University of Macau
Faculty of Science and Technology
Department of Computer and Information Science
SFTW371 Database Systems II
Syllabus
1st Semester 2013/2014
Part A – Course Outline

Elective course in Computer Science

Course description:
(3-2) 4 credits. This course provides an in-depth study of: query optimization and evaluation; transaction management; concurrency control; database recovery; database security and authorization; and distributed database systems. In which, information retrieval and web search engines are also briefly introduced and reviewed.

Course type:
Theoretical with substantial laboratory/practice content

Prerequisites:
- SFTW370

Textbook(s) and other required material:
- Raghu Ramakrishnan & Johannes Gehrke, ***Database Management Systems*****,** Third edition, McGraw Hill, 2003.

References:
- Ramez Elmasri & Shamkant B. Navathe, ***Fundamentals of Database Systems***, Sixth edition, Addison-Wesley, 2011.
- A. Silberschatz, H. F. Korth and S. Sudarshan, ***Database System Concepts***, Sixth edition, McGraw Hill, 2011.

Major prerequisites by topic:
- Fundamentals of database systems
- Computers and programming
- Data structures and algorithms.

Course objectives:
- Learn internal structures, query evaluation and optimization. [a, c, e, k]
- Learn database transaction management and crash recovery. [a, k, l]
- Learn distributed databases and the related query evaluation and transaction. [a, c, e, j, l]
- Learn security and authorization mechanism for the web-based database applications. [a, c, e, j, l]
- Construct and implement one or more internal database components with the case study problems. [a, c, e, j, k, l]

Topics covered:
- **External sorting (3 hrs) –** two-way merge sort, external merge sort, using B+ tree for sorting, minimizing I/O cost versus number of I/Os.
- **Query evaluation (5 hrs) –** introduction to system catalog, three common techniques to operator evaluation and access paths, algorithms for relational operations, overview of query optimization, alternative evaluation plans and estimating the cost of a plan.
- **Evaluating relational operators (6 hrs) –** selection operation, project operation, join operation, set operations, and aggregate operations.
- **Transaction management (6 hrs) –** The ACID properties, transactions and schedules, concurrent execution of transactions, performance of locking., transaction support in SQL, introduction to crash recovery and lock-based concurrency control, overview of ARIES.
- **Concurrency control (6 hrs) –** Two-Phase Locking (2PL), serializability, and recoverability, introduction to lock management, lock conversions, dealing with deadlocks, special locking techniques, concurrency control without locking.

- **Crash recovery (6 hrs) –** introduction to ARIES, introduction to log, other recovery-related structures, the Write-Ahead Log protocol, checkpointing, recovering from a system crash, media recovery.
- **Security and authorization (6 hrs) –** introduction to database security, access control, discretionary access control and mandatory access control, security for Internet applications, additional issues related to security.
- **Distributed databases (4 hrs) –** distributed DBMS architectures, storing data in a distributed DBMS, distributed query processing, distributed transactions, concurrency control and recovery.

Class/laboratory schedule:

Timetabled work in hours per week			No of teaching weeks	Total hours	Total credits	No/Duration of exam papers
Lecture	Tutorial	Practice				
3	1	1	14	70	4	1 / 3 hours

Student study effort required:

Class contact:	
Lecture	40 hours
Tutorial	28 hours
Mid-term exam	2 hours
Other study effort	
Self-study	20 hours
Homework assignment	10 hours
Project / Case study	22 hours
Total student study effort	122 hours

Student assessment:

Final assessment will be determined on the basis of:

In-class/Lab Exercise	5%	Homework	10%	Project	20%
Quizzes	15%	Mid-term	15%	Final exam	35%

Course assessment:

The assessment of course objectives will be determined on the basis of:

- Homework, project and exams
- Course evaluation

Course outline:

