University at Buffalo
CSE462/562 Datbase Systems (Spring 2024)

Instructor: Zhuoyue Zhao, zzhao35@buffalo.edu.
Lecture location & time: Monday 4:00 pm to 6:40 pm, Know 109.
Instruction Mode: Lectures.
Deliver Mode: In-person.
Grader: Shawn Zimmermann.
Office hours: Wednesdays, 1 pm - 3 pm. See course website for updates.
Course website: `https://cse.buffalo.edu/~zzhao35/teaching/cse562_spring24/`
Annoucement and discussion board: We will be using Piazza for announcements, discussion and Q&A outside the lecture and office hours: `https://piazza.com/buffalo/spring2024/cse462562/home`
Number of Credits: 3
Course Description:
Database Management Systems (DBMS) are an important piece of software widely used in various data intensive applications. No matter what data model or query language it adopts, there are principles and methodologies commonly found in virtually all DBMS, in order to support efficient and fault-tolerant data storage, query and management. In this course, we will examine the internals of a traditional Relational DataBase Management System (RDBMS) and discuss the common principles and methodologies that may be useful in many other database and data processing systems beyond RDBMS. The students will also gain practical experience of efficient data management through a semester-long project of implementing various components of a mini RDBMS.
The main topics include database storage management, query processing, query optimization, transaction processing, concurrency control and recovery in RDBMS. Please note that this is not a course on database design or database application development – those are covered in CSE460/560 Data Models and Query Languages. We will briefly review the necessary background on the relational data model and the SQL language in the first a few lectures but students are expected to understand relational algebra and SQL before this course.
Learning Outcomes:
The students should demonstrate mastery of the concepts/skills/knowledge expressed in the following ABET learning outcomes for computer science:

- (2) Design, implement and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.
- (5) Function effectively as a member or a leader of a team engaged in activities appropriate to the program's discipline.
- (6) Apply computer science theory and software development fundamentals to produce computing-based solutions.

Upon completion of this course, a student is expected to

- Be able to design and implement Relational DataBase Management Systems with basic storage and query processing capabilities (ABET 2; **Instructional Method**: Lectures; **Assessment Methods**: Homeworks, Exams, Projects);
- Be able to design and analyze efficient algorithms and data structures for data management systems (ABET 6; **Instructional Method**: Lectures; **Assessment Methods**: Homeworks, Exams, Projects);

- Be able to work in a team effectively for a complex software system project (ABET 5; **Instructional Method**: Lectures; **Assessment Methods**: Projects).

The Student Outcomes from the Computing Accreditation Commision (CAC) of the ABET have been adopted.
Program Outcome Support (Computer Science ABET Outcomes):

Program Outcome	1	2	3	4	5	6
Support Level	No coverage	Demonstrate mastery of skill/concept	No coverage	No coverage	Demonstrate mastery of skill/concept	Demonstrate mastery of skill/concept

Course Prerequisites:
Required: Soid background in programming and data structures. Programming experience with C++. For CSE462 students, CSE460 is a mandatory prerequisite.
No mandatory textbook: We will make all the lecture slides available on the course website.
Optional textbook:

- Database System Concepts, Seventh Edition, by Abraham Silberschatz and Henry Korth and S. Sudarshan.

Course Requirements:

