

Data Structures and Algorithms

CSCI 2270

Rhonda Hoenigman
January 12, 2015
Lecture 1

Today...

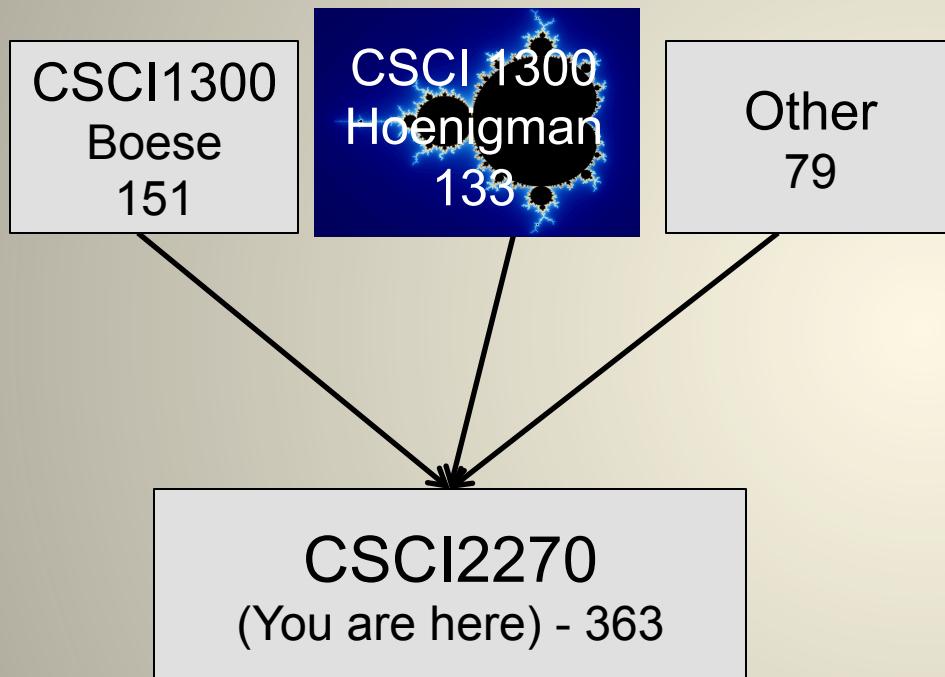
- Logistical details
 - Grading
 - Textbooks
 - Recitations
 - Virtual Image
 - Computer Science Moodle
- First...

What is Data Structures and Algorithms?

Why should I suffer through it?

Backup to Intro Programming...

Most of you coming from CSCI1300



Reasons for enrolling in CSCI1300

- Intentional interest
- Required for another major
- Your mom made you do it
- You got lost on your way to English 100
- Accidental

If you're here, presumably you enjoy programming and want to know more.

Data Structures and Algorithms

Catalog Definition:

- Studies data abstractions (e.g., stacks, queues, lists, trees) and their representation techniques (e.g., linking, arrays). Introduces concepts used in algorithm design and analysis including criteria for selecting data structures to fit their applications.
- Wow, really? That's a lot of words.

An interesting political problem...

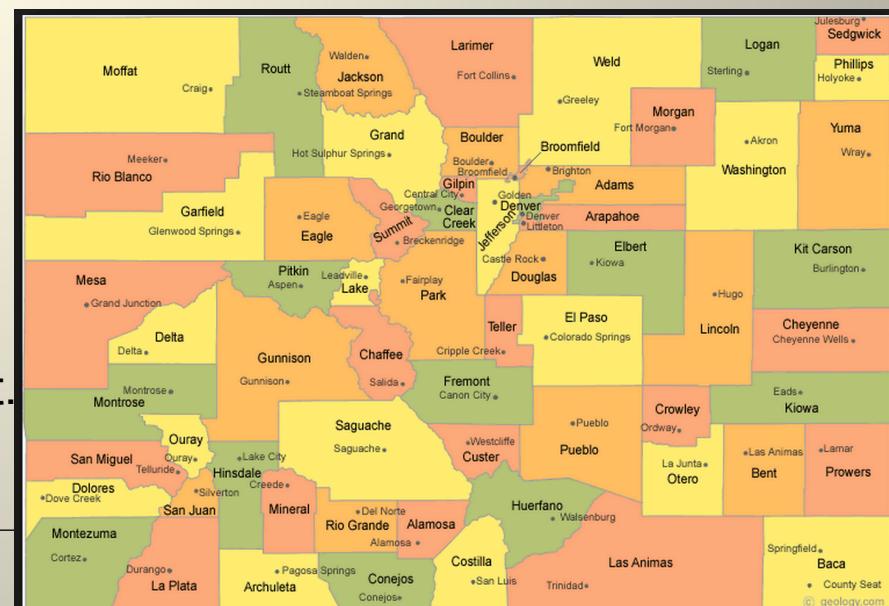
Gerrymandering – the process of manipulating district boundaries to gain political advantage.



