CAS CS 460/660: Introduction to Database Systems  
Boston University  
Fall 2013  

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 (indexing structures, concurrency control, recovery, and query processing); management of semistructured and complex data; and distributed databases.

Prerequisites: CAS CS 112, or equivalent

Instructor  
David G. Sullivan, Ph.D. (dgs @ cs . bu . edu, removing the spaces)  
office hours: Mon., 3-5 p.m.; Wed., 2-4 p.m., and by appointment  
office: Psychology Building (PSY), room 228D  
64 Cummington Mall (behind Warren Towers)

Teaching Fellow  
Haohan Zhu (zhu @ cs . bu . edu, removing the spaces)  
office hours: TBA

Lectures: Tues/Thurs, 9:30-11:00 a.m., KCB 106

Course Website: http://www.cs.bu.edu/courses/cs460

Requirements  
1. Five problem sets, including a combination of written exercises and programming problems.
2. Midterm exam
3. Final exam
4. Attendance at and participation in the lectures
CS 660 students will complete additional work.
The exams are closed-book.

Policies  
Lateness policy: Assignments must be submitted by the date and time listed on the assignment. There will be a 10% deduction for homework that is up to 24 hours late, and a 20% deduction for homework that is 24-48 hours late. We will not accept any homework that is more than 48 hours late. Plan your time carefully, and don’t wait until the last minute to begin an assignment.
Determining the final grade:

- attendance/participation 10%
- assignments 40%
- midterm 17%
- final exam 33%

Extensions and makeup quizzes/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.

**Collaboration Policy**

The collaboration policy for this class is as follows.

- You are strongly encouraged to collaborate with one another in studying the textbook and the lecture material.

- As long as it satisfies the following conditions, collaboration on the homework assignments is encouraged and will not reduce your grade:
  1. Before discussing each homework problem with anyone else, you must give it an honest half-hour of serious thought.
  2. You may discuss ideas and approaches with other students in the class, but not share actual code or solutions to other types of problems. In other words, the work you submit must be entirely your own, which you must complete without looking at other people's work, and you must not permit others to copy your work. You must also acknowledge clearly in the appropriate portion of your solutions (e.g., in the comments of your code) people with whom you discussed ideas for that portion.
  3. You may get help from the teaching staff and from tutors in the lab for specific homework problems. Don't expect them to do it for you, however.
  4. If you get really stuck with a bug in a program (defined roughly as over an hour of frustration), you are allowed to get help from a friend as long as you acknowledge that help clearly in your solutions (e.g., in the comments of your code).
  5. You may not work with people outside this class (but come and talk to us if you have a tutor), seek online solutions, get someone else to do it for you, etc.

- You are not permitted to collaborate on exams.

The last point is particularly important: if you don't make an honest effort on the homework but always get ideas from others, your exam score will reflect it.

*Thanks to Prof. Leo Reyzin, who wrote the original versions of the sections describing the collaboration policy and violations to that policy. I have made only minor modifications.*

David G. Sullivan, Ph.D.
Violations of Collaboration Policy*

Violations of collaboration policy fall into two categories: ones that are acknowledged at the time they occur (for example, in clearly marked comments in your code) and ones that are unacknowledged.

Acknowledged violations (e.g., using someone else's code for a method you didn't know how to write yourself, and stating clearly in your code that this is not your own work) will result in an appropriate reduction in the grade, but will not be considered cheating. You should send an email to cs460-staff@cs.bu.edu about all such violations, or anything that you think may possibly be considered as such.

Unacknowledged violations of the collaboration policy—for example, not stating the names of your collaborators, or any other attempt to represent the work of another as your own—will result in an automatic failing grade and will be reported to the Academic Conduct Committee (ACC). The ACC often suspends or expels students deemed guilty of plagiarism or other forms of cheating. We will assume that you understand the CAS Academic Conduct Code, copies of which are available in CAS 105.

If you are uncertain as to whether a particular kind of interaction with someone else constitutes illegal collaboration or academic dishonesty, please ask Dr. Sullivan before taking any action that might violate the rules; if you can't reach him in time, then at the very least include a clear explanation of what happened in your homework to avoid being treated as a cheater. Citing your sources is usually the easiest way out of trouble.

Textbooks

- Computer Science 460/660 coursepack. This will be available for purchase at Fedex Office at the corner of Comm Ave. and Cummington. More information will be given during the first lecture.
- Optional readings will be also given from the following book: *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, but you may find it useful to purchase it. Students in CS 660 will be required to do additional readings.
## Schedule (tentative)

<table>
<thead>
<tr>
<th>week</th>
<th>lecture dates</th>
<th>topics, exams, and special dates</th>
<th>optional readings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9/3, 9/5</td>
<td>introduction; database design and ER models</td>
<td>4.1-4</td>
</tr>
<tr>
<td>2</td>
<td>9/10, 9/12</td>
<td>the relational model; relational algebra and SQL</td>
<td>2.1-5, 6.1-5</td>
</tr>
</tbody>
</table>
| 3    | 9/17, 9/19    | more SQL, **Problem Set 1, part I due on 9/17**  
9/16: last day to add a class | see above |
| 4    | 9/24, 9/26    | storage and indexing, **Problem Set 1, part II due on 9/24** | 13.5, 13.7, 14.1-3 |
| 5    | 10/1, 10/3    | the logical-to-physical mapping; key-value stores; query processing | 15.1-3, 15.5-6 |
| 6    | 10/8, 10/10   | transactions; concurrency control, **Problem Set 2, part I due on 10/8**  
10/7: last day to drop without a 'W', and to change from credit to audit status | 6.6, 19.1-2, 18.1-8 |
| 7    | 10/17         | concurrency control (cont.), **Problem Set 2, part II due on 10/15**  
No lecture on Tuesday, 10/15 (Mon. schedule) | see above |
| 8    | 10/22, 10/24  | recovery and logging, **Midterm exam on 10/24** | 17.1-4 |
| 9    | 10/29, 10/31  | complex data; objects in database systems, **Problem Set 3, part I due on 10/29** | 4.9-10, 10.3-5 |
| 10   | 11/5, 11/7    | semistructured data; XML, **Problem Set 3, part II due on 11/5**  
11/8: last day to drop a class with a 'W' | 11.1-4, 12.1-2 |
| 11   | 11/12, 11/14  | distributed databases and replication, **Problem Set 4, part I due on 11/12** | 20.3-4, 20.5-6 |
| 12   | 11/19, 11/21  | noSQL databases, **Problem Set 4, part II due on 11/19** | |
| 13   | 11/26         | graph databases, **No lecture on 11/28 (Thanksgiving recess)** | |
| 14   | 12/3, 12/5    | performance tuning; other topics, **Problem Set 5, part I due 12/3** | |
| 15   | 12/10         | wrap-up and review, **Problem Set 5, part II due on 12/10**  
**No lecture on 12/12 (Study period)** | |
|      |               | **Final exam: Monday, 12/16, 9-11 a.m.** |