

San José State University
Computer Engineering Department
CmpE 130, *File Processing*, Section 1, Fall 2009

Instructor:	Dr. M. Robinson
Office Location:	Eng. 283B
Telephone:	924-3894
Email:	robinson@email.sjsu.edu
Office Hours:	M 1:30-3:20, TR 1:30-2:50 (subject to furlough restrictions)
Class Days/Time:	TR 1030-1145
Classroom:	Eng 337, lab periods in Eng 206 (as shown on schedule)
Prerequisites:	CmpE 126, ISE 130 (both with grade of "C" or better)

Faculty Web Page

Web page: www.engr.sjsu.edu/mrobins

Course Description

Physical and logical characteristics of data stored on external storage devices. File systems. Structures and access methods for data retrieval. Performance analysis.

Course Goals and Student Learning Objectives

1. To learn the techniques used to store information on rotating magnetic storage devices.
2. To learn basic techniques for storing information in a file.
3. To learn the problems involved in storing and retrieving very large quantities of data.

Course Content Learning Outcomes

Upon successful completion of this course, students will be able to:

1. learn the techniques used to store information on rotating magnetic storage devices. **(a, b, e, j, II)**
2. learn basic techniques for storing information in a file. **(a, b, c, e, I)**
3. learn the problems involved in storing and retrieving very large quantities of data. **(a, b, e, I)**
4. communicate with fellow engineers or customers through technical reports and design documentation. **(g, III)**

Required Texts/Readings

Textbook

File Structures, Second Edition, Michael J. Folk and Bill Zoellick, Addison-Wesley Publishing Company, 1991, **ISBN 10:** 0201557134, **ISBN 13:** 9780201557138.

or

File Structures: An Object-Oriented Approach with C++, Michael J. Folk, Bill Zoellick, and Greg Riccardi, Addison-Wesley Publishing Company, 1997, **ISBN:** 9780201874013 (hardback), 9780321210791 (paperback).

The book is out of print, so you will need to find a used copy. I recommend the 1991 edition over the 1997 edition, if you can find it. I do not recommend the 1987 edition.

Other Readings

Lecture notes will be available from Maple Press (San Carlos St. between 10th and 11th). The lecture schedule in the notes is not the same as this semester's lecture schedule, however the content is the same.

Classroom Protocol

Please be considerate of your fellow students and minimize distractions during the lecture.

Dropping and Adding

Students are responsible for understanding the policies and procedures about add/drops, academic renewal, etc. [Information on add/drops are available at http://info.sjsu.edu/web-dbgen/narr/soc-fall/rec-324.html](http://info.sjsu.edu/web-dbgen/narr/soc-fall/rec-324.html) . [Information about late drop is available at http://www.sjsu.edu/sac/advising/latedrops/policy/](http://www.sjsu.edu/sac/advising/latedrops/policy/) . Students should be aware of the current deadlines and penalties for adding and dropping classes.

Assignments and Grading Policy

Projects: Students have remote access to the department's Linux systems. Lab problems will be posted on the class webpage. The lab projects may be developed in any C or C++ environment, but must run under Linux. If you wish to develop your lab work in a different environment, allow extra time for the first lab project to discover the differences between your system and the Linux systems. Usually these are fairly minor, but it may take some time to discover what they are.

In order to receive credit for a lab problem you must submit correctly functioning and documented source code and a working makefile, and also turn in a lab report. For full credit, submission of the code and makefile must occur before midnight on the due date.

Your code must be documented according to standards which will be provided. Your lab report MUST follow the format specified in the problem statement. The assignment will not receive credit unless these requirements are met.

Examinations: There will be two midterm examinations and a comprehensive final examination.

Quizzes: There may be some short quizzes during the semester, announced or unannounced. The content of these quizzes may also cover the prerequisites. If there are quizzes, they will count as part of the homework total for the final grade.

Homework: Homework problems will be posted on the class webpage and announced in class. You may prepare your homework as a TEAM of up to THREE members. Make ONE submission as a team, and make sure that all team members sign the homework paper, and that all three names are PRINTED on the paper. If your name doesn't appear on an assignment, you won't receive credit for it.

Your homework will be checked for evidence of original effort, but not necessarily for correct answers.

The purpose of the homework is to enhance your learning effort so that you can do the lab work, pass the quizzes and pass the final examination. You should make an earnest attempt to solve the homework exercises on your own. Letting someone else do the work will not help your comprehension of the material. Make sure that YOU have made a good effort at solving EACH exercise.

Late homework will NOT be accepted.

The grading percentages will be approximately as follows. Any category may be changed by $\pm 10\%$ at the discretion of the instructor.