1812 political cartoon.
Boundary of a Mass Congressional district.
Candidates of Elbridge Gerry
Looks like a salamander

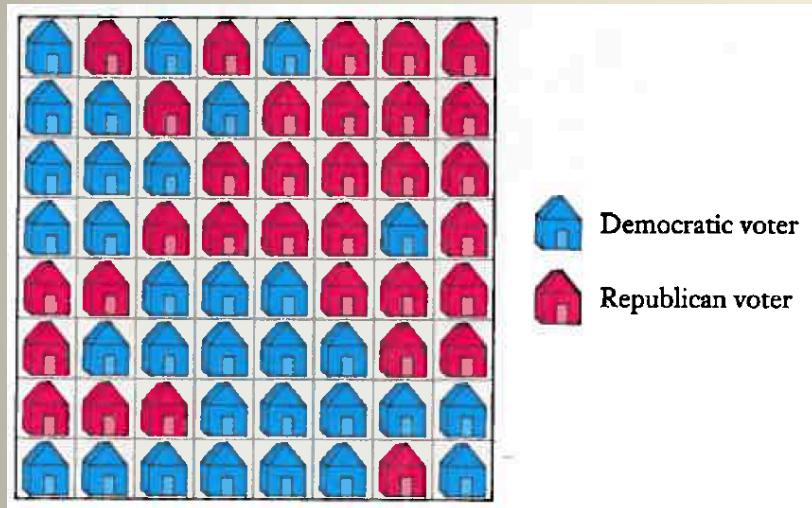
What is fair districting?

63 Colorado counties, 7 congressional
representatives



An interesting (computational) problem...

Write an algorithm to determine fair boundaries



Fictional neighborhood:

- 64 houses, divide into 8 districts of 8 contiguous houses

District winner is majority, Republican or Democrat

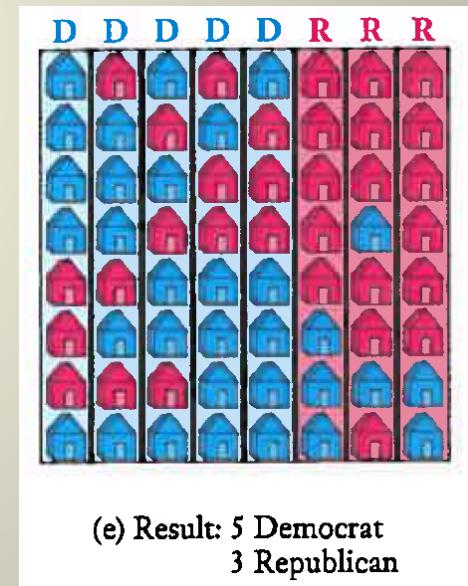
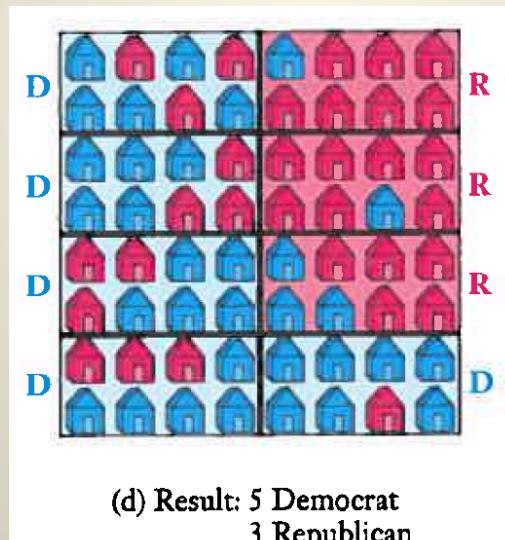
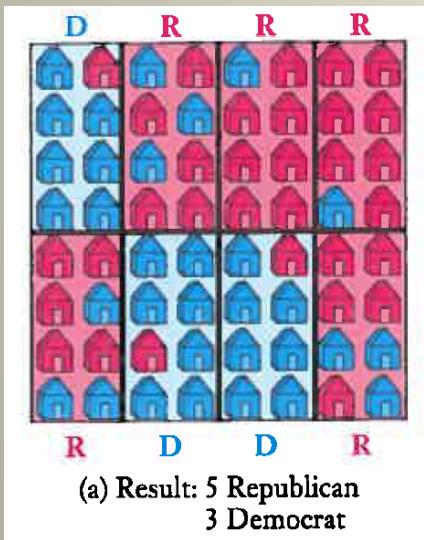


