Systems Development for Computational Science

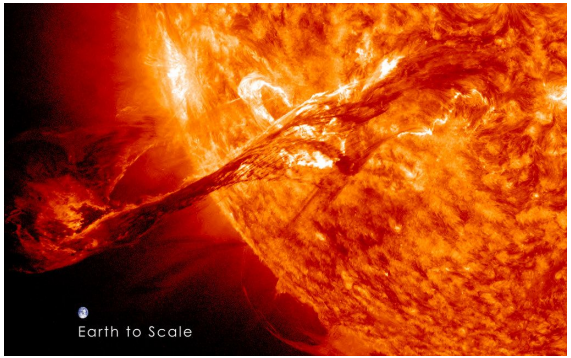
CS107 / AC207 / E-207

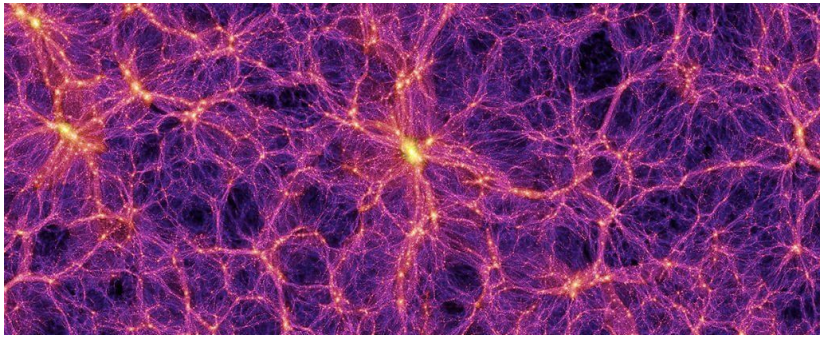
Dr. David Sondak

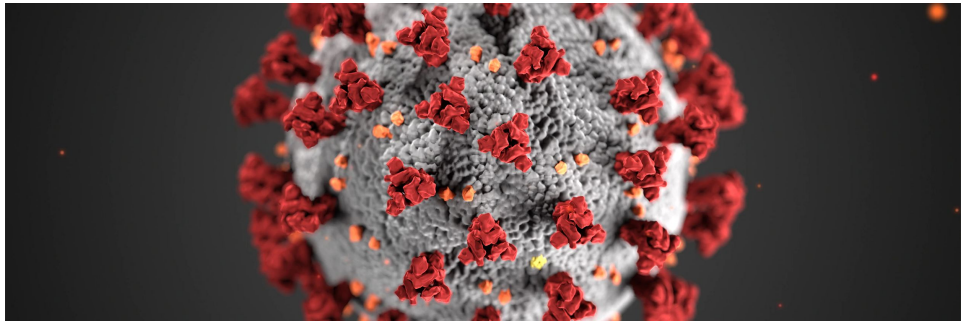
Part 1: Why?

Why should I take this class?

