CAS CS 460: Introduction to Database Systems

Boston University, Fall 2023

Syllabus

Description
This course covers the fundamental concepts of database systems. Topics include data models (ER, relational, and others); query languages (relational algebra, SQL, and others); implementation techniques of database management systems (index structures, concurrency control, recovery, and query processing); management of semistructured and complex data; distributed and noSQL databases.

Prerequisites
CAS CS 112, or the equivalent

Instructor
David G. Sullivan, Ph.D. (dgs@bu.edu)
See the course website for the instructor and TA office hours.

Other Course Staff
Edwyn Song (esong501@bu.edu), course facilitator and teaching assistant
Katie Rimey (rimey@bu.edu), teaching assistant
Junsun (Lucas) Yoon (lyoon02@bu.edu), course assistant

Lectures and Labs
lectures: MWF, 1:25-2:15 pm, KCB 101
lab: a weekly session; see your schedule for the time and location

Course Website: https://cs-people.bu.edu/dgs/courses/cs460
In addition, announcements and some course materials will be posted Blackboard.

Requirements and Grading
1. Five problem sets (30% of the final grade)
2. A midterm exam (25%) and a final exam (35%)
3. Participation (10%; see below)

To pass the course, you must have a passing average on the problem sets and a passing average across the two exams.

Course Materials
- **Required:** CS 460 Coursepack. This contains all of the lecture notes for the course. More detail will be provided in class and in Lab 0.
- **Optional**: *Database Systems: The Complete Book (2nd edition)* by Hector Garcia-Molina, Jeff Ullman, and Jennifer Widom (ISBN 978-0131873254, Pearson Prentice Hall, 2009). This book is *not* required.
- **Required:** We will be using the Top Hat Pro platform. More detail will be provided in class.

Collaboration Policy
You are strongly encouraged to collaborate with one another in studying the lecture materials and preparing for quizzes and exams.

Problem sets will primarily involve *individual-only* problems that you must complete on your own. We may occasionally include a *pair-optional* problem that you may complete either alone or with a partner.

For both types of problems, you may discuss ideas and approaches with others (provided that you acknowledge this in your solution), but such discussions should be kept at a high level and should not involve actual details of the code or of other types of answers. **You must complete the actual solutions on your own** (or, in the case of a pair-optional problem, with your partner if you choose to use one).

Rules for working with a partner on pair-optional problems:

- You may *not* work with more than one partner on a given assignment. (However, you are welcome to switch partners between assignments.)
- **You may *not* split up the work and complete it separately.**
- **You must work together** (at the same computer or via a Zoom meeting) for all problems completed as a pair, and your work must be a collaborative effort.
- You and your partner must *both* submit the same solution to each problem that you did as a pair, and you must clearly indicate that you worked on the problem as a pair by putting your partner's name at the top of the file.

Academic Misconduct
We will assume that you understand BU's Academic Conduct Code:
http://www.bu.edu/academics/policies/academic-conduct-code
You should also carefully review the CS department's page on academic integrity:
http://www.bu.edu/cs/undergraduate/undergraduate-life/academic-integrity

Prohibited behaviors include:

- copying all or part of someone else's work, even if you subsequently modify it; this includes cases in which someone tells you what to write for your solution
- viewing all or part of someone else's work (with the exception of work that you and your partner do together on a pair-optional problem)
- showing all or part of your work to another student (with the exception of work that you and your partner do together on a pair-optional problem)
- giving another student access to your laptop unless you monitor their usage
- consulting solutions from past semesters, or those found online or in books
- using ChatGPT or other forms of generative AI when writing code or solving other types of problems as part of your work on the homework assignments
- posting your work where others can view it (e.g., online), even after you complete the course
- receiving assistance from others or collaborating with others during an exam, or consulting materials except those that are explicitly allowed.

Incidents of academic misconduct will be reported to the Academic Conduct Committee (ACC). The ACC may suspend/expel students found guilty of misconduct. ***At a minimum, students who engage in misconduct will have their final grade reduced by one letter grade (e.g., from a B to a C).***

Other Policies

Laptops: Students taking CS courses are expected to have a laptop capable of running a currently supported version of Microsoft Windows, Mac OS X, or Linux. See this page for more info: https://www.bu.edu/cs/undergraduate/undergraduate-life/laptops

Late problem sets: Problem sets must be submitted by the date and time listed on the assignment (typically by 11:59 p.m.). There will be a 10% deduction for submissions up to 24 hours late. **We will not accept any homework that is more than 24 hours late.** Plan your time carefully, and don't wait until the last minute so you will have ample time to ask questions and obtain assistance from the course staff.

