Principles of Database Systems

CS 5423: Fall 2019
Department of Computer Science
Oklahoma State University

August 22, 2019

Class Time and Location

Wednesday 4:30 PM - 7:10 PM. MSCS 310

Instructor Information

- Instructor: Esra Akbas
    - Email: eakbas AT okstate.edu
    - Office: MSCS 216
    - Office hours: Monday - 10:30AM - 12:00 PM, Wednesday - 3:00 - 4:30 PM
- Teaching Assitant: Habib Boloorchi Tabrizi
    - Email: hboloor AT okstate DOT edu
    - Office: MSCS 312
    - Office hours:Monday 2:30-3:30 and 6-7 (pm), Tuesday 3-4 (pm)

Class Homepage

Course Website: `www.cs.okstate.edu/~eakbas/CS5423.html`

This page provides the schedule of lecture topics, reading materials, and important milestones/deadlines. Please note: as a plan, the schedule is subject to changes and adjustments). The class will also have a Canvas page. Assignments, lecture notes and announcements will be posted on Canvas.

Course Description

Do you know how many Turing Awards were given to the database field? For what contributions?

The dominating relational database system, which we take for granted now, was deemed impossible to implement and difficult to use in its early days. What breakthroughs turned the rather mathematical idea into reality? What are the classic milestones that pioneered and shaped the landscape of the data management field?

Do you know that, although relational database systems are almost commodity software now, researchers are even more fascinated by general data management issues? The new challenges naturally arise in the junction of databases vs. information retrieval, structured vs. unstructured data, querying vs. searching, the Web, and many more novel contexts. What are the current topics and future agenda?

To build the essential foundation as a stepping-stone to data management research, this course exposes students to advanced database system design principles and concepts. We will broadly explore the classic as well as more recent research work. Our agenda will be two fold:

- As an advanced course for databases, CS 5423 studies the fundamentals of implementing relational database management systems from the system perspective (i.e., how to design and implement a database system), the course will study data representation, indexing, query processing and optimization.

- As the first graduate course in the general data management field, CS 5423 studies the advanced topics beyond traditional relational database systems, such as parallel/distributed data processing, data mining and Web data analytic.

Textbook

- Database Systems: The Complete Book 2nd edition, by Hector Garcia-Molina, Jeff Ullman and Jennifer Widom. ISBN: 978-0131873254, Prentice Hall, 2008.

- Fundamentals of Database Systems 7th edition, by Elmasri and Navathe. ISBN: 9780133970777, 2016.

- Database Management Systems 3rd edition, by Raghu Ramakrishnan and Johannes Gehrke. ISBN: 978-0072465631, McGraw-Hill, 2002.

- Readings in Database Systems 5th edition, by Peter Bailis, Joseph Hellerstein, and Michael Stonebraker. ISBN: 978-0262693141, The MIT Press, 2005.

- Fundamentals of Database Systems 7th edition, by Elmasri and Navathe. ISBN: 9780133970777, 2016.

In addition to this you can use the lecture notes, third party online documentation and extra reading material handed out in class.

Prerequisites

CS4433: Introduction to Database Systems
CS3423: Data Structures and Algorithms or equivalents courses are required.
Furthermore, students should come with good programming skills. If you are not sure whether you have the right background, please contact the instructor. Note: We will not cover programming-specific issues in this course.

Format and activities

This course will draw (advanced) materials from the textbook as well as database literature. Students will study the materials, read classic papers and submit summaries, do a presentation in the class, complete a significant research project, and take a final exam at the end of the semester.

- **Lectures, class participation, and reading:** we encourage (and appreciate!) students to attend classes, because effective lectures rely on students' participation to raise questions and contribute in discussions. We will provide lecture notes and related papers before class, which will be posted on Canvas.

  Read the textbook and papers for the required reading before lectures, and study them more carefully after class. Please note that all the required readings are fair materials for exams. These materials may not be fully covered in lectures. Our lectures are intended to motivate as well as provide a road map for your reading– with the limited lecture time we may not be able to cover everything in the readings.

- **Questions:** We encourage students discussing their questions and problems first with their group peers and classmates. This way, you can get immediate help and also learn to communicate "professionally" with your peers. In any case for more thorough discussion, come to the office hours of TA's and the instructor's. Any announcement will be posted Canvas. Make sure to check it frequently enough to stay informed.

- **Assignment**: There will be two written assignments spaced out over the course of the semester. All the assignments should be done individually by the students. Assignments should be submitted through the canvas.

- **Exam**: There will be a final exam held in the end of the semester.

- **Paper summaries**: There will be a series of papers for careful studying, which represent our selection from both classic and state-of-the-art literature in database and data management areas. Each student will submit a one-or-two page summary for each paper in the due date through the canvas.

- **Paper presentation**: To enrich our channels of learning, we ask our students to choose one paper related to her/his project and present it in the class. The paper should be chosen from the designated conferences/journals and published no earlier than 2008. Selected paper needs to be approved by instructor. Each student will sign up and present in the second half of the semester. Each student will spend 20 minutes for presentation and 10 minutes for Q&A and will lead and moderate discussions in the class.

- **Projects**: There will be a semester-long project which involves significant database or data related research. The project will be structured with several milestones due in the course of the semester, leading to a conference/journal-paper-style final report (and auxiliary materials, such as source code and datasets) near the end of the semester.

Grading Policy

The final course grade will be computed as follows:

Assignments	15%
Paper Presentation	10%
Paper Summary	15%
Project	35%
Final exam	25%

Requests for regrading should be within a week of grades being posted on Canvas.

The final grade will be calculated according to your numerical average as shown in the table below.

A	90-100
B	80 - 89
C	70 - 79
D	60 - 69
F	0 - 59

This table indicates minimum guaranteed grades. Under certain limited circumstances (e.g., an unreasonably hard exam), we may select more generous ranges or scale the scores to adjust.

Late Assignment Policy

- There will be 40% deduction for One day late submission More than one day late assignments will not ordinarily be accepted. If, for some compelling reason, you cannot hand in an assignment on time, please contact the TA or instructor as far in advance as possible.

- No make-up exams (except under extremely unusual circumstances).

Academic Dishonesty

The Computer Science departmental policy for academic dishonesty and misconduct applies to this class. In addition, a student attempting to gain unfair advantage by keeping an examination paper longer than the time permitted is guilty of academic misconduct. Assignments/projects/exams are to be done individually, unless specified otherwise. It is a violation of the Academic Honor Code to take credit for the work done by other people. It is also a violation to assist another person in violating the Code. Examples of cheating behaviors include:

- Discuss the solution for a homework question.

- Copy programs for programming assignments.

- Turning in group work for assignments where you are expected to work as an individual.

- Use and submit existing programs/reports on the world wide web as written assignments.

- Submit programs/reports/assignments done by a third party, including hired and contracted.

- Plagiarize sentences/paragraphs from others without giving the appropriate references.

Penalty for violating the Academic Honesty: Students who do not comply with the described policy will receive a grade of F in the course. Furthermore, the case will be reported to the University Officials

Accommodation for Disabilities

The Computer Science departmental policy for students with disabilities applies to this class. Anyone who has a need for examinations by special arrangements should see the instructor as the earliest possible opportunity during scheduled office hours.

Syllabus Change Policy

This syllabus is a tentative guide for the course and is subject to change. You'll be informed in class if there's a change in the syllabus.

*Syllabus Attachment: `www.cs.okstate.edu/~eakbas/cs5423/SyllabusA.pdf`