

Database Management Systems

Winter 2003

CMPUT 391: Introduction

Dr. Osmar R. Zaiane



University of Alberta

Objectives of Lecture 1

Introduction

- Get a rough initial idea about the content of the course:
 - Lectures
 - Resources
 - Activities
- Mind refresher for Database Systems (CMPUT 291) (Students who are taking this course need to have knowledge about databases and expertise in structured programming, i.e., CMPUT 291 or equivalent is a course requirement)

Class and Office Hours

Classes for Section B1:

Mondays, Wednesdays and Fridays 11:00 to 11:50

CSC B2

Classes for Section B2:

Mondays, Wednesdays and Fridays 13:00 to 13:50

CSC B10



Office Hours:

Mondays and Wednesdays from 15:30 to 16:30

Office: ATH 352

By appointment:

E-mail: zaiane@cs.ualberta.ca Tel: 492 2860

Check appointment page on my web page.



Labs and TAs

Labs (CSC 219):



H01: Wednesday 14:00 to 16:50

H02: Thursday 08:00 to 10:50

H03: Thursday 11:00 to 13:50

H04: Thursday 14:00 to 16:50

H05: Friday 08:00 to 10:50

TAs:



Samer Nassar

(samer@cs.ualberta.ca)

Fan Deng

(fandeng@cs.ualberta.ca)



Mohammad El Hajj (mohammad@cs.ualberta.ca)

Jia Li

(jial@cs.ualberta.ca)

Jie Luo

(jieluo@cs.ualberta.ca)



TA office hours: T.B.A.



(Tentative, subject to changes)

Course Schedule

There are 14 weeks from January 6th to April 9th

- *Lectures:* cover the basic material for the course.
- *Tutorials:* complement the course and will be given during some lab hours. They contain information that is necessary to do the project.
- *Assignments and Project:* will be given later in the semester. You should work on them during lab hours (when there are no tutorials).
 - Implementation assignments will also be demonstrated during lab hours in the week following the assignment deadline.
 - The project demos will be at the end of the semester.
- There are no additional lab exercises.

Midterm (February 14th)

Final Exam (April 24th for section B1 and April 22nd for section B2)

Project Demos (last week of the semester)

Lab Tutorials

- Installation and use of Tomcat
- Java servlets
- Connectivity with databases
- Java Server Pages



Tentative

Course Calendar



- | | |
|--------------------------------------|-----------------------|
| • Introduction | Jan 06 |
| • Database Design Theory | Jan 08-10-13-15-17-20 |
| • Query Processing and Optimisation | Jan 22-24-27-29-31 |
| • Concurrency Control | Feb 03-05-07 |
| • Database Recovery and Security | Feb 10-12-24 |
| • Midterm | Feb 14 |
| • Object-Oriented Databases | Feb 26-28 |
| • Inverted Index for IR | Mar 03 |
| • Spatial Data Management | Mar 05-07-10 |
| • XML | Mar 12-14-17 |
| • Data Warehousing | Mar 19-21 |
| • Data Mining | Mar 24-26-28 |
| • Parallel and Distributed Databases | Mar 31- April 02 |
| • Project Demos | Apr 03 to Apr 09 |

Evaluation and Grading

Your final grade will depend on the entire profile of the grades in your lecture section (bell-curve distribution) and a particular composite score does not guarantee a particular final grade. However, your composite score will be computed using the following weights:



- | | | |
|------------------------|-----|--|
| • Assignments | 12% | (4 assignments 3% each) |
| • Mid-Term Examination | 23% | (Feb 14 th) |
| • Project | 35% | (demo at end of semester) |
| • Final Exam | 30% | (April 22 nd & April 24 th) |

! You have to pass the final exam in order to pass the course

More About Evaluation

Re-examination.

None, except as per regulation.



Collaboration.

Do Collaborate on assignments; do not merely copy.
Do not exchange machine-readable code (programs)

Plagiarism, cheating, misrepresentation of facts and participation in such offences are viewed as serious academic offences by the University and by the Campus Law Review Committee (CLRC) of General Faculties Council. Sanctions for such offences range from a reprimand to suspension or expulsion from the University.

Plagiarism.

Work submitted by a student that is the work of another student or a tutor is considered plagiarism. Read **Sections 26.1.4** and **26.1.5** of the University of Alberta calendar. Cases of plagiarism are immediately referred to the Dean of Science, who determines what course of action is appropriate.