University of Colorado
Boulder

Computational, and political, challenges

3 examples of where political boundaries could be drawn

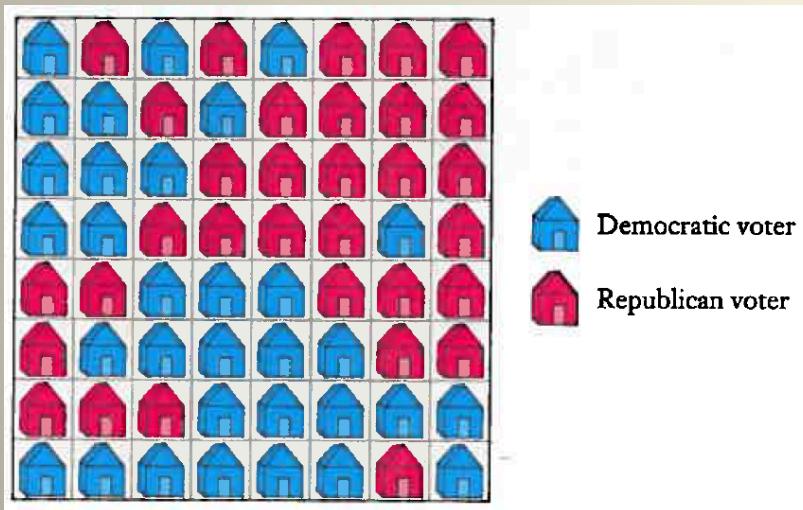
- Boundaries determine election outcome!
- Lots and lots of possible districts.



Draw boundaries computationally

Computationally generate all possible districts

- Humans can't do this manually
- Challenging computationally too
- Fairness could require this
- (We can't define fairness here.)



CSCI2270 Data Structures and Algorithms

Learn to generate and evaluate solutions to this type of problem

Algorithms

- Steps to solve problem
 - Generate possible districts
- Cost
 - How many operations? Time? Is one algorithm faster than another?

Data Structures

- Store partial solutions efficiently

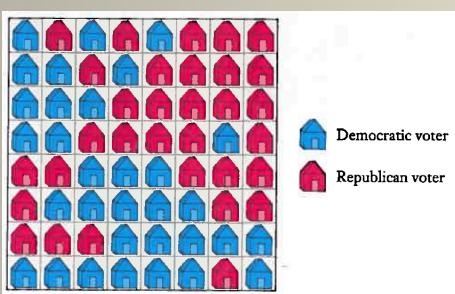


University of Colorado
Boulder

Assigning districts algorithmically

Assume partial district generated:

- 7 cells assigned. What are options for 8th cell?

