COMP9315 00s1
Database Systems Implementation
Course Introduction

Aims

This course aims to examine the kinds of data structures and algorithms that go to make up a database management system such as Oracle. An alternative title for the course could be "Data Structures and Algorithms for Data-intensive Systems".

At the end of this course, students should

- have some insight into how DBMSs function internally
- understand how to analyse the performance of data-intensive systems
- be familiar with a variety of programming techniques for large-scale data manipulation
- appreciate advanced transaction processing issues in database systems

The course will examine techniques from both the research literature and from commercial systems.

To avoid any confusion, here is what this course **is not about**:

- it will **not** teach you how to be an Oracle DBA
  (although a good understanding of the issues in this course will make you a more effective DBA)
- it will **not** teach you how to performance tune Oracle applications
  (although you will have the necessary theoretical background to learn how to do this effectively)

Prerequisites

This course assumes that you have a reasonable understanding of the following topics:

- the Java language and programming in Java (e.g. completed COMP9021)
- developing applications on RDBMSs (SQL, relational algebra e.g. COMP9311 or COMP3311)
- file organisation/manipulation (Unix `open,close,lseek,flock` e.g. COMP9024 or COMP2011)
- sorting/searching techniques (quick/merge sorts, binary trees, hash tables e.g. COMP9024 or COMP2011)

If you do not have this prerequisite knowledge, you should reconsider your enrolment in the course.

Staff Members

Name	Office	Phone	E-mail	Role
Dr. Xuemin Lin	K17-503	x6493	lxue	Lectuer in charge
Mr. Qing Zhang	K17-TBA	TBA	qzhang	Subject Admin

Lectures

The course has 3 hours of lectures per week, held each Tuesday from 6-9pm.

The following table gives the major topic areas in the syllabus and an approximate guide to how much time will be devoted to each. More fine grained syllabus details can be found on the course Web site.

Introduction	1 week
Storage Management	1 week
Indexing and Query Processing	5 weeks
Query Optimization	2 weeks
Transaction Management	4 weeks

The lecture notes for each week will be available on the course web site on the Monday of that week.

Assessment

There will be 3 assignments, 2 projects, and 1 final exam. The assessment components contribute to the overall course mark as follows:

Project 1	10%
Project 2	10%
Assignment 1	5%

Assignment 2	5%
Assignment 3	5%
Final Exam	65%

Text Books

The lecture notes for the course are more than 50% from the first and second text books.

1. Ramakrishan, *Database Management Systems*, McGraw-Hill, 1997.
2. Garcia-Molina, Jeffrey Ullman, and Widom, *Database System Implementation*, Prentice Hall, 2000.
3. Silberschatz, Korth & Sudarshan, *Database System Concepts* (3r d ed), McGraw-Hill, 1998.
4. Elmasri & Navathe, *Fundamentals of Database Systems* (3rd ed), Benjamin/Cummings, 1999.

References:

The following two papers cover much of the more advanced material in storage management and query processing:

1. Graefe, *Query Evaluation Techniques for Large Databases*, ACM Computing Surveys, 25(2), 73-170, 1993.
2. Gaede and Gunther, *Multidimensional Access Methods*, ACM Computing Surveys, 30(2), 170-231, 1998.

Both papers are available to download and print from the course web site.

Warning: there is a lot of inconsistency in notation among the textbooks and the papers. This is due to the area still being reasonably young and research-active. Keep this in mind when reading material from a variety of sources; you'll probably need to do some mental mapping from one notation/analysis-scheme to another.

More Information

For further information on this course, and to keep up to date with any changes, please consult the course web site (frequently):

http://www.cse.unsw.edu.au/~cs9315/

Xuemin Lin, 28 Feb 2002