- You are expected to understand and follow the academic integrity policy of UB, CSE and this course. We will *strictly* enfore the academic integrity policy (see below).
- In-person attendance is *required*. You should attend all the lectures and be familiar with all the course materials presented in class, which may appear in any course projects and exams. We will try to make the recording of the lectures available, but we cannot guarantee the availablility of those in case of technical difficulties. Please note that there is no live streaming of the lectures.
- There will be a mid-term exam and a final exam. They must be attended in person and there will be no make-up exam except for the exam/course conflicts made known to the instructor within two weeks of the spring semester (2/7/2024), or unusual circumstances on a case-by-base basis. The exams are open-book but you may only bring the paper-copy of lecture slides, the written assignments, the textbook or your lecture notes. You should not use any electronic device during the exame, except for a calculator that may be needed.
- There is a semester-long course project of building a mini RDBMS in C++, divided into six subprojects. You may work individually or collaborate in a team of two. The tasks in each project will include a coding component, which you may collaborate with your teammate, and a write-up, which you must complete independently without any collaboration or sharing with your teammate. The tasks of the projects remain the same regardless of the team size, and you will receive the same grade for the coding component of your team. The write-up will be graded individually. We will use Autolab for code submission and UBLearns for write-up submission. Please refer to the course website for details.
- There will be 6 graded written assignments. The wrriten assignments must be completed independently. Collaboration with classmates (including your project teammate) or using any matierals other than the lecture slides and the optional textbook is not allowed. The lowest two grades are excluded from the final grade.
- Late submission policy: you have up to 3 grace days in total for projects and written assignments, and you may only use at most 1 grace day for each project or written assignment. There is no penalty for late submissions that fall in the the allowed grace days. No credit will be given if you make a late submission that uses more than allowed grace days. If you have a project teammate and any of you make a late submission, any grace days used will count towards the used grace days of both team members.

Important Dates:

- Add/Drop deadline: January 31, 2024.

- Mid-term exam: March 27, 2024, 7:05 pm - 8:25 pm, Knox 104.
- Last day to resign from the course: April 16, 2024.
- Final exam: May 15, 2024, 3:40 pm - 5:20 pm, Knox 109.
- Project due dates: please refer to the course website.

Grading Policy:
The final grade will be broken down as follows:

- Mid-term exam: 20%.
- Final exam: 20%.
- Written assignments: 20%, 5% each, the lowest two excluded
- Projects: 40% + up to 10% extra bonus.

And the following is the (tentative) assignment of letter grades:

- $[90, 110]$: A
- $[80, 90)$: A-
- $[70, 80)$: B+
- $[60, 70)$: B
- $[50, 60)$: B-
- $[40, 50)$: C+
- $[30, 40)$: C
- $[20, 30)$: C-
- $[10, 20)$: D
- $[0, 10)$: F

The grades will not be curved and will not be differentiated between the sections.
Course Schedule:
The following is the tentative course schedule and may be changed throughout the semetser. Please refer to the course website for the latest schedule.

Lectures

Week#	Topics	Textbook chapter
1	Course introduction and logistics Physical storage	Ch. 1 Ch. 12
2	Buffer management Data storage layout	Ch. 13.5 Ch. 13.1, 13.2, 13.3
3	Access methods and indexing Hasing techniques	Ch. 13.6, 14.1, 14.2 Ch. 14.5, 24.5
4	Hasing techniques (cont'd) Hash index	 Ch. 14.3, 14.4
5	Tree index	 Ch. 14.1, 14.2
6	Index and Cost Analysis	Ch. 14.1, 14.2
7	Relational model and SQL	Ch. 2, 3
8	Spring recess, no lectures	
9	Query processing overview Midterm review Midterm exam	Ch.15.1, 15.2
10	Single-table query processing	Ch. 15.3, 15.6, 15.7
11	Single-table query processing (cont'd) External sorting	 Ch. 15.4
12	Join algorithms	Ch. 15.5
13	Query optimization	Ch. 16
14	Transaction processing Pessimistic Concurreny Control	Ch. 17 Ch. 17, 18.1 - 18.4
15	Pessimistic concurrency control (cont'd) Crash recovery	 Chapter 19
16	Final exam	

Written Assignments

Written Assignment	Release Date	Due Date
HW1: buffer and storage	2/5/2024	2/18/2024
HW2: hashing	2/19/2024	3/3/2024
HW3: tree indexes and cost analysis	3/4/2024	3/17/2024
HW4: query processing	4/8/2024	4/21/2024
HW5: query optimization	4/15/2024	4/28/2024
HW6: concurrency control and crash recovery	4/29/2024	5/12/2024