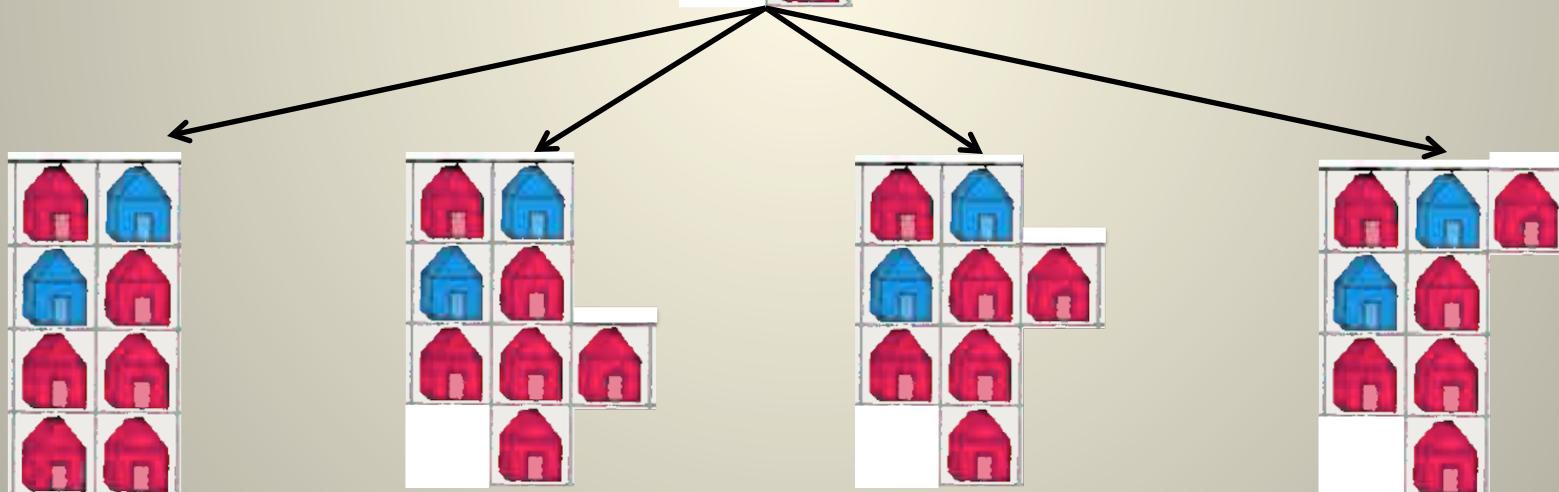


Democratic voter
Republican voter



Approximately 19,000 first district selections (8 cells)

- Algorithm generates all 19K
- How to store each selection?

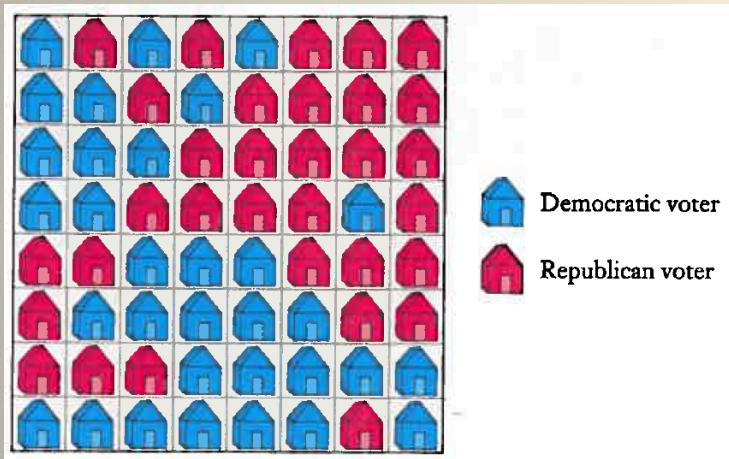


University of Colorado
Boulder

Storing districts, you need a data structure

Approximately 19,000 first groupings for district

- Ways to choose first 8 contiguous cells
- Subsequent selections depend on previous selections



Data structure stores each of those 19K first selections and all subsequent selections as possible solutions.

- e.g. list of lists, dictionary, array of objects
- Search data structure for what's been selected

Cost matters

AI Algorithm – Adversarial search

- Approximately 3 weeks to calculate solution in Artificial Intelligence class

Data Structures and Algorithms

Catalog Definition:

- Studies data abstractions (e.g., stacks, queues, lists, trees) and their representation techniques (e.g., linking, arrays). Introduces concepts used in algorithm design and analysis including criteria for selecting data structures to fit their applications.