Weeks	Topics	Course Work
1	**External Sorting** Talk about the difference between sorting data on disk and in-memory; introduce a simple two-way merge sort, and how does external merge-sort work. Study the way of using B+ tree for sorting, and observe the techniques of minimizing I/O cost versus number of I/Os.	Assignment 1
2-3	**Query Evaluation** Introduce system catalog, three common techniques to operator evaluation, algorithms for relational operations, overview of query optimization, and building alternative evaluation plans.	
3-4	**Evaluating Relational Operators** Introduce evaluation of various relational operators, such as selection, projection, join, set and aggregate operations.	Assignment 2
5-6	**Transaction Management** ACID properties will be described, talk about concurrent execution of transactions, some overviews of lock-based concurrency control and crash recovery.	
7-8	**Concurrency Control** Study 2PL, serializability and recoverability, introduction to lock	Assignment 3

Weeks	Topics	Course Work
	management, more details in locking techniques and dealing with deadlocks, and introduce the concurrency control without locking.	
9-10	**Crash Recovery** More details in ARIES and log, introduce other recovery-related structures, study the Write-Ahead Log protocol and checkpointing, learning the procedures for recovering from a system crash.	Assignment 4
11-12	**Security and Authorization** Introduce database security, and the main considerations in designing a database application. Learn the mechanisms that a DBMS provides to control a user's access to data, and how is discretionary access control supported in SQL. Study the role of encryption in ensuring secure access, and how is it used for certifying servers and creating digital signatures.	Assignment 5
13	**Distributed Databases** Understand the motivation for distributed DBMSs, learn how is the data distributed across sites, and how can we evaluate and optimize queries over distributed data. Study the merits of synchronous vs. asynchronous replication, and the transaction management in a distributed environment.	
14	**Final Review**	

Contribution of course to meet the professional component:
This course prepares students with advanced knowledge in core techniques to database design and management.

Relationship to CS program objectives and outcomes:
This course primarily contributes to the Computer Science program outcomes that develop student abilities to:
(a) an ability to apply knowledge of mathematics, science, and engineering.
(c) an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability.
(e) an ability to identify, formulate, and solve engineering problems.
(k) an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice.
(l) an ability to use the computer/IT tools relevant to the discipline along with an understanding of their processes and limitations.

This course secondarily contributes to the Computer Science program outcomes that develop student abilities to:
(j) a knowledge of contemporary issues.

Relationship to CS program criteria:

Criterion	DS	PF	AL	AR	OS	NC	PL	HC	GV	IS	IM	SP	SE	CN
Scale: 1 (highest) to 4 (lowest)	4	3	3			4	2				1		2	4

Discrete Structures (DS), Programming Fundamentals (PF), Algorithms and Complexity (AL), Architecture and Organization (AR), Operating Systems (OS), Net-Centric Computing (NC), Programming Languages (PL), Human-Computer Interaction (HC), Graphics and Visual Computing (GV), Intelligent Systems (IS), Information Management (IM), Social and Professional Issues (SP), Software Engineering (SE), Computational Science (CN).

Course content distribution:

Percentage content for			
Mathematics	Science and engineering subjects	Complementary electives	**Total**
0%	100%	0%	100%

Persons who prepared this description:
Dr. Sam Chao

_____

Part B – General Course Information and Policies

1st Semester 2013/2014

Instructor: Dr. Sam Chao
Office: R108
Office hour: Wed 11:00 am – 12:30 pm, Thu 2:30 pm – 4:00 pm, Fri 10:00 am – 12:00 pm, or by appointment
Phone: 8397 8051
Email: lidiasc@umac.mo

Time/Venue: Wed 9:00 am – 11:00 pm, U105 (lecture)
Thu 4:00 am – 6:00 pm, NG02 (tutor/laboratory)

Grading distribution:

Percentage Grade	Final Grade	Percentage Grade	Final Grade
100 - 93	A	92 - 88	A–
87 - 83	B+	82 - 78	B
77 - 73	B–	72 - 68	C+
67 - 63	C	62 - 58	C–
57 - 53	D+	52 - 50	D
below 50	F		

Comment:
The objectives of the lectures are to explain and to supplement the text material. Students are responsible for the assigned material whether or not it is covered in the lecture. Students who wish to succeed in this course should read the assignments prior to the lecture and should work all homework and lab assignments, as well as the course project. You are encouraged to look at other sources (other texts, etc.) to complement the lectures and text.