Motivation







The Pillars of Science

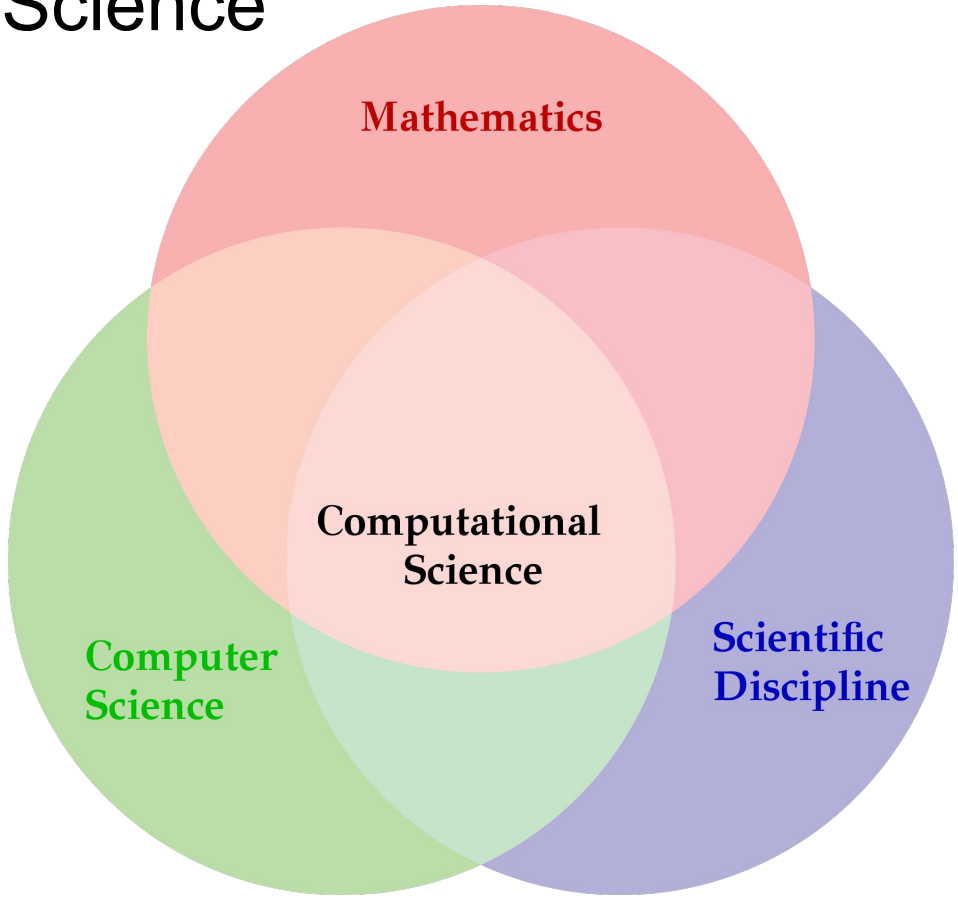
Science

Theory

Experiment

Computation

Computational Science

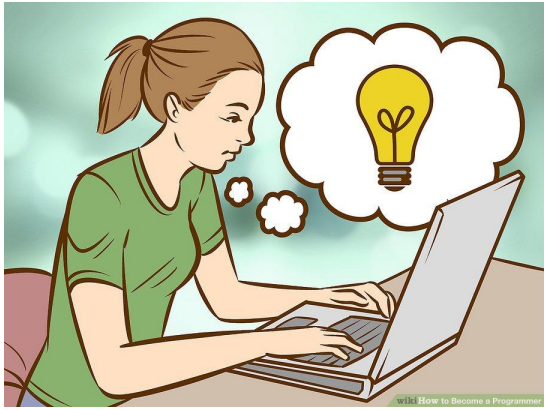


Why take this class?

- Scientific software is complex
- Your code needs to be:
  - Reusable
  - Portable
  - Robust
- Must go beyond “scripting”

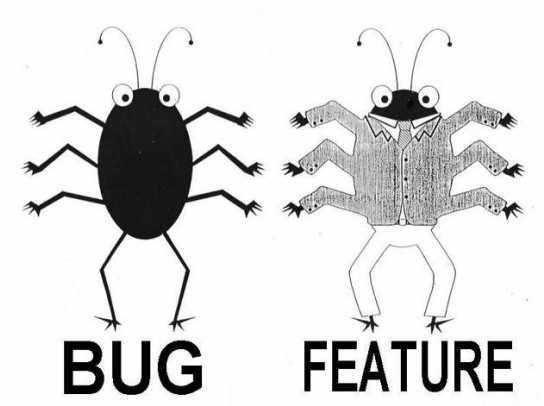
Course Objectives

To give students who may not have a traditional computer science background the knowledge and tools to develop and maintain effective software for computational and data science applications



Who should take this class?

- Any kind of scientist is welcome to take this class!
- This course is computer science for people who aren’t computer scientists
  - Data scientists
  - Biologists
  - Chemists
  - Engineers
  - Physicists
  - Mathematicians
  - Engineers
  - ⋮