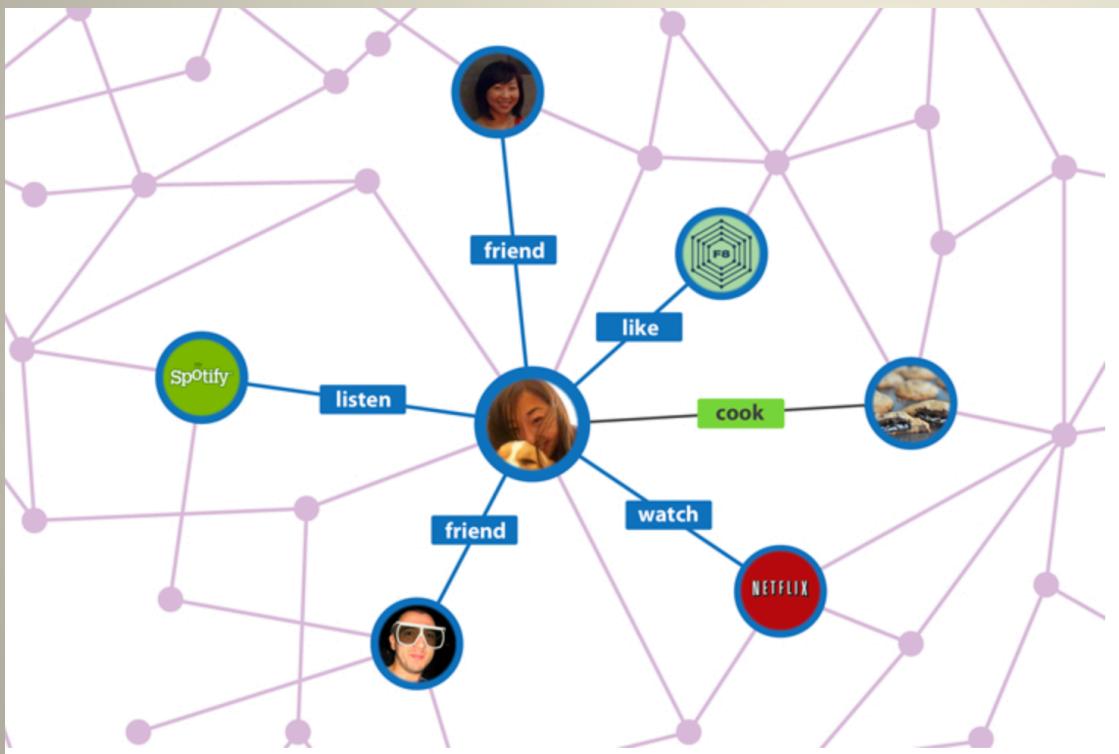
For districting problem

- Algorithm shown was exhaustive search
 - Someone else needs to determine fairness
- Districts organized using stack, queue, list, tree
- Analyze cost of these options, select most efficient



Different structures for different problems

- Complex variables with connections and order



Example: Network (graph)

- Objects with connecting edges
- Facebook
- Internet traffic routing
- Airplane routing

