CS 133 - Introduction to Computational and Data Science

Instructor: Renzhi Cao
Computer Science Department
Pacific Lutheran University
Spring 2017

About me

Renzhi Cao

Prof. Cao ? Cow

- Data Science
- Machine learning
- Bioinformatics

- Office: MCLT 248
- Office hours: In class website (cs.plu.edu/133)
- Office Phone: 535-7409
- Email: caora@plu.edu

Pictures from: https://www.google.com/search?

About guest speaker

Assistant Professor : Kyoungnam Catherine Ha

School of Business

- Marketing Finance Interface
- Marketing Analytics
- Customer Relationship Management

About you

- Names
- Where are you from?
- Your major?
- Hobbies? Movie? Song? Sports? Book? TV show? What you did over your break? …

What have I gotten myself into?

What is this course?

Mathematics?

Statistics?

Computer Science?

Programming?

Biology?

Economics?

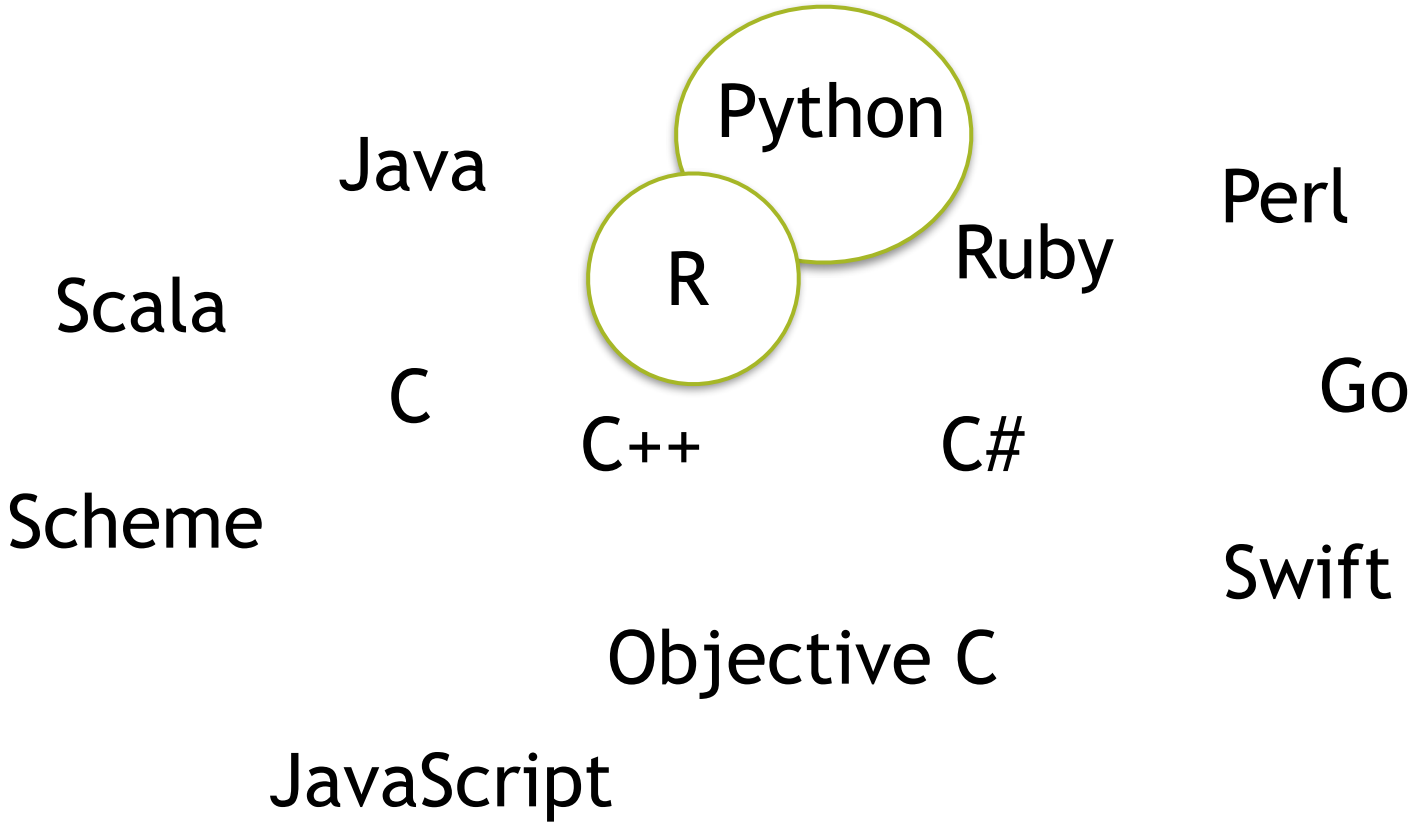
What is this course?

