Texas A&M University
College of Engineering
Computer Science Department

SYLLABUS

CPSC 603:600/310:501 Applied Database Systems
Fall Semester 2002

Instructor: Dr. Michael Thomadakis
Office-Hours: Fri 12:30–14:30 or by appointment
Location: H. R. Bright Building, room 315B
e-mail: `miket@cs.tamu.edu`
Tel: 845-3907

1 Course Information

Course Catalog Description `File structures and access methods; database modeling, design and user interface; components of database management systems; information storage and retrieval, query languages, high level language interface with database systems.`

Official Course Site and Newsgroup Class material will be posted on the official web site for these courses. It is most important to visit often the following sites:

- `http://courses.cs.tamu.edu/cpsc310/miket` and,
- `news:tamu.classes.cpsc310` or
- `news:tamu.classes.cpsc603`

for official announcements and up-to-the-minute information. **Visit often both sites**.

Course Description and Objective Introduction to concepts and design methodologies for database systems. Emphasis on data design and modeling, E. R. diagram modeling, E. F. Codd's relational model, relational algebra/calculus, ANSI SQL, normalization process, concurrency control, transaction processing, distributed, WEB-based and multi-tiered database systems; introduction to distributed databases; hands-on database design and application development using recent and open APIs and standards.

CPSC 310/603 have the following goals

- study design, implementation and pragmatics aspects of database management systems;
- study the conceptual, logical, and physical design of databases, and database applications;
- provide a balanced coverage of database theory, pragmatics and actual implementations (case studies);
- provide hands-on experience on state-of-the-art platforms and technologies, leveraging industrial and open database standards;

- allow students to develop their database modeling and development skills.

Prerequisites: CPSC 210. Familiarity with at least one imperative programming language, such as, `C/C++` or `JAVA`. Familiarity with operating systems (*e.g.*, UNIX), and basic machine organization would be helpful but is not required.

Main Textbooks

- (**Required**) [GaUlWi2002], Hector Garcia-Molina, J. D. Ullman and J. Widom, *Database Systems: The Complete Book*, Prentice Hall, 2002, 1152 pp; `ISBN: 0-13-031995-3`. Publisher's URL:
  `http://vig.prenhall.com/catalog/academic/product/1,4096,0130319953,00.html`
  Current errata list can be found at:
  `http://www-db.stanford.edu/~ullman/dscb/errata.html`
- (**Recommended**) [Kolo2000] George B. Koch and Kevin Loney, *Oracle 8i: The Complete Reference*, Osborne/McGraw-Hill, May 23, 2000; ISBN 072123648. [Book and CD Edition.] Our database development will take place on ORACLE on Unix and on Windows NT/2000.

Other Reference Textbooks In the course of our discussion we will occasionally refer to material from the following textbooks.

- [SKS2001] A. Silberschatz, H. Korth and S. Sudarshan, *Database Systems Concepts*, Fourth Edition, McGraw-Hill, July 2001. ISBN 0-07-228363-7.
- J. D. Ullman and J. Widom, *A First Course in Database Systems*, Prentice-Hall, 1997.
- J. D. Ullman, *Principles of Database and Knowledge-Base Systems*, Vol. 1, Computer Science Press, 1988.
- J. D. Ullman, *Principles of Database and Knowledge-Base Systems*, Vol. 2, Computer Science Press, 1989.
- C. J. Date, *An Introduction to Database Systems*, Vol. 1, Seventh Edition, Addison-Wesley, 2000.
- [ElNa] Ramez Elmasri and Shmkant B. Navathe, *Fundamentals of Database Systems*, Third Edition, Addison-Wesley, 2000; ISBN 0–8053–1755–4.

Extended Course Description CPSC 310 is an introductory undergraduate course on database design and database management systems. CPSC 603 is an introductory graduate course. No prior knowledge on databases is expected. All the necessary material will be covered in class. The graduate course will include some additional requirements with respect to the technical depth of the material and the level of projects. Both theoretical and pragmatic aspects of database design, as well as, important problems dealt with in modern database management management systems will be emphasized.

