COMP 421: Files and Databases

Fall 2025

The University of North Carolina at Chapel Hill

Course Information

Credit Hours: 3 credit-hours

Pre or Co-Requisites: Prerequisites, COMP 210, 211, and 301; or COMP 401, 410, and 411; a grade of C or better is required in all prerequisite courses.

Target Audience: CS majors who have completed the intro sequence, and who have an interest in systems

Meeting Pattern: Monday, Wednesday; 12:20-1:35;

Instructional Format: In-person

Classroom or Location: Murphey 0116

Instructor Information

Name: Benjamin Berg (Ben) [he/him]

Email Address: ben@cs.unc.edu

Office Location: FB 336

Office Hours: TBD

Course Content

Course Description

This is an introduction to modern database systems. This class does not emphasize database applications, but rather focuses on the design and implementation of database management systems (DBMS). We will focus on the design of disk-based relational DBMSs. Topics will include: the relational algebra, storage management, memory management, data structures used in modern DBMSs, indexing, query execution, query optimization, transaction processing, recovery, concurrency control, and distributed DBMS designs.

The course will consist of lectures, midterm and final exams, and several large programming assignments aimed at implementing pieces of a DBMS. Assignments will be written in modern C++.

While SQL will be covered briefly for the sake of completeness, this is not a class on writing SQL queries or developing database applications.

Course Texts & Materials

The official course textbook is [Database Systems Concepts (7th edition)](), Silberschatz, et al.

For additional background on OS concepts, we recommend [Operating Systems: Three Easy Pieces](), Arpaci-Dusseau & Arpaci-Dusseau.

For references on the C++ language, we recommend [A Tour of C++]() and [Effective Modern C++](), two classic books that are widely available.

Students will be referred to other additional reading in lectures and on the course web page.

Class Expectations

The central expectation of the course is that students and instructors adhere to the [Reasonable Person Principle]().

Students should attend lectures and try to answer any questions either during Q&A periods of lecture or by coming to office hours. Solving the course projects will almost certainly require students to come to office hours. If a student cannot attend posted office hours, appointments can be made with the instructors. However, the instructors will not answer questions via email.

AI and Collaboration Policies

There is no group collaboration in this course. Students should submit their own work for both course projects and the midterm and final exams. Plagiarism will not be tolerated, and the UNC honor code will be enforced. These rules extend to sharing or publishing of solutions before, during, or after the course. Releasing your solutions in any form is an honor code violation. Additionally, because this course is based on open source course materials, releasing your solutions is harmful to CS Academic community beyond UNC. Please be a good citizen of the larger CS community.

Instructors/TAs/LAs will help students with high level advice on projects and will assist students in learning how to use tools to debug their code. **However, instructors will not debug student code**. Similarly, students shall not give or receive debugging help from other students. Discussions should be conceptual in nature, not focused on debugging specific issues.

The use of generative AI in this course is (i) probably inevitable (ii) permissible in some cases. The permissible uses fall broadly into two categories. First, students may use generative AI to answer high-level questions. For example "What is the default storage layout in postgres?". Second, limited use of AI-based coding assistance is permissible subject to the following condition.

In all cases, students will be held to the following standard regarding generative AI, collaboration, or other outside sources: Students are responsible for ensuring that submitted work is not substantially similar to existing publicly or privately available materials. Submitting work that is substantially similar to existing work is an honor code violation. In particular, students use generative AI tools at their own risk.

As a word of advice, current AI coding assistants are now good enough to generate a lot of code, very quickly, that may do approximately what you ask for. However, the assignments in this class are generally complex enough that generated solutions are unlikely to work immediately. If you generate a lot of code without understanding what it does, you will ultimately struggle to debug your program. Remember, **instructors will not help you debug your code**, so you are only hurting yourself by doing this.

Finally, the above policies do not apply to the midterm or final exams. Both exams will be completed without the use of any computing device. The use of AI or any computing device during the completion of an exam is an honor code violation.

Course Goals & Student Learning Outcomes (SLOs)

Beyond gaining general knowledge in the area of database systems, there are three high-level goals of this course.

First, through the lecture material, students should learn to evaluate and compare the designs of large-scale, real-world computer systems. This includes examining choices made by system designers in industry, as well as integrating concepts from the vast academic literature on database systems. Students should leave the class knowing how discuss and evaluate state-of-the-art work on databases and other computer systems. A student's ability in this area will be tested on the midterm and final exams.

Second, through the course projects, students should become comfortable writing performant, concurrent code in a lower-level language (C++). This will require students to understand how the software they write interacts with a modern operating system and modern hardware.

Third, to achieve a good grade on the course projects, students will be asked to both debug and optimize their code using industry-standard debuggers and performance profilers. These skills are broadly applicable to the implementation of any large-scale computer system, not just databases. Unlike some other CS classes, students will be asked to debug and optimize code independently, with minimal TA assistance or group work. This will ensure that the course projects help students develop competence in these tasks.

Course Assignments & Assessments

Assignment Descriptions

There will be a midterm exam, a final exam, and four programming projects which are expected to work-intensive for the students. If time allows, there may be a fifth programming project.