Homework policy:
The completion of homework is a powerful learning experience; therefore:
- There will be approximately 4-6 homework assignments.
- Homework is due one week after assignment unless otherwise noted, no late homework is accepted.
- All assignments must be word-processed in A4 size, and all the diagrams must be developed by a drawing tool. No handwritten work will be accepted or counted for grading.
- There will be occasional in-class assignments during the class.
- The course grade will be based on the average of the HW grades.

Course project:
The project is probably the most exciting part of this course and provides students with meaningful experience to design and implement an application system with topics covered in the course:
- The project is either individual or in a group of 2.
- The application domain will be discussed further in class.
- The project will be presented by the groups at the end of the semester.

Exam:
- One 2-hour mid-term exam, one 3-hour final exam and 2-3 quizzes will be held during the semester. All exams are closed book examinations.

Note:
- Laboratory session is important part of this course and attendance is strongly recommended.
- Check course web pages for announcement, homework and lectures. Report any mistake on your grades within one week after posting.
- No make-up exam is given except for CLEAR medical proof.
- Cheating is absolutely prohibited by the university.

Appendix:

Rubric for Program Outcomes

Rubric for (a)	5 (Excellent)	3 (Average)	1 (Poor)
Understand the theoretic background	Students understand theoretic background and the limitations of the respective applications.	Students have some confusion on some background or do not understand theoretic background completely.	Students do not understand the background or do not study at all.
Use a correct model and formulation correctly	Students choose a model correctly and properly apply correct techniques.	Students choose a wrong model sometime, use a wrong formula, or a different technique.	Students use a wrong model and wrong formula, or do not know how to model.
Rubric for (c)	**5 (Excellent)**	**3 (Average)**	**1 (Poor)**
Design capability and design constraints	Student understands very clearly what needs to be designed and the realistic design constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability.	Student understands what needs to be designed and the design constraints, but may not fully understand the limitations of the design constraints.	Student does not understand what needs to be designed and the design constraints.
Process to meet desired needs	Student understands very clearly the process of the design.	Student understands what the needs of the process design, but may not fully understand the limitations of the design constraints.	Student does not understand the process.
Rubric for (e)	**5 (Excellent)**	**3 (Average)**	**1 (Poor)**
Modeling, problem formulation and problem solving	Students choose and properly apply the correct techniques.	Students model correctly but cannot select proper technique or model incorrectly but solve correctly accordingly.	Students at loss as to how to solve a problem.
Rubric for (j)	**5 (Excellent)**	**3 (Average)**	**1 (Poor)**
Relevance to the present time	Student displays an understanding of the theoretical or practical impact and an ability to correlate a subject, perception, communication, association and reasoning from a global and societal perspective.	Student is able to display an understanding of current topics and issues with some knowledge regarding their impact in a bigger global and societal sense.	Student has difficulty demonstrating an awareness or familiarity with current topics and issues relevant to most current global and societal affairs.
Rubric for (k)	**5 (Excellent)**	**3 (Average)**	**1 (Poor)**
Use modern principles, skills, and tools in engineering practice	Student applies the principles, skills and tools to correctly model and analyze engineering problems, and understands	Student applies the principles, skills and tools to analyze and implement engineering problems.	Student does not apply principles and tools correctly and/or does not correctly interpret the results.

	the limitations.		
Rubric for (l)	**5 (Excellent)**	**3 (Average)**	**1 (Poor)**
Use modern computer/IT tools relevant to the discipline	Student uses computer/IT tools relevant to the engineering discipline, and understands their limitations.	Student uses computer /IT tools relevant to the engineering discipline.	Student does not use computer/IT tools relevantly, and does not understand their limitations.