In the recent years, the field of databases has grown to encompass several fully-fledged sub-fields, starting from the high-level data modeling down to the field of physical database design. Database management systems (DBMS) are fairly large and complex programming systems designed for efficient data storage and retrieval of large volumes of data. The underlying database theory (*e.g.*, concurrency control, transaction processing, serializability *etc.*) have application on a wide range of fields, including databases, multi-processor design and distributed systems. At the same time, the sheer volume of data requires specialized data storage and organization techniques. Data retrieval and modification requires specialize indexing and retrieval algorithms which are geared towards efficiency. Recent database applications require the storage of different non-traditional types of (voluminous) objects, including, images,

Table 1: Course Topics and Corresponding Textbook Chapters

#	**Topic**	**Textbook Chapter**
1	Introduction to Database Systems Architecture and Concepts	Ch 1
2	Conceptual Data Modeling with Emphasis on Entity-Relationship Models	Ch 2
3	Introduction to the Relational Data Model	Ch 3
4	Relational Algebra	Ch 5
5	ANSI SQL and Mappings from ER, EER into Relational Data Model	Ch 4
6	Constraints and Triggers in DBMS	Ch 7
7	System Aspects of SQL	Ch 8
8	Functional Dependencies and Normalization Theory for Relational Databases	Ch 3
9	Introduction to Object Oriented Databases, Object Relational Databases and XML	Ch 4, 9 (partially)
10	Storage and File Structures	Ch 11
11	Indexing and Hashing	Ch 12, 13 (partially)
12	Introduction to Query Processing and Optimization	Ch 15 (partially)
13	Introduction to Concurrency Control Techniques	Ch 18 (partially)
14	Introduction to Transaction Processing	Ch 19
15	Introduction to Database Recovery Techniques	Ch 17
16	Distributed Databases and Client/Server Architectures	Ch 18, 19 (partially)
17	Applications and three-tiered systems	Ch 20 (partially)
18	Data Warehousing and Data Mining	Ch 20 (partially)
19	Other DB Topics, such as Parallel DBMSs, Advanced Data Types, Advanced Transaction Processing	Notes
20	Case Studies	Notes
21	Hierarchical and Network Data Models	Notes

or video and audio contents. Finally, databases are increasingly being found in the "back-end" as data servers, with WWW based applications in the front-end interacting with users over the Internet.

Several programming mini-projects will be given to corroborate the concepts and algorithms covered in the lectures. Familiarity with an imperative programming language such as `C/C++` is expected. Knowledge or familiarity with basic operating systems mechanisms or programming are not required but they will be helpful.

2 Course Topics

This course will focus on the topics shown in Table 1. These may not be covered in this order, or not at all if time does not permit it.

Schedule and Reading Assignments Students are responsible to read the class textbook [GaUlWi2002] and other assigned material ahead of the corresponding lecture.

Table 2: Course Requirements and Corresponding Weight

1.	programming and term projects	35%
2.	HW assignments	05%
3.	2 mid-term exams	30%
4.	Final exam	30%

Table 3: Grading Brackets, V is your effective class average

$V \geq 90$	A
$80 \geq V < 90$	B
$70 \geq V < 80$	C
$60 \geq V < 70$	D
$V < 60$	F

3 Grading Scheme and Course Requirements

3.1 Grading Scheme

The course consists of lectures, reading, homework, programming assignments and examination periods. The weights and the grade brackets are shown in Table 2 and 3, respectively.

3.2 Examinations

All major examinations will be held in class with exact dates determined later. The exams will generally test lecture material, as well as, lab, homework, and project material submitted with other partners. You must be familiar with the work that your partners contributed in group assignments. All exams will be closed book and closed notes (unless otherwise stated).

3.3 Assignments

Homework will be assigned weekly or bi-weekly and will typically consist of end-of-chapter exercises and other problems. Homework exercises can be carried out by one or two people, unless stated otherwise.