Resources



Course home page:

<http://www.cs.ualberta.ca/~zaiane/courses/cmput391/>

Contains links to course notes, detailed course calendar and other resources

Newsgroup

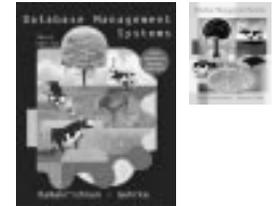
<news://news.srv.ualberta.ca/ualberta.courses.cmput.391>

Textbook:

Database Management Systems (Third Edition)

by Raghu Ramakrishnan and Johannes Gehrke

McGraw-Hill, 2002, ISBN: 0-07-232206-3



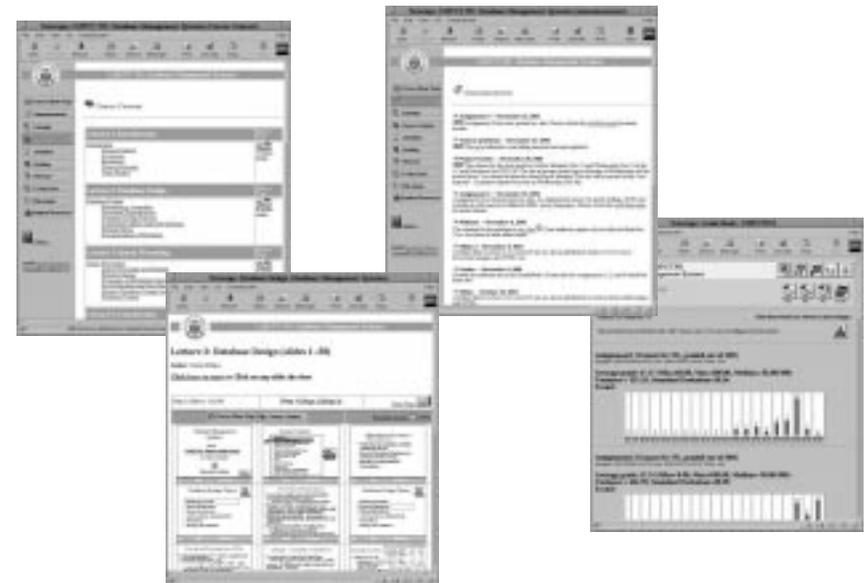
Other recommended textbooks:

- P.M. Lewis, A. Bernstein and M. Kifer, Databases and Transaction Processing, Addison-Wesley, 2002, ISBN: 0-201-70872-8.
- R. Elmasri and S. Navathe, Fundamentals of Database Systems, 3rd Edition, Addison-Wesley, 1999, ISBN: 0-8053-1755-4.

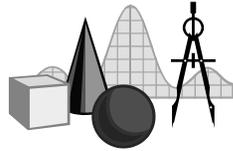
On-line Resources



- CMPUT 391 web page
- Course slides
- Web links
- Glossary
- Student submitted resources
- Student spaces
- U-Chat
- Frequently asked questions



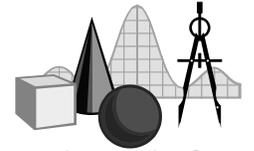
Course Project



- The objectives of the course project are to gain hands-on experience in design and implementation of Web-based information systems that use a database management system for storage and management of data.



Course Project



- Projects will be demonstrated in class at the end of the semester.
- The idea is to build a web-based application from the ground up with technologies such as:

ORACLE-8, Java, Servlets/JSP, JDBC, HTML forms, etc.

The topic of the project is a management information system for a fictive “distributed” Yacht-club.

Past Projects



Student registration



On-line store



Bibliog. DB



Car Rental



Auto registration

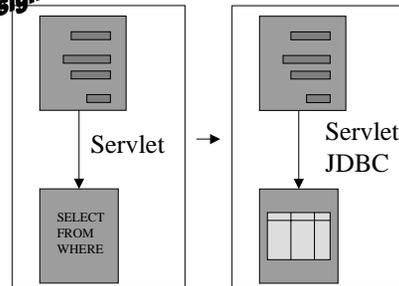


Medical Lab

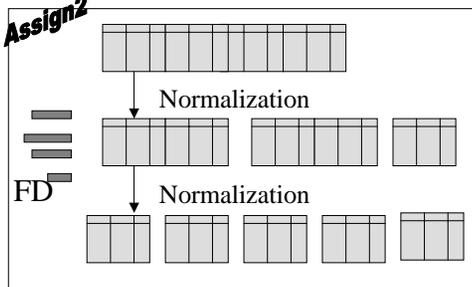


Assignments

Assign1



Assign2



Assign3



Assign4



Objectives for CMPUT 391

- To **understand** the fundamental concepts underlying database management systems:
 - database design methodology (normalization,...)
 - database management systems (query optimisation, concurrency, recovery, security,...)
- To **learn** about additional DB support for special data types such as XML documents and Spatial Data
- To **get acquainted** with data analysis issues such as data mining, data warehousing and information retrieval;
- To **gain** hands-on experience with database application systems and commercial database management systems.
 - developing an application system using ORACLE & web technology

What you studied in CMPUT 291

ER Model
Relational Model
Relational algebra
Relational Calculus
SQL

Database Design/Normalization

Disk and File Structures
Indexing
Tree-structured indexes
Hashing

The main objective for CMPUT 291 was:
Ensure that the student becomes a knowledgeable user of database management technology

- Understand how database management differs from file processing;
- Learn how to model data and build relational databases;
- Use query languages to access stored data.