Projects

Assignment	Release Date	Due Date
Project 1 - Lab 0: project and team sign-up	1/29/2024	2/1/2024
Project 1 - Lab 1: File I/O	1/29/2024	2/4/2024
Project 2: Storage Layer	2/7/2024	2/25/2024
Project 3: B-tree	3/4/2024	3/24/2024
Project 4: Query processing: basic operators and external sorting	4/1/2024	4/14/2024
Project 5: Query processing and optimization: joins	4/22/2024	5/12/2024

Academic Integrity Policy:
Academic integrity is critical to the learning process. It is your responsibility as a student to complete your work in an honest fashion, upholding the expectations your individual instructors have for you in this regard. The ultimate goal is to ensure that you learn the content in your courses in accordance with UB's academic integrity principles, regardless of whether instruction is in-person or remote. Thank you for upholding your

own personal integrity and ensuring UB's tradition of academic excellence. You should get familiar with the departmental and the university academic integrity policies and procedures for *graduate students* and for *undergraudate students*

In this course, you may **NOT** discuss/share code with/copy code from anyone about your course projects except your project teammate. For project write-ups, you must submit your **independent** write-ups. We also require all students, whether enrolled, dropped or resigned from the class, to keep your course project repository inaccessible to public indefinitely, and never share it with any current or future students who may take the course. Examples that we consider as academic integrity violations include but are hot limited to: 1) copying any part of other team's code implementation/code found online in the course project, regardless of whether it is a verbatim copy or a modified copy; 2) collaborate with other teams in the course project; 3) making your course project repository or any of its copy available to any current or future students of the course; 4) submitting work that is not created by you or your team, including deriving your solutions based on materials from search engines, tutoring services, Q&A websites, and/or those generated by generative AI software (e.g., ChatGPT); 5) cheating or referring to any material not permitted in the exams.

Consistent with the CSE policy, we have **zero tolerance** towards academic integrity violations. Any academic integrity violation will result in an F grade for all students involved, and will be referred to the Office of Academic Integrity. The only exception is when the violation is accidental and does not provide any unfair advantage to any of the students involved, in which case you may receive a reduction in grade in the particular assignment/exam/project.

Campus resources:

Accessibility Resources. If you have any disability which requires reasonable accommodations to enable you to participate in this course, please contact the Office of Accessibility Resources in 60 Capen Hall, 716-645-2608 and also the instructor of this course during the first week of class. The office will provide you with information and review appropriate arrangements for reasonable accommodations, which can be found on the web at: http://www.buffalo.edu/studentlife/who-we-are/departments/accessibility.html.

Counseling Services. As a student you may experience a range of issues that can cause barriers to learning or reduce your ability to participate in daily activities. These might include strained relationships, anxiety, high levels of stress, alcohol/drug problems, feeling down, health concerns, or unwanted sexual experiences. Counseling, Health Services, and Health Promotion are here to help with these or other concerns. You learn can more about these programs and services by contacting:

Counseling Services: 120 Richmond Quad (North Campus), phone 716-645-2720
202 Michael Hall (South Campus), phone: 716-829-5800
Health Services: Michael Hall (South Campus), phone: 716- 829-3316
Health Promotion: 114 Student Union (North Campus), phone: 716- 645-2837

Sexual Violence. UB is committed to providing a safe learning environment free of all forms of discrimination and sexual harassment, including sexual assault, domestic and dating violence and stalking. If you have experienced gender-based violence (intimate partner violence, attempted or completed sexual assault, harassment, coercion, stalking, etc.), UB has resources to help. This includes academic accommodations, health and counseling services, housing accommodations, helping with legal protective orders, and assistance with reporting the incident to police or other UB officials if you so choose. Please contact UB's Title IX Coordinator at 716-645-2266 for more information. For confidential assistance, you may also contact a Crisis Services Campus Advocate at 716-796-4399.