Mini-projects are more substantial assignments that will typically require about one to three semester weeks each to complete. There will be three to five mini-projects, each carried out by a group of two people. *You are strongly advised to form groups as early as possible.* Projects will involve database programming in `SQL` and variants. Students will be required to carry out projects in stand-alone, interactive ("scripting") SQL, and in SQL embedded in `C/C++/Java`.

All DB tools run on departmental UNIX systems, so you will need to get a student computer account immediately.

Assignment handouts will be placed on the course web page ahead of time.

3.4 Deadline Policy

Late work will not be accepted, in general. Turn in all work by the established deadline. In case you have difficulties finishing an assignment contact the TA or the Instructor before the deadline. Late work can be accepted only under circumstances beyond student's control and after arrangement with the Instructor, prior to the deadline.

> **HINT**: work turned-in on time is eligible for partial credit. It will always be better to turn work in by the deadline, as trying to "perfect" it and turn it in late will give you no points at all.

3.5 Re-grading Policy

A student can request re-grading of assignments and exams, if he/she believes that the points assigned are inconsistent with the quality and merits of the submitted work. To request re-grading you have to follow the guidelines below.

1. Re-grading requests must be submitted AT MOST ONE WEEK AFTER the item has been graded and returned to the student. After this time limit NO re-grading requests will be honored.
2. Re-grading requests must be as specific as possible and must be accompanied by a reasonable amount of justification and documentation. Requests must be in written form for major assignments and exams.
3. E-mails must be sent to TAs and graders or the instructor within the one week time limit.

3.6 Excused Absences and Make-up Policy

Make-ups for assignments and exams will be given only under circumstances beyond student's control (*a university sanctioned excuse*). Prior arrangements with the instructor must be made when feasible and official verification of circumstances necessitating the absence will be required. The correct approach is to start working on assignments as early as possible and contact us when you encounter difficulties. In general, the closer to the deadline you request our assistance, the harder may be to obtain our help.

3.7 Group Work

Each assignment will state the number of people who can work together in a group. Some assignments will be done by students individually. Partners will turn in a single assignment paper (with each partner's name and section number on it) and each partner will receive the same grade. You are also free to work individually.

3.8 Submission of Work

All assignments deliverables must be submitted electronically, by the due date, using the `turnin` method. Late assignments are not going to be accepted in general, unless a University sanctioned excuse is provided ahead of time. A student will earn points when he/she submits the assignment on time, by the partial credit policy. Note that `email` submissions will not be accepted (they will be ignored without notice). You have to follow the submission and media policies and guidelines published on the web.

4 Teaching Personnel and Resources

The Instructor and the Teaching Assistants are your most valuable resource. We are available during our published office hours or at other times by appointment. Walk in any time during the office hours to see the instructor or the TAs. We can meet at other times but students need to arrange for an appointment. Unannounced visits outside office hours may not be honored. Make it a habit of yours to discuss with your TA or the Instructor homework problems and other assignments. Contact the TAs or the Instructor to collect your scores of all graded items. Observe the re-grade period mentioned above.

The Instructor and the TAs will provide all necessary resources for the course to progress smoothly, including, handouts, notes, programming code files, among others. Most items are posted to the web page for the students to down-load.

People having difficulties with either the material or finishing assignments, are strongly encouraged to discuss the problems with the Instructor and the TAs *as early as possible*. Things can be corrected when problems are addressed early. The Instructor will propose a recovery plan for the student. The recovery plan is a course of action that will assist the student boost his understanding of the material and perform better. As a rule of thumb, the more a student delays to contact the Instructor concerning problems with the material, the more intense effort will be required of the student in order to recover.

5 Scholastic Dishonesty

Plagiarism is the passing of someone else's work as one's own, without giving the original author due credit. Scholastic dishonesty will be treated very strictly as per Texas A&M University rules. Typically, the given incidence **has to be reported to the Department Head as per Texas A&M University rules**. The Department Head will then determine the type of punitive actions, including, 0 points to the assignment, an `F` grade in the course, academic suspension, or even expulsion from Texas A&M University.