It's a little bit of all of those….

What is this course?

- Computer programming
  - Write software using Python and R
- Understand Computer Science as a discipline.

Programming languages



DO I HAVE WHAT IT TAKES TO SUCCEED AT PROGRAMMING?

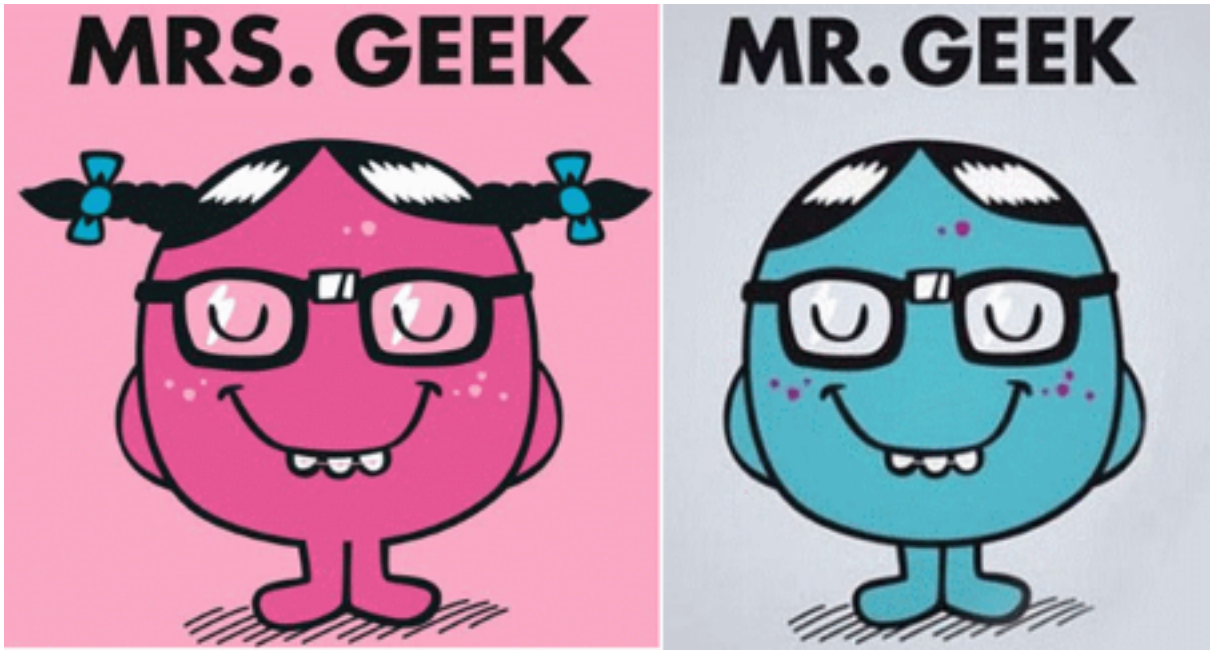


Image source: http://www.saveursdunet.com/wp-content/uploads/2012/12/Mr-and-Mrs-geek.png

