San José State University
College of Science
Department of Computer Science
CS 257, Database System Principle
Section 1, Fall 2016

Course and Contact Information

Instructor:	Tsau Young Lin
Office Location:	MH 214
Telephone:	(408)924-5121
Email:	tylin@cs.sjsu.edu
Office Hours:	TR 10:15-10:45 Friday (appointment only)
Class Days/Time:	TR 9:00 – 10:15
Classroom:	MH 222
Prerequisites:	CS 146 (with a grade of "C-" or better) or instructor consent

Course Description

CS257 is different from CS157A/B. In CS157A/B you learn how to use database and extended systems, but in CS257 you will learn how database and extended systems work internally.

Catalog Description:
(1) Design management and performance issues on: file organization and access methods, buffer management and storage management. (2) Query processing and query optimization, transaction management, recovery, and (3) concurrency control techniques. Reliability, (4) protection and integrity techniques. (5) Extensive programming project. Prerequisite: CS 157B or instructor consent.

Course Goals

The focus is on project; each team has different schedules

The impact of Big Data to CS257 is tremendous, so this semester; we focus on project that continues

(1) the building of a text/web based non SQL DBMS that was built in CS257-267 for several years (Big Data DBMS). In this part (1) and (2) will be addressed; note that the results of text/web analysis are stored in MS SQL. (3) Concurrent control are replaced by Map-Reduce; (4) due to time limit, we do not plan to cover

protection. However, we will do extensive data analysis that may address some integrity issues. In addition, a new data analysis technique that are derived from algebraic topology will be explained.

Programming outcomes:

- to use commercially available DBMS, such as Oracle, DB2, MS SQL Server and etc (we will use MS SQL Server)
- tohandle the interactions between DBMS and a third generation programming language, such as Java, C++.
- Ability to organize a team projectto build a “real world” modern database system(e.g., non-SQL DBMS)

Course Learning Outcomes

Important topics, database design theory and diagnosis are added.

1. Student learning outcomes: Upon successful completion of this course, students shouldhave a knowledge on

- the basicof system and user aspects of database systems.
- thebasic of database modeling including relational model, relation algebra and constraints, such as key constraints
- thebasic of database diagnosis and design theory (discovering the data anomaly and normalizations)
- thebasic skill in writing SQL, including embedded and dynamic SQL.

2. Programming outcomes:

- to use commercially available DBMS, such asOracle, DB2, MS SQL Server and (we will use MS SQL Server)
- tohandle the interactions between DBMS and a third generation programming language, such as Java, C++.
- Ability to organize a team project(using waterfall model) to build a “real world” database system (complexity is near a real world database system) that is based on a DBMS and a programming language

Required Texts/Readings

Textbook

Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer D. Widom.

Database Systems: The Complete Book. Prentice Hall. 2nd Ed ISBN-13: 978-0131873254. 2008.

Other Readings

Data Mining: Concepts and Techniques, Third Edition (The Morgan Kaufmann Series in Data Management Systems) By Jiawei Han and Micheline Kamber, 2012 Data Mining: Introductory and Advanced Topics by Margaret D. Dunham. Publisher: Prentice Hall, 2002. ISBN: 0130888923

Research papers on Rough Sets and Granular Computing

(Rough set theory is a name for algebraic theory of relational databases)

Course Requirements and Assignments

1. Projects: Understand “our” Concept Deep Data Analysis include Data Mining Engine
2. Exams/programing: 4 exams/short programs
3. Quizzes: Many unscheduled quizzes. Missing quizzes (up to 4 quizzes) can be made up by attending public technical talks sponsored by professional organizations, such as IEEE, ACM, AMS, and etc.
4. Homework:
5. Class Participation: Present some interesting topics in class, give some demos of short programs, or explaining hard home works in class will be properly awarded.

Final Examination or Evaluation

University Policy S06-4 (http://www.sjsu.edu/senate/docs/S06-4.pdf) An accumulative exam that includes lectures and project will be given at the time scheduled by the university.

Based two university policies:

- *University Syllabus Policy S16-9 (http://www.sjsu.edu/senate/docs/S16-9.pdf)*
- *University policy F15-12 (http://www.sjsu.edu/senate/docs/F15-12.pdf)*

Determination of Grades

Projects & home works	30%
All Exams and Quizzes (see below) will be returned and may discuss the details of exam grade	30%
Quizzes (Class average set 80 linearly)	10%
Final Exam (include lecture and project)	30%
Total	100%
90-92; 93-96;97-100	A
80-82; 83-86;87-89.99	B
70-72; 73-76;77-79.99	C
60-62; 63-66;67-69.99	D
<60	F

Classroom Protocol

I expect you to arrive promptly for every class meeting. If you do come in late, please take a seat quietly. Do not talk on a cell phone during class. If your phone rings, turn it off or leave the room. I would appreciate it if you

would refrain from talking to your neighbors while I am talking or while a classmate is trying to talk to me. A lot of people making tiny noises makes it very hard for me to hear.

University Policies

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/"

Course Schedule

Week	Date	Topics, Readings, Assignments, Deadlines
1		Overview of the class, policies (Green sheet) Project Description
2		Query Compiler (Introduction and Query Algebra)
3		Query Compiler (Introduction and Query Algebra)
4		Query Compiler (Introduction and Query Algebra)
5		Secondary Storage Management
6		Secondary Storage Management
7		Secondary Storage Management
8		Secondary Storage Management
9		Concurrency Control
10		Concurrency Control
11		Concurrency Control
12		Query Execution
13		Query Execution
14		Query Execution
15		Project demo
16		Project demo
17		Project demo
18		Review and Discussions
19	Friday, December 16 0715-0930	Final Exam