Basic Notions

Data: any information

| | |
|---------------|--------------------------|
| Manufacturing | Product data |
| University | Student data, courses |
| Hospital | Patient data, facilities |
| Bank | Account data |

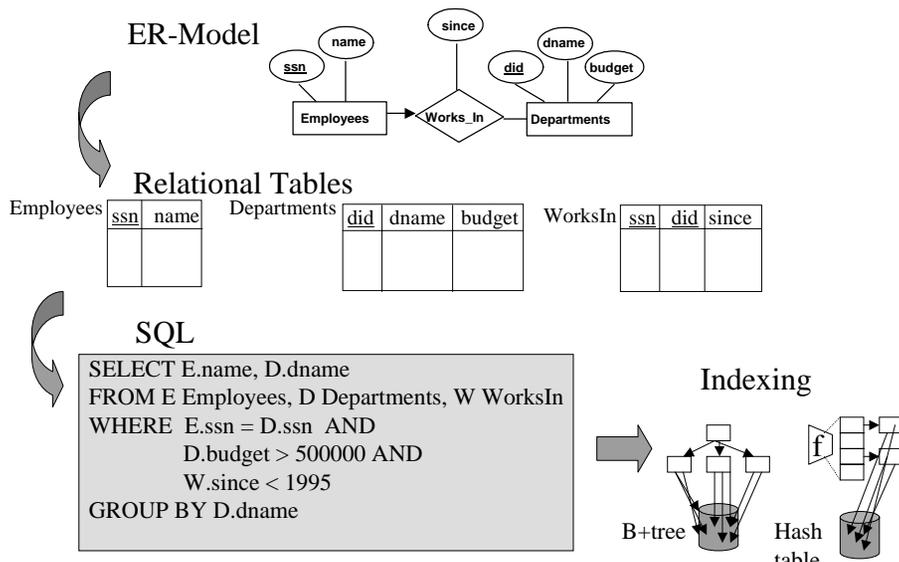
What is a database?

Database: a large collection of data
an integrated collection of data

What is a database management system?

Database management system:

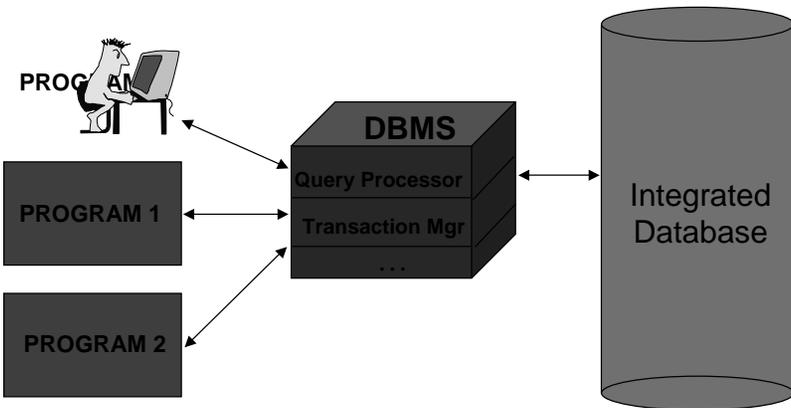
a software system that provides an efficient as well as convenient environment for accessing data in a database.



Functionalities of DBMSs

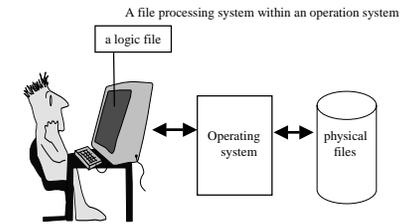
- Specifying the database structure
 - data definition language
- Manipulation of the database
 - query processing and query optimisation
- Integrity enforcement
 - integrity constraints
- Concurrent control
 - multiple user environment
- Crash recovery
- Security and authorization

Database Approach



Why file management systems?
efficient file accesses

Why database management system?
efficient and convenient data access



A file system sets up mappings between logical files and physical files.

File management systems:

- goal: efficiency
- problem: high speed ram vs. low speed disk access
- solution: complicated file structures

Database management systems:

- goal: efficiency as well as convenience
- problem: conflicts between efficiency and convenience
- solution: data independence supported by various database models

Database Management System Users

Naïve and casual database users

- a person who knows nothing or not much about the database and accesses the data via forms and pre-built queries embedded in programs.

Advanced users and application programmers

- a person who knows about the structure of the database and about query languages, and can embed queries in programs.

Expert users and application programmers

- a person who knows how to write complex queries and advanced database programs based on knowledge on DBMS intricacies.

Database administrator

- a person who is responsible for the database design, scheme modification, user authorization, etc..

CMPUT 291

CMPUT 391

Database models:

- conceptual tools used to describe:
 - data
 - data relationships
 - data semantics
 - data constraints

Major database models

- E-R model: a logic foundation for conceptual database design
- network model: a set of records connected by links
- hierarchical model: a set of database trees
- relational model: a set of tables

Data independence

Capability of changing a database schema without having to change the schema at the next higher level

