divided into multiple phases – one phase corresponding to each database component – with each phase building on top of the previous ones. Each project phase (database component) is evaluated separately, while the project deliverable is expected to be a fully-functional database engine. Overall, the degree to which the learning outcomes are accomplished is assessed at three different levels:

1. The understanding of specific algorithms implemented in each database component is evaluated through weekly quizzes. This represents the theoretical part of the assessment and is standard exam evaluation.

2. The understanding of the role and functionality of each database component is assessed in the corresponding project phase that requires the implementation of the component. This represents one of the practical parts of the assessment and requires the actual implementation of the algorithms presented during the lectures. The evaluation strategy is as follows: students are provided input data sets and expected output (these are called battery tests) on which to evaluate their implementation. The final score for each part of the project is given by testing the component on additional data sets that are not provided to the students at first.

3. The ultimate goal of the course – build a fully-functional database engine – is assessed through the project deliverable—a database engine that integrates all the components into a single functional system. The components are designed in such a way that they have clear interfaces and they can be tested in isolation (this allows for the independent grading of each component). Evaluating the overall system is a matter of gluing all the components together by the means of the standard interfaces. Given this separation, it is possible to obtain a working system by just plugging-in the necessary components (the instructor insures this is the case by providing the missing components when needed).

Academic integrity policy. Each student in this course is expected to abide by the University of California, Merced’s Academic Honesty Policy. Any work submitted by a student in this course for academic credit will be the student’s own work. Students are encouraged to study together and to discuss information and concepts covered in lectures. Students can provide/receive "consulting" to/from other students. However, the permissible cooperation should never involve one student having possession of a copy of all or part of the work done by someone else, in the form of an email, an email attachment file, a storage device, or a hard copy. Should copying occur, both the student who copied work from another student and the student who gave material to be copied will receive zero credit for the corresponding assignment. Penalty for violation of this Policy can also be extended to include failure of the course and University disciplinary action. During examinations, each student has to do only their own work. Talking or discussing is not permitted, nor students comparing their papers, copying from others, or collaborating in any way. Any collaborative behavior during examinations will result in failure of the exam and may lead to failure of the course and University disciplinary action.

Disability service information. Accommodations for students with disabilities: The University of California, Merced is committed to ensuring equal academic opportunities and inclusion for students with disabilities based on the principles of independent living, accessible universal design diversity. I am available to discuss appropriate academic accommodations that may be required for students with disabilities.
Requests for academic accommodations are to be made during the first three weeks of the semester, except for unusual circumstances. Students are encouraged to register with the Disability Services Center to verify their eligibility for appropriate accommodations.

**Grading.**

- Project: 60%
- Quizzes: 40%