CpSc 462/662: Database Management Systems (DBMS) (TEXNH Approach)

Introduction

James Wang

Attribution

- Materials and lecture notes in this course are adapted from various sources, including the authors of the textbook and references, Internet, instructor's personal notes, instructor's friends, etc.
- The instructor has tried to attribute all authors of the course materials.
- If you think that the instructor may overlook something, please tell the instructor.

2

What is DBMS?

- Database is an integrated collection of data that models a real-world enterprise.
  - Entities (e.g., students, courses)
  - Relationships (e.g., Jane is taking CpSc 462)
- A Database Management System (DBMS) is a software package designed to store and manage databases.
- Most Modern DBMS systems are:
  - Relational Database Systems.
  - Online accessible: It is hard to imagine that any database system designed in the Internet era cannot be accessed online.

3

DBMS Applications

- Registration System
- Course Management System
- Billing System
- Online Bookstore
- Digital Library
- Google
- YouTube (http://www.youtube.com)
- More ...

4

Why do we study DBMS?

- Shift from ***computation*** to ***information***
  - at the "low end": scramble to Web (a mess!)
  - at the "high end": scientific applications such as post genomic biomedical research.
- Datasets increasing in diversity and volume.
  - Digital libraries, Interactive video, Human Genome project, Online stores
  - ... need for DBMS exploding
- DBMS encompasses most of CS
  - OS, languages, theory, "AI", multimedia, logic
- Most CS jobs are database related.

5

Course Objectives

- understand conceptual modeling concepts and be able to use ER Model to design database applications.
- be familiar with relational data models and be able to design relational database schemas from ER diagrams.
- be able to use an industry standard query language (SQL) to query the relational databases.
- understand the basic concepts of query optimization and learn simple query optimization techniques.
- gain experience in designing and implementing web-based database systems.
- improve oral and written communication skills through written and oral presentation of their projects.

6

Instruction Approach

- **The problem-based instruction approach, TEXNH, will be used in this course.**
- **Students are expected to learn database concepts through a semester-long multimedia database project.**
- **Specifically, students will be required to implement an online multimedia database system, MeTube, while learning the theories and techniques through lectures.**
  - MeTube system is essentially a modified version of the popular YouTube system (http://www.youtube.com).
  - The modification to YouTube system includes adding the support of more media types, including graphics objects, video, audio, and animation clips.

7

MeTube System Development and Class Topics

- **The entire instruction is centered with the development of the semester-long multimedia database project, MeTube.**
- **The database concepts and techniques will be acquired through the course of solving the problems faced in the MeTube system development.**
- **Feedbacks of the intermediate project results will help detect the problems.**
- **The entire course consists of 6 integrated phases in which new theories or techniques will be learned along with the project development.**

8

Phase 1: Database Design

- **Explore the YouTube systems to identify system requirements.**
- **Finish the conceptual design of the database using ER model.**
- **Concepts and techniques:**
  - Business requirement identification
  - Design specification
  - Data modeling
  - ER model and diagram.

9

Phase 2: Relational Schema and Table Design

- **Convert conceptual designs into relational schemas and then design the actual database tables for the MeTube system**
- **Theories and techniques:**
  - Relational data model
  - relational tables and their properties
  - Normalization
  - SQL language
  - MySQL database management system.
- **Question:**
  - how to store the multimedia data for the MeTube system?
    - In file? Or in database?

10

Phase 3: SQL Queries

- **Design complex SQL queries to retrieve information from the multimedia database.**
- **Evaluate and optimize the queries.**
- **Theories and techniques:**
  - Convert English queries into the corresponding SQL statements.
  - Execute the queries on the multimedia databases using command line interface and phpMyAdmin.
  - Performance evaluation and and query optimization.

11

Phase 4: Web Interface Design

- **Design the web interfaces to add and retrieve information from the MeTube database.**
- **Allow users to upload media files into the MeTube system.**
- **Retrieve information from the MeTube database and properly present the information online.**
- **Techniques:**
  - Web Server, HTTP, and forms.
  - HTML/XHTML.
  - PHP, MySQL database access through PHP.
  - Web script for multimedia presentation.

12

Phase 5: Testing and Improvement

- **Test the MeTube system using various use cases.**
- **Improve the MeTube system.**
- **Evaluate the system performance.**
- **Optimize the database and PHP program.**
- **Concepts and techniques:**
  - Test case design
  - Performance evaluation
  - System optimization.
  - views, triggers, and stored procedures.

13

Phase 6: Demo, Presentation and Final Report

- **Students need to give an oral presentation about their projects.**
- **Students need to give a demonstration to the instructor.**
- **Students need to submit a final report containing:**
  - Design consideration
  - ER diagrams
  - Queries
  - Implementation details
  - Problems solved
  - etc.

14

General Information

- **Class web:**
  http://www.cs.clemson.edu/~jzwang/1101462/cpsc462.htm
- **Class hours:**
  - 11:00 AM - 12:15 PM, Tuesday and Thursday.
- **Classroom:**
  - 415 Daniel Hall.
- **Office hours:**
  - 9:30 AM - 10:30 AM, Tuesday and Thursday. or by appointment
- **Q&A:**
  - Send questions to jzwang@cs.clemson.edu.

15

Attendance

- **It is very important to attend the class since the lecture contents may be drawn from sources (internet, technical manual, etc.) other than the textbook.**
- **Failure to attend the course most likely leads to missing some information that won't show in lecture notes or textbook.**

16

Grading Policy

- **Midterm Exam (20%):** Cover the content studied in first half of the semester.
- **Final Exam (20%):** A comprehensive exam on content studied in the entire semester.
- **Quizzes (20%):** Up to ten quizzes will be given throughout the semester.
- **Project (40%):** A team of 2 or 3 students will be assigned to implement a MeTube system.
- **Penalty Points (-10%):** The instructor may penalize those who fail to attend the class without instructor's approval. Each tardy will result in 1 penalty point.
- **Grading:** A (90 - 100), B (80 - 89), C (70 - 79), D (60 - 69), F (0 - 59). (The scale may be curved down at the end of the semester)
- CpSc 662 Students: more requirements on project and additional questions in exams.

17

Textbook and References

- **An Introduction to Database Systems, Eighth Edition, C. J. Date, Addison Wesley, 2004, ISBN: 0-321-19784-4 (Textbook).**
- **Fundamentals of Database Systems, Fifth Edition, Ramez Elmasri, Shamkant B. Navathe, Addison Wesley, 2007, ISBN: 0321369572.**
- **Database Systems: The Complete Book, First Edition, Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer D. Widom, Prentice Hall, 2002, ISBN: 0-13-031995-3.**
- **Database Management Systems, Third Edition, Raghu Ramakrishnan and Johannes Gehrke, McGraw-Hill, 2002, ISBN: 0072465638.**

18

Website and Blackboard

http://www.cs.clemson.edu/~jzwang/1101462/cpsc462.htm
http://bb.clemson.edu

- Check website or blackboard regularly for latest announcements, updates, and assignments.
- All lecture notes and assignments will be posted on website or blackboard before or after the class.
- Useful links will be posted up there too.

19

Important Issues

- A successful project is the essential part of the success in this class.
- Start the project from the first day.
- Ask questions if you have any doubts on anything.
- If you are confused with some parts of the project description, ask questions before actual implementation.
- Tests will be limited to those content discussed in the classroom.

20