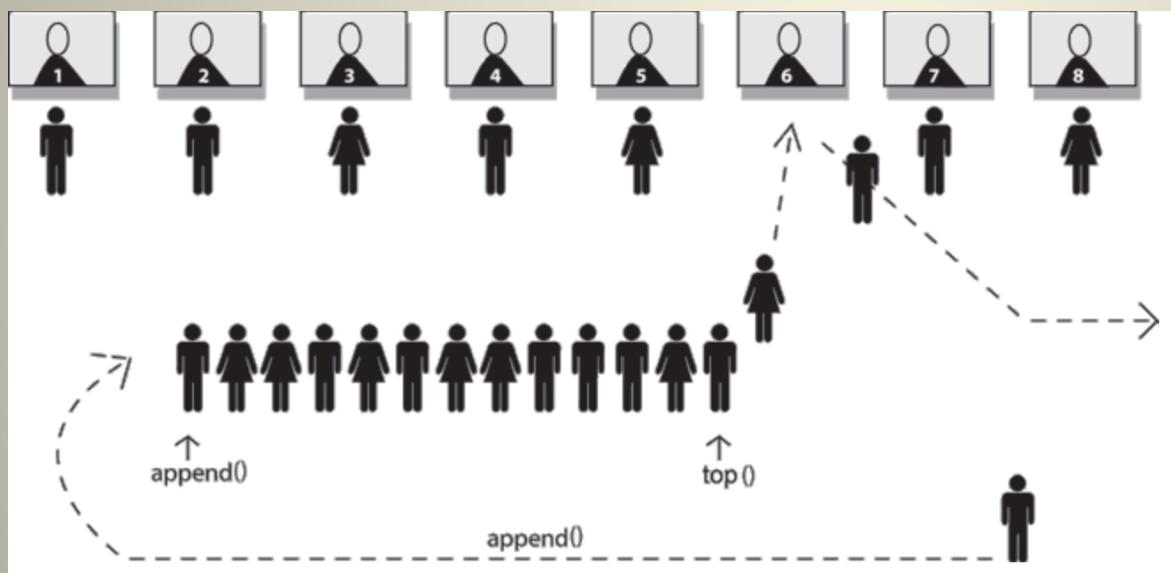
Study how to:

- Search
- Sort
- Remove
- Traverse



Another data structure – queue

- Other data structures:
 - Stacks, queues, linked lists, trees
 - Searching and sorting
 - Algorithmic complexity

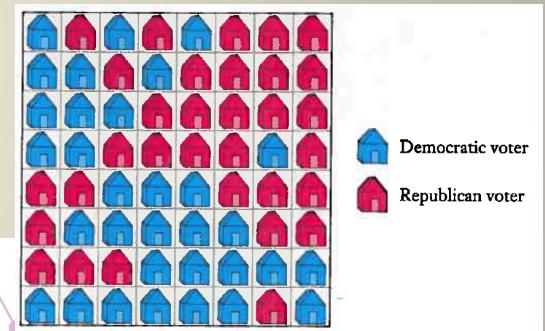
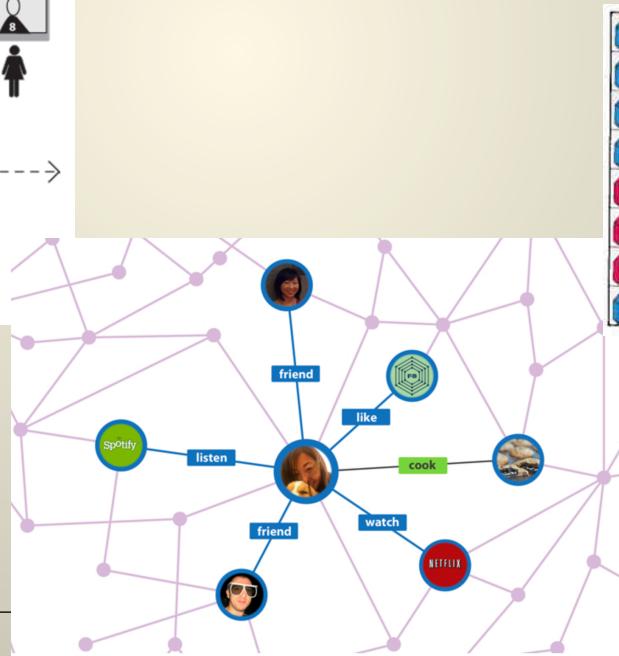
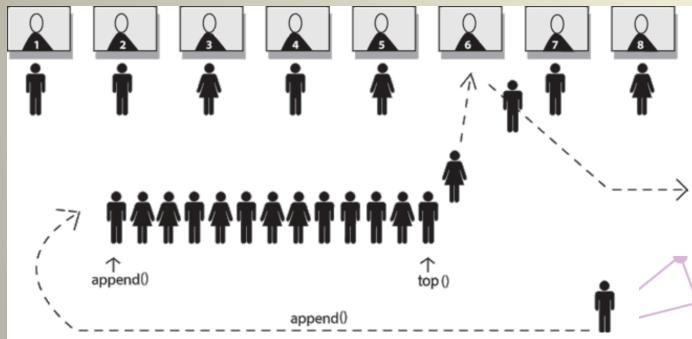