Pre-lecture preparation: To help you prepare for lecture, you will often be required to read or review some online materials. You may also be required to complete an online quiz to demonstrate that you have completed the preparation. Your work on these quizzes will not typically be graded for correctness, but it should demonstrate that you have adequately prepared for lecture. The pre-lecture tasks must be submitted by the specified date and time. **Late pre-lecture work will not be accepted**.

The *participation* portion of your grade will be based on your completion of the pre-lecture quizzes and in-lecture questions, and on your consistent participation in the lab sessions. You will receive full credit for participation if you earn at least 85% of the points for the pre-lecture/in-lecture questions and participate in at least 85% of the lab sessions. If you earn x% of the pre-lecture/in-lecture points or participate in x% of the lab sessions for a value of x that is less than 85, you will get x/85 of the possible points.

Absences: The above participation policy is designed to allow for occasional absences due to illness or other special circumstances – including ones stemming from isolation for Covid. We will be recording the lectures and making the recordings available to everyone in the class. If you need to miss a lecture for any reason, you should simply watch the recording for that lecture as soon as possible after it is posted. In addition, you should keep up with the pre-lecture tasks and the current assignments. **Please do not email your instructor for absences of this type.**

The final exam will replace your lowest problem-set grade if doing so helps your final grade. The final exam will also replace your midterm-exam grade if doing so helps your final grade. Regardless of whether any such replacements occur, the final exam itself will always count for at least 35% of the final grade.

The final grades are *not* curved. The performance of the class as a whole is taken into account in assigning letter grades, but this can only improve your grade, not harm it.

Extensions and makeup exams will only be given in *documented* cases of serious illness or other emergencies.

You cannot redo or complete extra work to improve your grade.

Incompletes will not be given except in extraordinary circumstances.

Schedule (tentative)

week	lecture dates	topics, exams, assignments, and special dates
0	9/6, 9/8	Course overview and introduction Database design and ER diagrams *No labs this week.*
1	9/11, 9/13, 9/15	The relational model Relational algebra and SQL
2	9/18, 9/20, 9/22	SQL (cont.) *9/18: last day to add a class* **Problem Set 1, part I due on 9/20**
3	9/25, 9/27, 9/29	Storage and indexing **Problem Set 1, part II due on 9/27**
4	10/2, 10/4, 10/6	Indexing (cont.) Implementing a logical-to-physical mapping Query processing
5	**10/10**, 10/11, 10/13	Transactions and schedules Concurrency control basics *No lecture on 10/9 (Indigenous Peoples' Day)* *Lecture on 10/10 (Mon. schedule)* *No labs this week* *10/10: last day to drop without a 'W'* **Problem Set 2, part I due on 10/11**
6	10/16, 10/18, 10/20	Concurrency control (cont.) **Problem Set 2, part II due on 10/18**
7	10/23, 10/25, 10/27	Concurrency control wrap-up Semi-structured data and XML databases **Problem Set 3, part I due on 10/25**
8	10/30, 11/1, 11/3	XML databases (cont.) Distributed databases and replication **Midterm exam on 11/1**
9	11/6, 11/8, 11/10	Distributed databases (cont.): map-reduce **Problem Set 3, part II due on 11/8**
10	11/13, 11/15, 11/17	NoSQL *11/13: last day to drop a class with a 'W'*
11	11/20	NoSQL (cont.) *No lecture on 11/22 or 11/24 (Thanksgiving)* **Problem Set 4 (all) due on 11/20 (Mon)**
12	11/27, 11/29, 12/1	Recovery and logging
13	12/4, 12/6, 12/8	Recovery and logging (cont.) Performance tuning **Problem Set 5 (all) due on 12/6**
14	12/11, 12/13	Wrap-up and review *12/14: Study period* **Final exam: Friday, 12/15, 12:00-2:00 p.m.**