The midterm exam will cover material from the first ~15 lectures. The final exam will be cumulative, but will emphasize material from the second half of the course. The programming projects are (subject to change):

0. HyperLogLog, A C++ primer: where students implement a data structure commonly used in database systems as a means to become more familiar with features of C++.

1. Buffer Pool Manager: implement the part of a disk-oriented DBMS that manages the movement of pages between disk and DRAM.
2. B+ Tree: a concurrent implementation of the ubiquitous B+ Tree algorithm used for indexing in databases.
3. Query Executor: students implement the query operators, execution mechanisms, and some example optimizations used in query execution.
4. (There may be a fifth project on multi-version concurrency control depending on how these first projects proceed. If so, the below grading scheme will be rescaled.)

The value of the coursework will be assigned as follows:

Project 0: 10%
Project 1: 20%
Project 2: 20%
Project 3: 20%

Midterm exam: 10%
Final exam: 15%

Participation: 5%

Grading Scale & Schema

Late Work

Students will receive 4 late days to use at their discretion. Late days can be used for any project **besides project 0.** Outside of this late day policy, no late work will be accepted without a university-approved excuse.

Grading Rubrics

Programming assignments will be autograded on gradescope using a mixture of available and hidden test cases. Other code formatting requirements will be detailed on the course page.

Exam responses will be graded using a rubric for each question that will be made available upon request.

Grading Scale

Grades may be rescaled at the instructors' discretion to improve letter grades. This rescaling will never decrease your grade.

Course Schedule

Lecture	Date	DoW	Topic	Project Info
0	8/18	M	Course Overview, Files	P0 Released
1	8/21	W	Relational Algebra	
2	8/25	M	SQL	
2.5	8/27	W	SQL II	

	9/1	M	LABOR DAY, NO CLASS	
3	9/3	W	Storage I	P0 Due
4	9/8	M	Storage II	P1 Released
5	9/10	W	Storage Models, Compression	
	9/15	M	WELL BEING, NO CLASS	
6	9/17	W	Memory Management	
7	9/22	M	Hash Table	
8	9/24	W	B Trees	P2 Released
9	9/29	M	Tries, Bloom Filters	P1 Due
10	10/1	W	Concurrent Data Structures	
11	10/6	M	Extrenal Sort, External Aggregation	
12	10/8	W	Joins	
	10/13	M	Midterm Review	
	10/15	W	Midterm Exam	
13	10/20	M	Query Execution I	
14	10/22	W	Query Execution II	P3 Released
15	10/27	M	Query Optimization	
16	10/29	W	Concurrency Control	P2 Due
17	11/3	M	Two phase locking	
	11/5	W	NO CLASS	
18	11/10	M	Timestamp Ordering	
19	11/12	W	MVCC	
20	11/17	M	Logging	
21	11/19	W	Recovery	
22	11/24	M	Distributed Databases	P3 Due
23	11/26	W	THANKSGIVING, NO CLASS	
	12/1	M	Distibuted OLTP	
24	12/3	W	Distributed OLAP	
25	12/5	M	Final Review Q+A	
	12/12	F	Final Exam	

Policy Statements

Academic Policies

University Class Attendance Policy

University Policy: As stated in the University's Class Attendance Policy, no right or privilege exists that permits a student to be absent from any class meetings, except for these University Approved Absences:

1. Authorized University activities: University Approved Absence Office (UAAO) website provides information and FAQs for students and FAQs for faculty related to University Approved Absences
2. Disability/religious observance/pregnancy/short-term military service, as required by law and approved by the University Compliance Office , or in the case of short-term military service, the Dean of Students
3. Significant health condition and/or personal/family emergency as approved by the Office of the Dean of Students, Gender Violence Service Coordinators, and/or the University Compliance Office.

Code of Conduct

All students are expected to adhere to University policy and follow the guidelines of the UNC Student Code of Conduct. Additional information can be found at https://studentconduct.unc.edu/.

Artificial Intelligence (AI) Use Policy

All Carolina students are expected to follow these AI guidelines:

1. AI should help you think, not think for you. You may be able to use these tools to brainstorm ideas, research topics, and analyze problems, but you must decide what's appropriate and accurate.
2. Engage responsibly with AI. You must evaluate AI-generated outputs for potential biases, limitations, inaccuracies, false output, and ethical implications. Do not put personal or confidential data into these tools.
3. The use of AI must be open and documented. You should declare, explain, and cite any use of AI in the creation of your work using applicable standards (e.g., APA, MLA, course guidelines). Understand that you are ultimately 100% responsible for your final product.
4. Follow specific AI guidelines in this syllabus. If you are unsure, check with me. Guidance offered in this syllabus would be referenced should an issue be referred to Student Conduct for alleged academic misconduct.
5. *[Add details that apply to specific circumstances, assignments, or activities within the course. Additional guidance for faculty is available on the AI @ UNC website.]*

Syllabus Changes

The instructor reserves the right to make changes to the syllabus including project due dates and test dates. These changes will be announced as early as possible.