WITH THE RIGHT ATTITUDE

AND A BIT OF HARD WORK

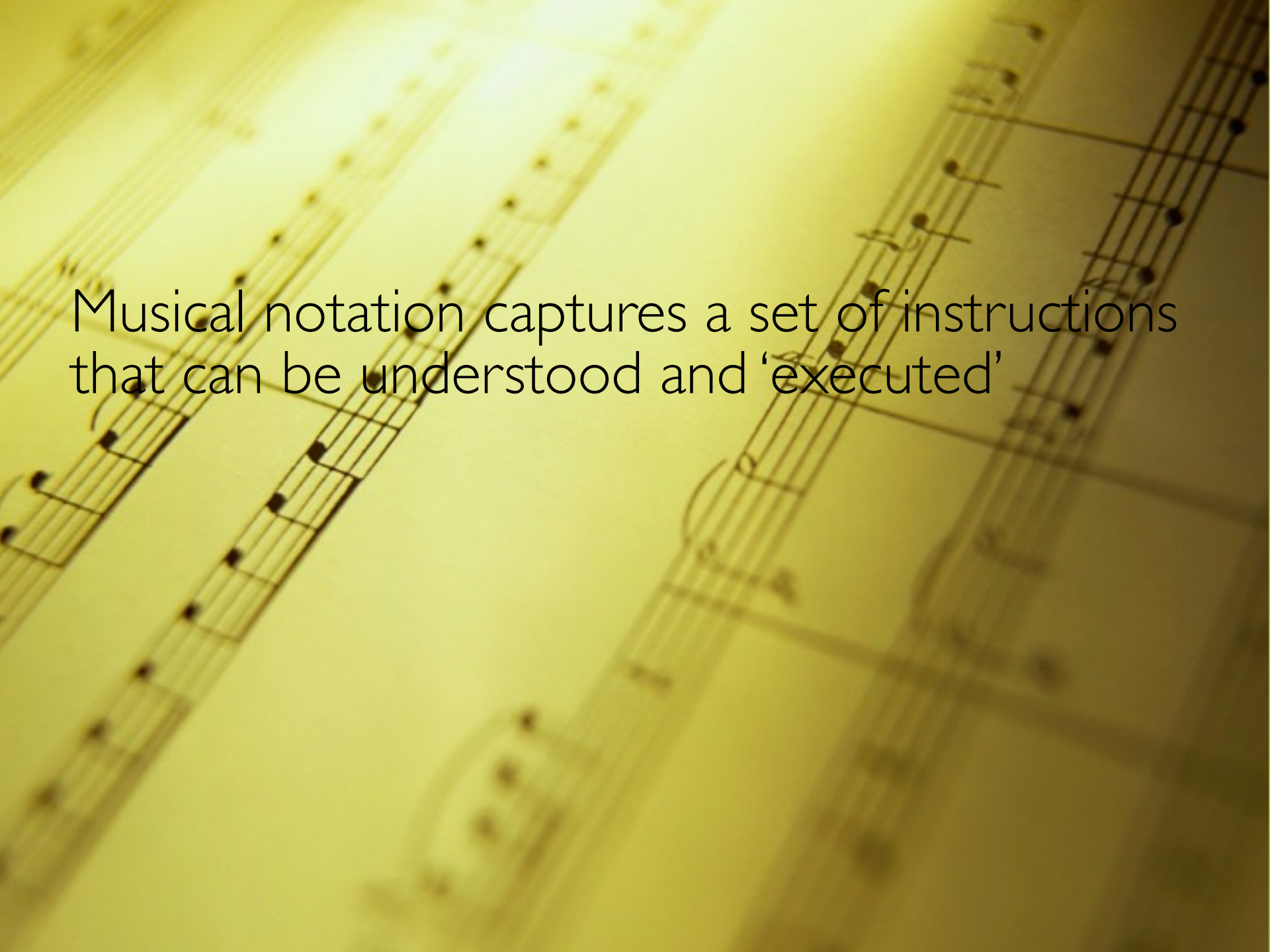
YOU CAN LEARN PROGRAM

… AND EVEN ENJOY IT!

```
# For loop on a list
>>> numbers = [2, 4, 6, 8]
>>> product = 1
>>> for number in numbers:
...     product = product * number
...
>>> print('The product is:', product)
The product is: 384
```

```
# Python 3: Fibonacci series up to n
>>> def fib(n):
>>>     a, b = 0, 1
>>>     while a < n:
>>>         print(a, end=' ')
>>>         a, b = b, a+b
>>>     print()
>>> fib(1000)
0 1 1 2 3 5 8 13 21 34 55 89 144 233 377 610 987
```