Example: Queue

- Line at the store
- Jobs on a processor

Algorithms for searching, sorting

- Common tasks to search, sort data
 - Algorithms – simple and slow, complicated and fast
 - How long does it take search/sort a list
 - Number of operations
 - How does an algorithm grow with data size (complexity)



University of Colorado
Boulder

Administrative details

Course syllabus on Moodle

<http://moodle.cs.colorado.edu/>

CSCI 2270 – Hoenigman – Data Structures

Enrollment key: csci2270

Login with your identikey

Class times:

MWF, 3-3:50pm. CHEM140

Recitation:

Weekly, 1.25-hour meetings. ECCS112C

Recitation activity. Ask questions about assignments and get extra help.

Administrative details

TAs:

Madhumitha Soundararajan (Madhumitha.soundararajan@colorado.edu)
Kannan Subramanian (Kannan.subramanian@colorado.edu)
Nilesh Jagnik (Nilesh.jagnik@colorado.edu)
Brett Israelsen (Brett.israelsen@colorado.edu)
Matthew Bubernak (Matthew.bubernak@colorado.edu)
Santhanaraman Ganesh (Santhanaraman.ganesh@colorado.edu)

LAs:

Multi-course LAs only. Full schedule on Moodle
(A few of them have been LAs for this class before.)

Office hours:

Mine: M-W-F, 1:30-2:30pm. ECOT 738

TA hours: Full schedule on Moodle

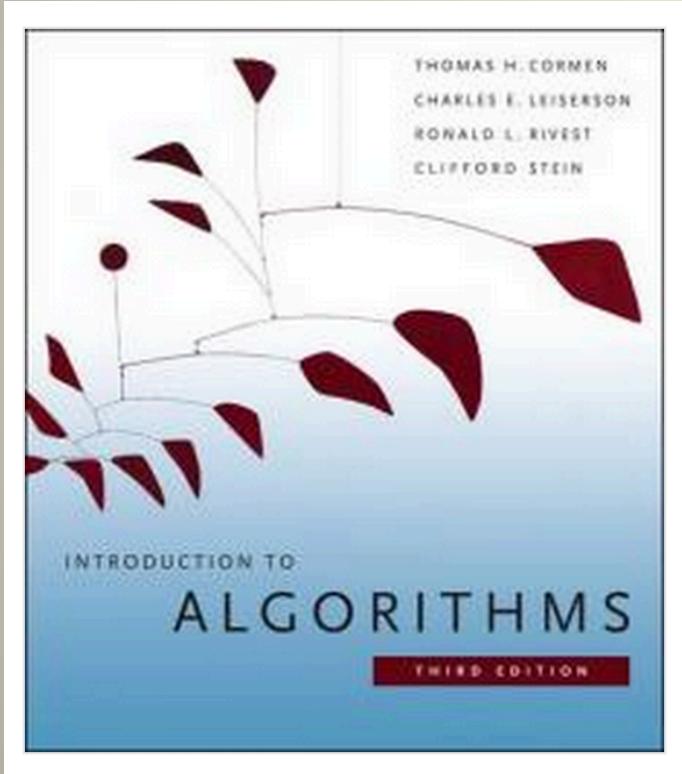
Lecture format

- Bring your laptop to class everyday
- Lecture notes posted on Moodle before class
 - Review them before coming to class
- Class is theory and implementation
 - Pseudo-code for algorithms
 - Implement algorithms in lecture
 - Follow conventions from your textbook



Textbook

- Cormen, Leiserson, Rivest, and Stein: Introduction to Algorithms, third edition
- Various online C++ resources



Classic computer science reference book.

Only has pseudocode

Part of this class will be learning to read this book

“We’re using CLRS” at your next party with computer scientists.