Lab Projects	15%
Homework & Quizzes	10%
2 Midterm Exams	35%
Final Exam	40%
Total	100%

- **Note 1:** There will be a penalty for late programming projects of 1% per day. After the drop dead date, late projects will not be accepted. The drop dead date is generally the class meeting before the next project is due.
- **Note 2:** The lab projects have different weights according to the amount of work required.
- **Note 3:** Late homework will NOT be accepted.
- **Note 4:** No makeup quizzes or exams will be given. If you must miss an exam for reasons beyond your control, speak to me first. Do not ask after the exam.

Grading calculations

Final scores are evaluated with respect to your peers, although absolute standards will be imposed as well. There are several conditions that will result in receiving a FAILING GRADE in this course:

- Not receiving a sufficiently high score, calculated as shown above.
- Not having completed the required prerequisite courses before taking the course.
- Academic dishonesty (see below).

University Policies

Academic integrity

Students should know that the University's [Academic Integrity Policy is available at http://www.sa.sjsu.edu/download/judicial_affairs/Academic_Integrity_Policy_S07-2.pdf](http://www.sa.sjsu.edu/download/judicial_affairs/Academic_Integrity_Policy_S07-2.pdf). Your own commitment to learning, as evidenced by your enrollment at San Jose State University and the University's integrity policy, require you to be honest in all your academic course work. Faculty members are required to report all infractions to the office of Student Conduct and Ethical Development. The website for [Student Conduct and Ethical Development is available at http://www.sa.sjsu.edu/judicial_affairs/index.html](http://www.sa.sjsu.edu/judicial_affairs/index.html).

Instances of academic dishonesty will not be tolerated. Cheating on exams or plagiarism (presenting the work of another as your own, or the use of another person's ideas without giving proper credit) will result in a failing grade and sanctions by the University. For this class, all assignments are to be completed by the individual student unless otherwise specified. If you would like to include in your assignment any material you have submitted, or plan to submit for another class, please note that SJSU's Academic Policy F06-1 requires approval of instructors.

Campus Policy in Compliance with the American Disabilities Act

If you need course adaptations or accommodations because of a disability, or if you need to make special arrangements in case the building must be evacuated, please make an appointment with me as soon as possible, or see me during office hours. Presidential Directive 97-03 requires that students with disabilities requesting accommodations must register with the DRC (Disability Resource Center) to establish a record of their disability.

CmpE130, File Processing, Fall 2009 Course Schedule

The schedule is subject to change with fair notice on the class web page. The lab periods are a new idea this semester and it may become clear that they could be better scheduled on different days. This schedule may be changed due to university policies that have not been fully determined at the time of writing.

Lec.	Date	Topic	Reading	HW	Lab
1	25 Aug..	Introduction	Ch. 1		
2	27	Files—User view	Ch. 2	1. UNIX	
3	1 Sept.	<i>Lab</i>			1. Binary tree
4	3	Disk performance	Ch. 3	2. Dsk. Perf.	
5	8	File systems and I/O	Ch. 3		
6	10	File structure concepts	Ch. 4		
7	15	Managing files of records	Ch. 5	3. File Struc.	
8	17	<i>Lab</i>			2. Fields
9 ¹	22	gdb debugger	Notes		
10	24	Performance issues	Ch. 6	4. Perf.	
11	29	Random access—Indexes	Ch. 7		
12	1 Oct.	<i>Lab</i>		5. Index	3. Index
13	6	Random access—Secondary indexes	Ch. 7		
14	8	Review			
15	13	Midterm Exam 1			
16	15	B-Trees—intro	Ch. 9		
17	20	B-Trees—algorithms	Ch. 9	6. B-trees	
18	22	<i>Lab</i>			4. B-tree
19	27	B-Trees—extensions to algorithms	Ch. 9		
20	29	B+ Trees	Ch. 10	7. B+-trees	
21	3 Nov.	O(1) access (hashing)	Ch. 11	8. Hashing	
22	5	O(1) access (hashing)	Ch. 11		
23	10	Review			
24	12	Midterm Exam 2			
25	17	<i>Lab</i>			5. Hashing
26	19	Extendible hashing	Ch. 12		
27	24	Extendible hashing	Ch. 12	9. Ext. hash	
	26	<i>No class—Thanksgiving</i>			
28	1 Dec.	Cosequential processing of large files	Ch. 8		
29	3	Cosequential processing of large files	Ch. 8		
30	8	<i>Lab</i>			
Final exam	14	Eng 337, 0945-1200			

¹ Faculty furlough day designated by the President. No class.