PROGRAMMING HAS A LOT IN COMMON WITH THINGS PEOPLE DO EVERY DAY


Musical notation captures a set of instructions
that can be understood and ‘executed’

Knitting languages describe different types of stitches, how many rows to knit, etc.

PATTERN:

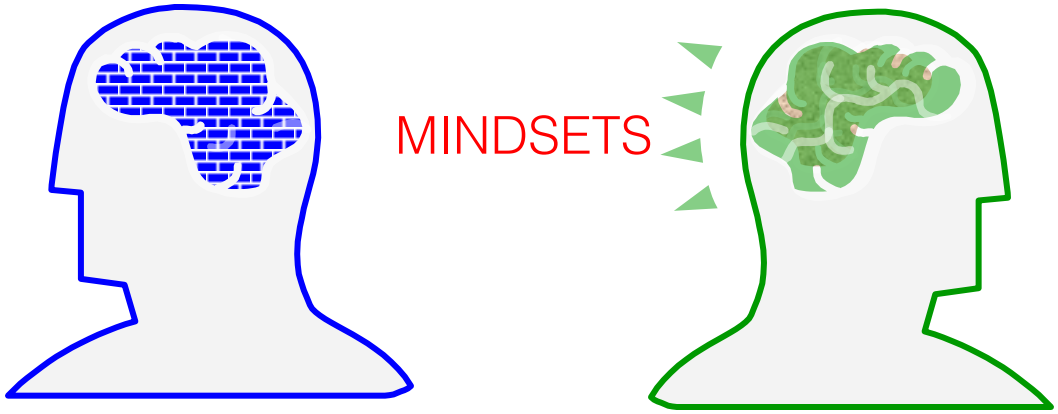
1. **Knit one row.**
2. **Purl one row.**
3. **Pattern Row: K1, K2 together, K6, WF K1, WF K6, K2 together twice, K6, WF K1, WF K6, K2 together, K1**
4. **Keep repeating these rows until you reach your desired length.**


Cooking also has special instructions and notations


Similar to learning to drive, programming
requires learning many new things at once

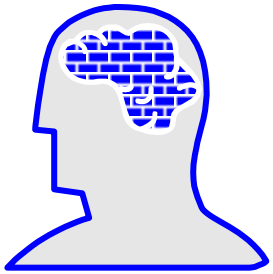
Much like reading maps, we often think of our programs abstractly and from different perspectives

