

Wright State University

CORE Scholar

Computer Science & Engineering Syllabi

College of Engineering & Computer Science

Fall 2012

CS 4700/6700: Introduction to Database Management Systems

Soon M. Chung

Wright State University - Main Campus, soon.chung@wright.edu

Follow this and additional works at: https://corescholar.libraries.wright.edu/cecs_syllabi



Part of the [Computer Engineering Commons](#), and the [Computer Sciences Commons](#)

Repository Citation

Chung, S. M. (2012). CS 4700/6700: Introduction to Database Management Systems. .
https://corescholar.libraries.wright.edu/cecs_syllabi/343

This Syllabus is brought to you for free and open access by the College of Engineering & Computer Science at CORE Scholar. It has been accepted for inclusion in Computer Science & Engineering Syllabi by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.

CS 4700/6700 Introduction To Database Management Systems, Fall 2012

Description: Survey of logical and physical aspects of database management systems. Entity Relationship and relational models for databases are presented. Physical database design methods, formal DB design concepts, and operating principles of database systems are discussed.

Prerequisite: CS 3100 Data Structures and Algorithms, or equivalent.

Instructor : Dr. Soon M. Chung, 403 Russ Center
(937)775-5119, soon.chung@wright.edu, <http://www.cs.wright.edu/~schung>

Class: M. W. F. 3:35-4:30 pm at 154 Russ Center

Office hour: M. W. 4:45-5:45 p.m. at 403 Russ, or by appointment.
*use e-mail for short questions.

Text Book: R. Elmasri and S. B. Navathe, Fundamentals of Database Systems, 6th edition, Addison Wesley, 2010.

Topics: DBMS concepts and architecture (Chap 1, 2)
Entity-Relationship model (Chap 7)
Relational data model (Chap 3)
ER to Relational mapping (Sec. 9.1)
Relational algebra (Chap 6)
SQL - a relational database language (Chap 4, 5)
File organizations and hashing (Chap 17)
Index structures for files (Chap 18)
Relational DB Design Theory and Normalization (Ch. 15, 16)
Query Processing and Optimization (Ch. 19)
Transaction Processing, Concurrency Control and Recovery (Ch. 21, 22, 23)

Grading: A:[85,100], B:[75,85), C:[65,75), D:[55,65), F:[0,55)

- There is no homework, but solutions of some exercise questions will be provided.
- 1st Midterm 15% (10/5, F), 2nd Midterm 15% (11/2, F), Project 30%,
Final 40% (12/12, W. 2:45-4:45 pm)
- Project is paper-review, programming, or DB design. Select one by 10/17.
- The final report (around 25 pages in double space) is due on the last class, 12/12.
 - (1) small database design and SQL programming using MS-Access or some other DBMS
{ description of problem 4%, ER design 6%, Relational Schema 5%,
SQL queries and results 8%, discussion 7% }
 - (2) programming
Extendible hashing (Ref. Sec. 17.8.3) simulation using a high-level
programming language (C, C++, Java, etc.),
{ design 5%, documentation 5%, correctness 15%, discussion 5% },
 - (3) paper-review project
{ papers reviewed 6%, technical quality 8%,
written presentation 7%, discussion 9% }
- submit the topic and a list of at least 5 selected papers by 10/17.

CS 4700/6700 DB Design Project

The DB design and implementation project can be performed by the following steps:

1. pickup an application problem.
 2. Entity-Relationship model design.
 3. Convert the ER design to an equivalent Relational DB schema.
 4. create the relations using MS Access, Oracle, or some other DBMS.
 5. Insert sample tuples into the relations.
 6. Write some meaningful SQL queries and run against the sample relations.
 7. Print out the results.
 8. Add discussion.
- Your ER design should include some relationships, so that some of your SQL queries can involve more than one relations. In general, a relational DB design including more than 4 relations is okay.
 - Try most of the typical SQL commands to see how they work.
 - Your final report should include the description of the problem, data requirements, ER design, relation schemas and instances, query descriptions and SQL queries, query outputs, and discussion.
 - The discussion section can include anything that you want to share with other students in our class, such as your comments on the logical DB design, implementation, DBMS functionality and performance, future work, etc.
 - A guideline for using MS Access is given at <http://www.cs.wright.edu/~schung/cs6700.htm>

CS 4700/6700 Programming Project

Simulation of Extendible Hashing.

1. The hash value of the hash-key attribute of each record is given as a randomly generated integer value.
2. Each data bucket (block) can store 40 records and is dynamically allocated. Each directory entry stores a pointer to a data bucket.
3. Generate up to 10^4 hash values (to simulate up to 10^4 records), and for every 10^2 hash values, display followings:

$$\text{utilization of the directory} = \frac{\text{number of data buckets}}{\text{number of entries in the directory}}$$

$$\text{utilization of the data buckets} = \frac{\text{number of total records}}{\text{number of records that can be stored in all the data buckets}}$$

Note: You can use any programming language, and you can plot the outputs if you want.

Reference:

1. Section 17.8.3 of the text book.
2. R. Fagin, J. Nievergelt, N. Pippenger, and H. R. Strong, Extendible Hashing: A fast access method for dynamic files, *ACM Trans. on Database Systems*, 4(3), 1979, pp. 315-344.

CS 4700/6700 Paper Review Project

1. Choose a topic and select at least 5 relevant technical papers. High-quality journal papers are preferred.
2. Summarize and compare the papers, and then add your own discussion.
3. Submit the working title and the list of candidate papers. (due 10/17)
4. Submit the report and the papers you studied. (due 12/12)
5. Size of the report is around 30 pages in double-space.

Possible Topics

- Database models
- Database access mechanism (such as indexing, hashing, etc)
- Query optimization
- Concurrency control and recovery
- Parallel algorithms for query processing
- Performance evaluation of DBMS
- Parallel database systems
- Distributed database
- Multidatabases (federated databases)
- Expert database
- Logic and database
- Multimedia database
- Object-Oriented database
- Image database
- Engineering database
- CAD/CAM database
- Text retrieval system
- Data mining
- Data warehousing
- Other relevant topics

Reference Sources

- IEEE Trans. on Software Engineering
- IEEE Trans. on Knowledge and Data Engineering
- Computer (IEEE Computer Magazine)
- Communications of ACM
- ACM Trans. on Database Systems
- ACM Trans. on Knowledge Discovery from Data
- ACM Trans. on Information Systems
- Information Systems
- Multimedia Systems (Journal by ACM and Springer International)
- IEEE Multimedia (Magazine by IEEE)
- Data and Knowledge Engineering (Journal)
- Data Mining and Knowledge Discovery (Journal)
- Knowledge and Information Systems (Journal)
- Proc. of IEEE Int'l Conf. on Data Engineering
- Proc. of ACM Conf. on Management of Data (SIGMOD Conference)
refer to the volumes of SIGMOD RECORD
- Proc. of ACM Symp. on Principles of Database Systems (PODS)

- Proc. of Very Large Data Bases (VLDB) Conference
- IEEE Trans. on Parallel and Distributed Systems
- ACM Computing Surveys
- Proc. of Int'l Conf. on Knowledge Discovery and Data Mining
- Proc. of IEEE Int'l Conf. on Data Mining
and Others