Data Structures
Spring 2021
Course Information

Last updated: 2021-Feb-09

Be sure to read **all** of this document!

1 The topics

The way in which a program represents and organizes its data deeply affects how it then moves through and uses that data. Is the data numbers? Strings? Objects? Are the data ordered? Are values inserted or removed often? Do values need to be found quickly? Are there hundreds of values, or billions? The choice of how to structure data is intimately related to what a body of data represent, and how it is to be used; more, it often represents something important about how values relate to one another, and how those solving the problem understand it.

We assume, thus far, that the primary ways in which you have structured data are *linear*: arrays/vectors, lists, queues, stacks. While each of these provide different ways of accessing data, it is all stored in a "straight line": one value after another. In this course, we discover that the organizational options are much richer, providing useful ways to organize and manipulate data in ways that change how problems are solved.

Here is a list of the topics on which we will spend our time, not necessarily in this order:

- Linear structures (review)
- Asymptotic analysis (Big-O)
- Binary trees (unbalanced and balanced)
- Priority queues and heaps
- Skip lists and probabilistic structures
- Hash tables
- Graphs

This course will mix abstract concepts, concrete implementation, and some mathematical analysis. You will implement the structures, use them in solving problems, and analytically compare them for different purposes. At the end of this course, you will have a rich collection of possible organizations, as well as some practice in knowing when to apply each.

2 Lectures, labs, and help

Lectures and labs: This class will meet on **Tuesdays and Thursdays**, from **10:10 am to 11:30 pm**. Our class meetings will be entirely on Zoom. You are expected to be present and prepared for **all class meetings**.

Office hours and meetings: If you seek assistance, reinforcement, review, or other opportunities to discuss the course material or assignments, you should see me. There is a link on the course web page for scheduling times for these meetings.

Communications: You should certainly feel free to send **email** (<sfkaplan@amherst.edu>) with your questions or thoughts. I also encourage the use of *Slack*. We will have a channel for the class, in which you will have the ability to post questions either anonymously or, if you're willing, onymously.

3 Texts and materials

There is no textbook for the course. However, the website `opendatastructures.org` is a good reference for the topics will we address.

4 Assignments and exams

There will be a mixture of assignments for this course. Most of them will be programming assignments (in Java), in which you will implement key aspects of these data structures, code that uses the structures, and code that tests your implementations. There will also be some handwritten homeworks for our work on analysis and probability.

There will also be **two mid-term exams**, as well as a final exam or a final project (TBD!).

5 Grading

At the end of the course, I will assess your mastery of the course material. To do so, I will consider your assignments (written and coding), and your performance on the mid-term exams and on the final exam/project. All of these elements of the course are opportunities for you to demonstrate your understanding of the material.

This unspecific description is intentional. Rather that contrive a formula, this approach allows me to account for variations in your performance: a single bad day on one mid-term exam should be balanced by consistent performance in other aspects of the course. How to balance these is always a matter of judgment; any pre-determined formula will fail for some case to reflect the mastery of the material.

6 Academic honesty

Discussing the material and assignments with your peers is a valuable part of the academic experience. I encourage you to work with others, study together, and reinforce each other's understanding of the material.

However, **the work that you submit must be wholly your own.** You may not look at someone else's code or solutions, and you may not show your code/solutions to someone else. You can discuss the assignment, discuss the questions and assignments, work together on understanding any code distributed as part of the assignment, and ask one another questions about the topics. However, **seeing or showing code** is not allowed.

If you are unsure of where the boundary lies—where "working together" lapses into "academic dishonesty"—then play it safe and ask me for clarification.