University of Arizona  
CSc 630: Advanced Topics in Database Systems  

Fall 2001  (call# 70427)

Instructor:  Bongki Moon (bkmoon@cs.arizona.edu)  
Office:  Gould-Simpson 715, 621-4326 (voice), 621-4246 (fax)  
Office hours:  Mon 3:00-4:00 and Wed 11:00-12:00; or by appointment.

Class schedule:  M/W 4:00-5:15, Gould-Simpson 701

Overview and Topics:  CSc630 is an advanced graduate-level course in database systems that will cover recent and on-going research issues. The focus will be on non-traditional large-scale database applications. The main topics that this course will investigate this year include the following.

- Fundamentals: GEMINI paradigm, $L_p$ metrics, (content-based) similarity match (nearest neighbor search). [1, 2, 3, 4, 5]
- Time series data (e.g., stock, ECG, random walk): dimensionality reduction techniques (DFT, DWT, SVD, PCA), time warping, indexing and retrieval. [6, 7, 8, 9, 10, 11]
- String data (e.g., string attributes, DNA sequences): exact and approximate match, substring selectivity estimation, string join. [12, 13, 14]
- XML and semi-structured data: relational vs. non-relational data model, indexing XML data, processing regular path expression queries. [15, 16, 17, 18, 19, 20]
- Multidimensional aggregation (e.g., spatial or OLAP data): (incremental) temporal aggregation, progressive approximate aggregation. [21, 22, 23, 24]
- Architectural issues: cache-conscious index structures, optimization of main memory accesses. [25, 26, 27]
- Mobile and broadcasting data: location management, tracking mobile objects, broadcasting disks, selective dissemination of data. [28, 29, 30, 31, 32]

This course will consist of lectures by the instructor and student presentations of research articles based on a reading list. The reading list will include recent research papers as well as classic, and is expected to evolve during the course reflecting the interests and accomplishments of the class.

Prerequisites:  CSc342 (Data Structures) and CSc460/560 (Database Systems); or instructor’s permission.

Recommended texts:
- There is no required textbook for this course. Instead, a collection of research articles will be provided by the instructor.

Grading:  This course will be offered for three credits. The grading policy will be Research project (50%), Reading assignments and Class participation (25%) and three-quarter-term quiz (25%). If you do a research project in a group, except for special circumstances, all members in the group will receive an identical grade for each part of the project.
Research project: Students are required to complete a research project. A list of sample or suggested projects will be provided. Students are also encouraged to pick up their own project topics provided that the topics are relevant to the course and challenging enough. All the project proposals are subject to the instructor’s approval. Projects will include extensive literature surveys, prototype implementations, performance measurements, and several reports. The final reports of the projects should be of quality both in depth and breadth equivalent to refereed conference papers.

Reading assignments and Class participation: You are to select research articles (from the reading list or any articles approved from the instructor) for approximately two class sessions. Give an overview of the main issues in the articles (for less than 25 minutes) and lead the discussion. You are expected to read every reading assignment before class. Background materials will be provided by the instructor if necessary, and then the major points of the day’s reading assignment will be discussed in class.

For each reading assignment, you will be required to do something to indicate that you have done the assigned reading and that you understand most of it. You will be often asked to turn in a summary of key points before class; sometimes you will be called on to answer questions; sometimes you will be given a short quiz; However, if you have carefully read each assignment, you will be in good shape.

Class URL and Listserv: http://www.cs.arizona.edu/classes/cs630/fall01/630-2

The class web page is the primary source for information on the projects and other important announcements. We may also use the class listserv to exchange questions and tips about the projects or to discuss any issues that arise in class.

Class accounts: You will be given an account on lectura (solaris) and linux machines, and an access card to use terminals in Gould-Simpson 228. These accounts are for instructional purposes only. To obtain an account, you will need to go to the 7th floor of Gould-Simpson during the first week of classes. There should be signs on the 7th floor directing you to the location of the electronic account application process.

Code of Academic Integrity: Students are responsible for understanding and complying with the University’s Code of Academic Integrity (http://w3.arizona.edu/~studpubs/policies/cacaint.htm). Violations of the Code will, at minimum, result in loss of credit for a graded item. An egregious first violation or any second violation will minimally result in failure of the entire course. See http://www.cs.arizona.edu/admin/CAI.policy/memo.html for the minimum penalties for code violations adopted by the Department.
A Tentative Reading List

Ordered By Class Sessions