Acceptable Use Policy

By attending the University of North Carolina at Chapel Hill, you agree to abide by the University of North Carolina at Chapel Hill policies related to the acceptable use of IT systems and services. The Acceptable Use Policy (AUP) sets the expectation that you will use the University's technology resources responsibly, consistent with the University's mission. In the context of a class, it's quite likely you will participate in online activities that could include personal information about you or your peers, and the AUP addresses your obligations to protect the privacy of class participants. In addition, the AUP addresses matters of others' intellectual property, including copyright. These are only a couple of typical examples, so you should consult the full Information Technology Acceptable Use Policy, which covers topics related to using digital resources, such as privacy, confidentiality and intellectual property.

Additionally, consult the [Safe Computing at UNC](#) website for information about data security policies, updates, and tips on keeping your identity, information, and devices safe.

Data Security & Privacy

UNC-Chapel Hill is committed to fulfilling its responsibilities of transparency as a state-sponsored institution of higher learning, protecting certain types of information, and using information Carolina collects only for appropriate purposes. Consult the [UNC-Chapel Hill Privacy Statement](#) for additional information.

Grade Appeal Process

If you have any concerns with grading and/or feel you have been awarded an incorrect grade, please discuss it with me as soon as possible. If we cannot resolve the issue, you may talk to our director of undergraduate studies or department chair.

Services & Student Support Policies

University Compliance Office (formerly Equal Opportunity and Compliance) - Accommodations

University Compliance Office (UCO) Accommodations Team ([Accommodations - UNC Equal Opportunity and Compliance)](#) receives requests for accommodations for disability, pregnancy and related conditions, and sincerely held religious beliefs and practices through the University's Policy on Accommodations. UCO Accommodations team determines eligibility and reasonable accommodations consistent with state and federal laws.

Counseling & Psychological Services (CAPS)

UNC-Chapel Hill is strongly committed to addressing the mental health needs of a diverse student body. The [Heels Care Network](#) website is a place to access the many mental health resources at Carolina. CAPS is the primary mental health provider for students, offering timely access to consultation and connection to clinically appropriate services. Go to the [CAPS website](#) or visit their facilities on the third floor of the Campus Health building for an initial evaluation to learn more. Students can also call CAPS 24/7 at 919-966-3658 for immediate assistance.

Title IX Resources

Any student who is impacted by discrimination, harassment, interpersonal (relationship) violence, sexual violence, sexual exploitation, or stalking is encouraged to seek resources on campus or in the community. Reports can be made [online to the UCO](#) or by contacting the [University's Title IX Coordinator](#), Elizabeth Hall, or the [Report and Response Managers](#) in the University Compliance Office (UCO) (formerly the Equal Opportunity and Compliance Office). Please note that I am designated as a Responsible Employee, which means I must report to the UCO any information I receive about the forms of misconduct listed in this paragraph. If you'd like to speak with a confidential resource, those include Counseling and Psychological Services, the University's Ombuds Office, and the [Gender Violence Services Coordinators](#). Additional resources are available at [safe.unc.edu](#).

Policy on Non-Discrimination

As set out in the University's [Policy Statement on Non-Discrimination](#), the University is committed to providing an environment where all members of our community can learn, work, and thrive. Consistent with these principles and applicable laws, it is therefore the University's policy not to discriminate on the basis of age, color, disability, gender, gender expression, gender identity, genetic information, national origin, race, religion, sex, sexual orientation or veteran status as consistent with the University's [Policy on Prohibited Discrimination, Harassment and Related Misconduct](#). No person, on the basis of protected status, shall be excluded from participation in, be denied the benefits of, or be subjected to unlawful discrimination, harassment, or retaliation under any University program or activity, including with respect to employment terms and conditions. The University will consider only relevant factors such as individual

abilities and qualifications in admissions, hiring, disciplinary action, and all other decisions and will apply consistent standards of conduct and performance.

If you are experiencing harassment or discrimination, you can seek assistance and file a report through the Report and Response Coordinators (email reportandresponse@unc.edu or see additional contact info at safe.unc.edu) or the University Compliance Office. Please note that I am designated as a Responsible Employee, which means that I must report to the UCO any information I receive about harassment or discrimination. If you'd like to speak with a confidential resource, those include Counseling and Psychological Services and the University's Ombuds Office.

Undergraduate Testing Center

The College of Arts and Sciences provides a secure, proctored environment in which exams can be taken. The Center works with instructors to proctor exams for their undergraduate students who are not registered with ARS and who do not need testing accommodations as provided by ARS. In other words, the Center provides a proctored testing environment for students who are unable to take an exam at the normally scheduled time (with pre-arrangement by your instructor). For more information, visit the testing center website.

Learning Center

Want to get the most out of this course or others this semester? Visit UNC's Learning Center to make an appointment or register for an event. Their free, popular programs will help you optimize your academic performance. Try academic coaching, peer tutoring, STEM support, ADHD/LD services, workshops and study camps, or review tips and tools available on the website.

Writing Center

For free feedback on any course writing projects, check out UNC's Writing Center. Writing Center coaches can assist with any writing project, including multimedia projects and application essays, at any stage of the writing process. You don't even need a draft to come visit. To schedule a 45-minute appointment, review quick tips, or request written feedback online, visit UNC's Writing Center online.