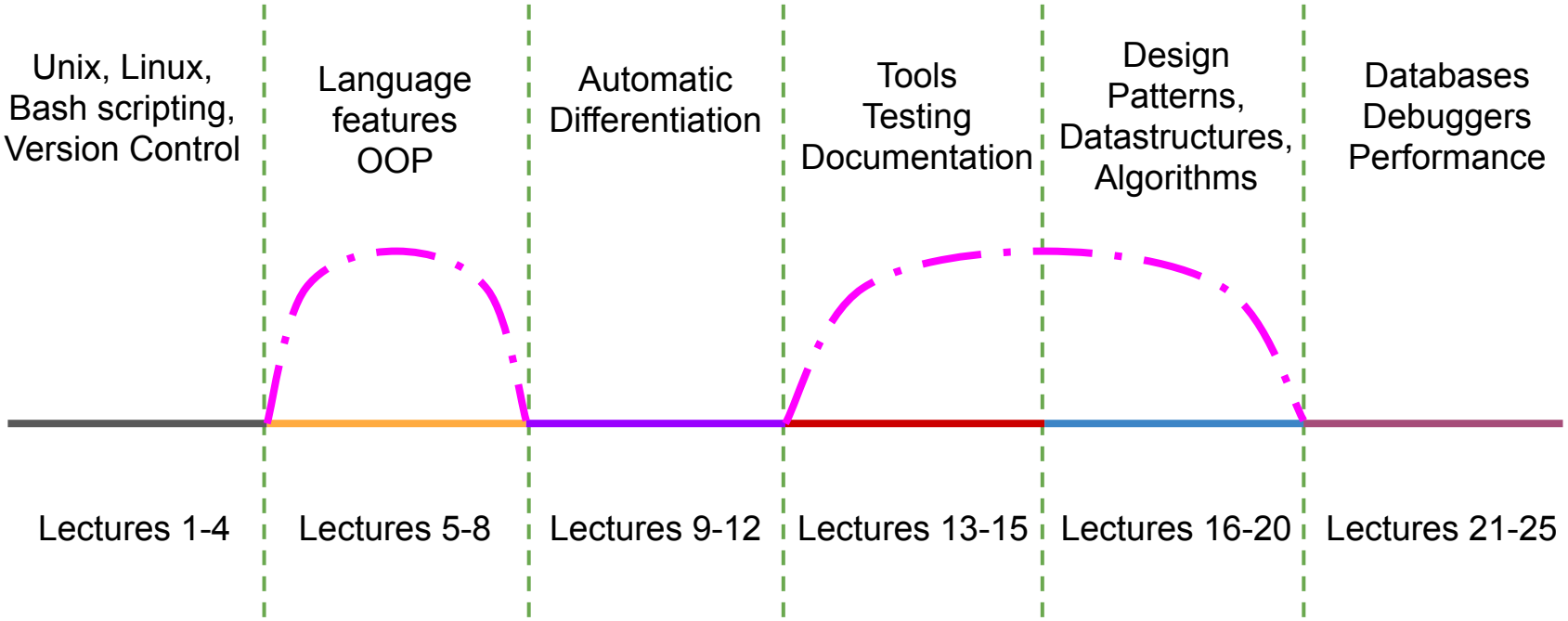


- Computer scientists who want to develop scientific software
- Students who need to know effective and modern software practices

Part 2: What

What’s in this class?

Topics and Timeline



Course Structure

Monday	Tuesday	Wednesday	Thursday	Friday
PP Sections 7:30 AM 12:00 PM 4:30 PM	Lecture 12:00 PM		Lecture 12:00 PM PP Sections 6:00 PM	PP Sections 9:00 AM 12:00 PM

- Lectures
  - Optional (but strongly encouraged)
  - Tues / Thurs. from 12:00 PM - 1:15 PM
  - Recorded and posted to Canvas --- watch before pair-programming section!
- Pair-Programming (PP) sections (times subject to change)
  - Must attend at least one per week

Pair-Programming Sections

- You must attend one per week
- You can choose which one
- You can switch it up week-to-week depending on your schedule
- Led by TFs (but I will come by from time-to-time)

- Group exercises
  - 1-2 exercises per section
  - Work with your project groups if possible
- Use Deepnote for pair-programming
  - See https://harvard-iacs.github.io/2020-CS107/pages/coursework.html for specifics
- Exercises are graded for completion
  - Turn them in on time and do your best
  - These are fairly easy points

Optional C++ Modules

- Some topics will offered in separate C++ lectures
- This is *entirely* optional
  - Students with prior experience many enjoy these sessions
  - A lot of scientific software is in C++ (e.g. Tensorflow and Pytorch as well as big scientific codes)
- If you attend a C++ lecture, then you do not need to attend the regular lecture
- You have the option of completing homework assignments in C++ or Python
  - You are not required to do all assignments for the semester in a single language
  - However, any individual assignment must use only one language
    - e.g. Do assignment 2 in Python and assignment 3 in C++, but assignment 2 must be done completely in Python and assignment 3 must be done completely in C++
  - Assignments in C++ will get +5 points added to the assignment grade

Grading Breakdown

- Homework: 20%
  - 7 assignments
- Participation: 20%
  - Piazza
  - Pair-programming
- Quizzes: 10%
  - 7 quizzes???
- Project: 50%
  - Three milestones
  - Each milestone is split into 3 parts

Project

- Develop a software library for automatic differentiation
- Work in groups of 3 - 4 (assigned by teaching staff)
- What is automatic differentiation?
  - A way to calculate and evaluate derivatives of functions and computer programs
  - Compute derivatives to *machine precision*!
- Why automatic differentiation?
  - Pervasive throughout science and gaining steam
    - Neural nets and backpropagation
    - Hamiltonian Monte Carlo methods
    - Jacobian (and Jacobian-free) calculations
  - Encapsulates many ideas in software design
    - Object-oriented programming
    - Datastructures

AD Teaser

Suppose we want to differentiate the function

$$y = \exp\left(-\sqrt{x + \cos^2(x)}\right) \sin\left(x \ln\left(1 + x^2\right)\right)$$

The symbolic derivative is

$$y' = \exp\left(-\sqrt{x + \cos^2(x)}\right) \cos\left(x \ln\left(1 + x^2\right)\right) \left(\frac{2x^2}{1 + x^2} + \ln\left(1 + x^2\right)\right)$$
$$- \exp\left(-\sqrt{x + \cos^2(x)}\right) \left(\frac{1 - 2\cos(x)\sin(x)}{2\sqrt{x + \cos^2(x)}}\right) \sin\left(x \ln\left(1 + x^2\right)\right)$$

And that's only the first derivative!!!

Q / A

- Much more detailed information can be found on the main course website:

https://harvard-iacs.github.io/2020-CS107/

  - It is still being updated, so check back often!

Shopping Week Office Hours: https://harvard.zoom.us/my/dsondak

- Tuesday, August 18th, 3:00 PM - 4:00 PM (Boston time)
- Friday, August 21st, 9:00 AM - 10:00 AM (Boston time)