

The University of Michigan - Dearborn
Department of Computer and Information Science

CIS 656 Advanced Database Systems Winter 2012

Instructor:	Dr. Qiang Zhu	Office hours:	M. 3:40 - 4:40pm & 9:00 - 10:00pm
Office:	213 CIS		W. 12:30 - 1:30pm
Phone:	(313) 593-4998		Also by appointments
Email:	qzhu@umich.edu	Web:	http://www.engin.umd.umich.edu/~qzhu

The database field has experienced rapid and incessant growth since the development of centralized relational databases. This course provides an in-depth discussion of some advanced database areas. In this term, it focuses on studying spatial databases and distributed databases. Topics covered include spatial database concepts, spatial data modeling and representation, spatial database queries, spatial data model with constraints, spatial access methods, spatial query processing, distributed database design, distributed query processing and optimization, distributed concurrency control, peer-to-peer data management, and cloud data management. An overview of other modern database technologies, such as parallel databases, multimedia databases, active databases, temporal databases, data warehousing and data mining, deductive databases, fuzzy databases, mobile database, real-time databases, transactional workflows, streaming data processing, and bioinformatics, is also given. (3 credit hours).

Your grade in this course will be determined by your scores for two exams (a midterm and a cumulative final) and one course project (including a proposal, a term paper and a presentation). The midterm exam will count for 30%, the final exam will count for 40%, and the project will count for 30%. Late work will be penalized (5% each working day, weekends count as 5%), as will evidence of cheating in any form¹.

Prerequisite: CIS556 or CIS 421 or equivalent

Class schedule: Mondays 6:10 - 9:00pm in Room 40 SB

Texts:

- T1. *Spatial Databases*, by Philippe Rigaux, et al., Morgan Kaufmann Publishers, Inc., 2002
- T2. *Principles of Distributed Database Systems*, 3rd Ed., by M. T. Ozsu and P. Valduriez, Springer Science+Business Media, LLC, 2011
- T3. *CIS 656 Lecture Notes*, Winter 2012

¹The University of Michigan - Dearborn values academic honesty and integrity. Each student has a responsibility to understand, accept, and comply with the University's standards of academic conduct as set forth by the Code of Academic Conduct, as well as policies established by the schools and colleges. Cheating, collusion, misconduct, fabrication, and plagiarism are considered serious offenses. Violations will not be tolerated and may result in penalties up to and including expulsion from the University.

Tentative Contents:

1. Overview of modern database technologies (reading: lecture notes)
 - DB technology trends in data model, database hardware, user interface, program interface, presentation and display, and nature of processing
 - modern DB technologies: object-relational databases, active databases, distributed databases, parallel databases, deductive databases, fuzzy databases, data warehousing and data mining, spatial and temporal databases, multimedia databases, bioinformatics, mobile databases, streaming data processing, internet related DB technology, long-duration transactions, real-time transactions, transactional workflows, peer-to-peer data management, and cloud data management
2. Spatial databases (reading: T1: chap.1, 2, 3, 4, 6, 7)
 - spatial database concepts and terminology
 - representation of spatial objects: modes, models, formats, and standards
 - logical model and query language: reference schemas and queries
 - constraint data model: modeling with constraints, linear constraint data model, modeling entity-based data, modeling field-based data and moving objects
 - spatial access methods: grid file, quadtree, R trees
 - spatial query processing: spatial join algorithms, complex query processing
3. Distributed databases (reading: T2: chap.1, 3, 7, 8, 11, 12, 16, 18.2)
 - what and why distributed databases
 - architecture and classification
 - distributed database design: strategies, fragmentation, and allocation
 - distributed query processing: query decomposition, localization of distributed data, distributed cost model, join and semijoin, distributed query optimization algorithms
 - distributed concurrency control: distributed transaction processing model, taxonomy of concurrency control mechanisms, locking-based concurrency control algorithms (centralized 2PL, primary copy 2PL, decentralized 2PL)
 - peer-to-peer data management: infrastructure, schema mapping, query, replica consistency
 - cloud data management: taxonomy, architectures, data management
4. Project presentations

Midterm: Monday, March 12th, in class (tentative)

Final Exam: Wednesday, April 23rd, 6:30 - 9:30pm (please confirm with the official announcement near the end of the term).