University of Colorado
Boulder

Interview grading

It's something we do in computer science

- Submit assignment electronically
- Make appointment for 15-minute meeting with your TA to discuss your assignment
- Grade depends on technical correctness and explanation
- Why: computer science code everywhere, easy to download a solution. Also provides time to discuss your work and help you.
- We will have some free-form assignments
 - Given some set of requirements, such as use five loops, write a program of your own design that meets the requirements.
 - Interview grading great for these.



Grading policy

Recitation 15%

- Weekly, generally due by 3pm on Friday

Homework 25%

- Weekly assignments
- Due by 3pm on Friday

Project 10%

- End of semester

Exams (midterm and a final, equally weighted) 50%

- Must get a 65% average on your exams to get a C- in the class
- C- required to take next class in computer science sequence
- 90, 45, 100 = 78% OK
- 70, 45, 70 = 61% Not Ok



Virtual machine development environment

Lots of different operating systems

- They all work differently.
- Debugging code on every OS is a nightmare.

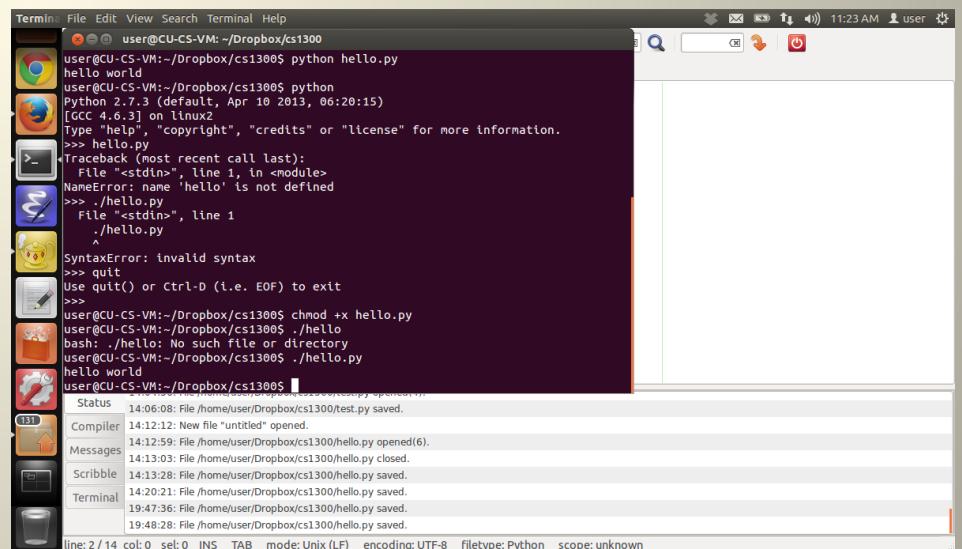
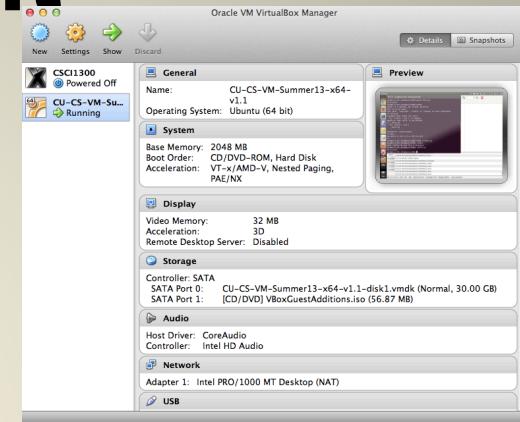
The virtual machine is another “pretend” computer running on your computer, regardless of your OS.

Provides consistency.

Uses a Virtual Machine, and An image (OS, programs)

Start here:

<http://foundation.cs.colorado.edu/sde/>



University of Colorado
Boulder

Due this week

Recitation:

Install the virtual machine on your laptop. New, faster version for Spring.

<http://foundation.cs.colorado.edu/sde/>

Create an account on the Moodle

<http://moodle.cs.colorado.edu/>

Enrollment key: csci2270

Submit Moodle exercise –New Moodle quiz type

Meet other students in your class to start a study group

C++ review exercises outside of class

Next time: C++ and cost of algorithms