CS 425 Database Organization – Fall 2019

6:25pm – 9:05pm, Mondays, Robert A. Pritzker Science Ctr 121
Instructor: Dr. Yuan Hong
E-mail: yuan.hong@iit.edu Phone: 312-567-5168
Office Hours: 4:30-6:00pm Mondays or by appointment (Office SB-216C)
TA: TBD
TA Office Hours: TBD (Office SB-013)

Course Description

Databases management systems are a crucial part of most large-scale industry and open-source systems. This course familiarizes students with important concepts of database systems and design. We will learn how to design a database using the Entity-Relationship model, how to query and modify a database using the declarative SQL language, and study APIs for writing application programs that use a database system to persist data. Furthermore, the course gives an overview of important database systems techniques such as indexing, query optimization, execution, concurrency control, and recovery.

Course Objectives

After taking the course, students should be able to:

1. Analyze application requirements and develop a database model based on such requirements:
   - Understand the Entity Relationship model and design Entity Relationship diagrams based on application requirements.
   - Understand functional dependency theory.
   - Analyze and construct functional dependencies.
   - Analyze a set of functional dependencies to identify candidate keys.
   - Understand normal forms and be able to normalize a database schema.
   - Define referential integrity constraints.
   - Be able to translate an Entity Relationship model into a relational schema.
2. Understand formal relational languages such as relational algebra and write queries in these languages.
3. Write SQL queries, updates (DML), and data definition (DDL) statements.
4. Develop a basic understanding of database extensibility mechanisms and procedural SQL extensions including triggers and stored procedures.
5. Develop a general understanding of important advanced database topics such as: query optimization, query execution, concurrency control, recovery, index structures, data warehousing, key-value stores, and big data platforms.
6. Implement applications using a commercial or open source relational database management system such as Oracle. That is, design and implement an application that uses a relational database management system for storage and provides a user interface for insertion, deletion, update and querying of data.

Course Material

- The following textbook is required reading material for the course:
  Silberschatz, Korth, and Sudarshan, Database System Concepts, 6th Edition, McGraw Hill, 2010
- Other recommended introductory books on databases are:
  Elmasri and Navathe, Fundamentals of Database Systems, 6th Edition, Addison-Wesley, 2003
  Ramakrishnan and Gehrke, Database Management Systems, 3nd Edition, McGraw-Hill, 2002
  Garcia-Molina, Ullman, and Widom, Database Systems: The Complete Book, 2nd Edition, 2008

The slides will be posted on the Blackboard (before each class).

Prerequisites

Courses: One of CS 331, CS 401, or CS 403

Students with Disabilities

Reasonable accommodations will be made for students with documented disabilities. In order to receive accommodations, students must obtain a letter of accommodation from the Center for Disability Resources. The Center for Disability Resources (CDR) is located in 3424 S. State St., room 1C3-2 (on the first floor), telephone 312 567.5744 or disabilities@iit.edu.

Course Details

The following topics will be covered in the course:

- The relational data model
- Database modeling and design
  - The Entity-Relationship (ER) model
  - Database design and normalization
- SQL
  - Data-definition language (DDL)
  - Data-manipulation language (DML)
- Formal relational languages
  - Relational algebra
- Database Architecture
- Database System Concepts
  - Transactions processing and concurrency control
  - Recovery
  - Indexing
  - Query processing and optimization
  - Security and access control

Homework

There will be some homework assignments during the course. The main objective of these assignments is for you and the instructor to evaluate how well you internalized the topics covered in the course (e.g., Relational Algebra, SQL, and Database Modeling). The homework assignments should be completed individually, and submitted electronically on the Blackboard by 11:59PM on the due date.

Course Project

During the second half of the semester, you will have a project assignment. The purpose of the project is for you to gain experience with physical implementations of databases. You will be taking a well-known SQL benchmark and trying to tune your database to improve performance.

There will be a group project (2-3 students). The class will collectively choose a topic in the first few weeks. In the project, you will design a database schema for a given application domain, derive a relational schema (and implement an SQL script to create this schema), and write an application accessing the database. The project will be split into three tasks:

- **Design an ER database model**: You will analyze the requirements of the application and then design an Entity-relationship diagram modelling the domain.
- **Implement an SQL script that creates the database**: You should use the techniques thought in class to derive a relational schema from the ER model and implement an SQL script that creates this schema using your Oracle account.

- **Implement the application**: You will an application program that enable a user to store, manipulate, and query the database.

Databases

Every student will get an account for an Oracle database running on one of our Unix servers. You can use this account to test SQL queries and DDL statements for the homework assignments and from class. You can also install Oracle in your personal computer (Windows, Mac OS).

Exams

There will be a midterm and final exam (cumulative) covering the topics of the course.

Grading Policies and Scale

- Homework Assignments: 20%
- Course Project: 25%
- Midterm Exam: 25%
- Final Exam: 30%

- **A** 85-100
- **B** 70-84
- **C** 50-69
- **E** less than 50

We will check for plagiarism. Plagiarism will result in zero points for the assignment and possibly academic sanctions.

Late policies: -20% per day late

Tentative Schedule

Week	Dates	Topics (Chapters in the textbook)
1	8/19	Syllabus and Introduction (Chapter 1)
2	8/26	The Relational Data Model (Chapter 2), Formal Relational Query Languages (Chapter 6)
3	9/2	Labor Day (No Class) Homework 1
4	9/9	SQL – Introduction (Chapter 3) Homework 1 Due
5	9/16	
6	9/23	SQL – Intermediate: Views, Integrity Constraints, Access Control (Chapter 4)
7	9/30	SQL – Advanced: APIs for SQL Access, Procedural Constructs (Chapter 5)
8	10/7	Fall Break Day (No Class) Homework 2
9	10/14	**Midterm Exam**
10	10/21	ER Model (Chapter 7) Homework 2 Due
11	10/28	Database Design and Normal Forms (Chapter 8)
12	11/4	Transactions (Chapter 14) Homework 3
13	11/11	Concurrency Control (Chapter 15) Homework 3 Due
14	11/18	Storage and Index Structures (Chapter 10 & 11)
15	11/25	Data Warehousing and Mining (Chapter 20 and extra materials)
16	12/2	**Project Demo**
17	TBA	**Final Exam**