One big difference with computers though, is that you must be extremely precise



Based on the work of
Stanford Psychologist Carol Dweck

intelligence is static

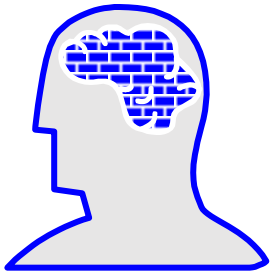


Fixed Mindset

intelligence can be developed

Growth Mindset

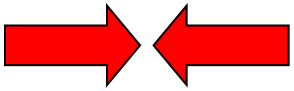
intelligence is static



Fixed Mindset

…avoid challenges

CHALLENGES

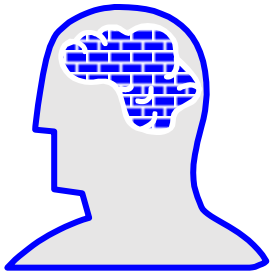


Growth Mindset

intelligence can be developed

…embrace challenges

intelligence is static



Fixed Mindset

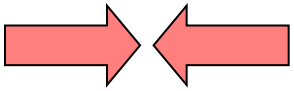
intelligence can be developed

Growth Mindset

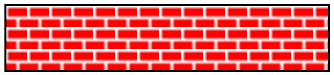
…avoid challenges

…give up easily

CHALLENGES



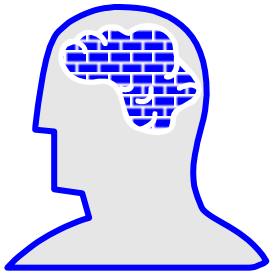
OBSTACLES



…embrace challenges

…persist in the face of setbacks

intelligence is static



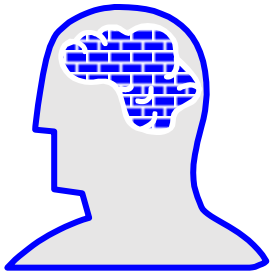
Fixed Mindset

intelligence can be developed

Growth Mindset

Fixed Mindset		Growth Mindset
…avoid challenges	**CHALLENGES**	…embrace challenges
…give up easily	**OBSTACLES**	…persist in the face of setbacks
…see effort as fruitless or worse	**EFFORT**	…see effort as the path to mastery

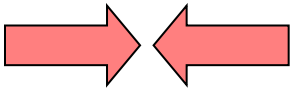
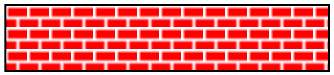
intelligence is static



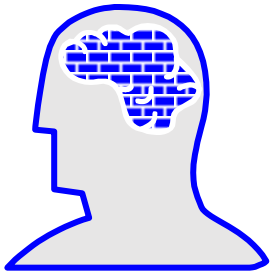
Fixed Mindset

intelligence can be developed

Growth Mindset

Fixed Mindset		Growth Mindset
…avoid challenges	**CHALLENGES** 	…embrace challenges
…give up easily	**OBSTACLES** 	…persist in the face of setbacks
…see effort as fruitless or worse	**EFFORT** 	…see effort as the path to mastery
…ignore useful negative feedback	**CRITICISM** 	…learn from criticism

intelligence is static



Fixed Mindset

intelligence can be developed

Growth Mindset

Fixed Mindset		Growth Mindset
…avoid challenges	**CHALLENGES**	…embrace challenges
…give up easily	**OBSTACLES**	…persist in the face of setbacks
…see effort as fruitless or worse	**EFFORT**	…see effort as the path to mastery
…ignore useful negative feedback	**CRITICISM**	…learn from criticism
...feel threatened by the success of others	**SUCCESS OF OTHERS**	…find lessons and inspiration in the success of others

Based on a graphic by Nigel Holmes available from :http://www.pvusd.net/Departments/GATE/dweck/

What mindset do you have?

Syllabus

Attendance

Attendance

- *Expected to attend every class*
- *YOU are responsible for missed materials*

Classroom Conduct

- Come to class on time
- Turn off electronic devices
- Refrain from private conversations (voice or electronic)
- Refrain from activities unrelated to current tasks in class
- Treat others with respect and dignity

Text book and meeting times

Data Science from Scratch: First Principles With Python. First Edition. Joel Grus.

R Programming for Data Science. Roger Peng

```
Section 1:  Tuesday, Thursday   09:55-11:40, Morken #203 (Dr. Cao)
```

Course website

https://www.cs.plu.edu/~caora/cs133/

Course Goals

- Develop important problem solving skills by programming
- Explore the Python and R programming language
- Better understand Computer Science as a discipline
- Have fun writing computer programs and analyzing data!

Course Grade

Innovation project and class participation - 10%

- *Use knowledge learned from the class, propose ideas and analyze the data*
- *Participate in each lecture*

Projects - 20%

- *Around two projects during the semester*
- *Late assignments will be docked 25% per day (every 24 hours).*

Exams - 15%

- *One mid-term exam*

Quizzes and homework - 40%

- *5 to 8 quizzes or homework*
- *drop lowest score*
- *no makeup quizzes*

Final Project --15%

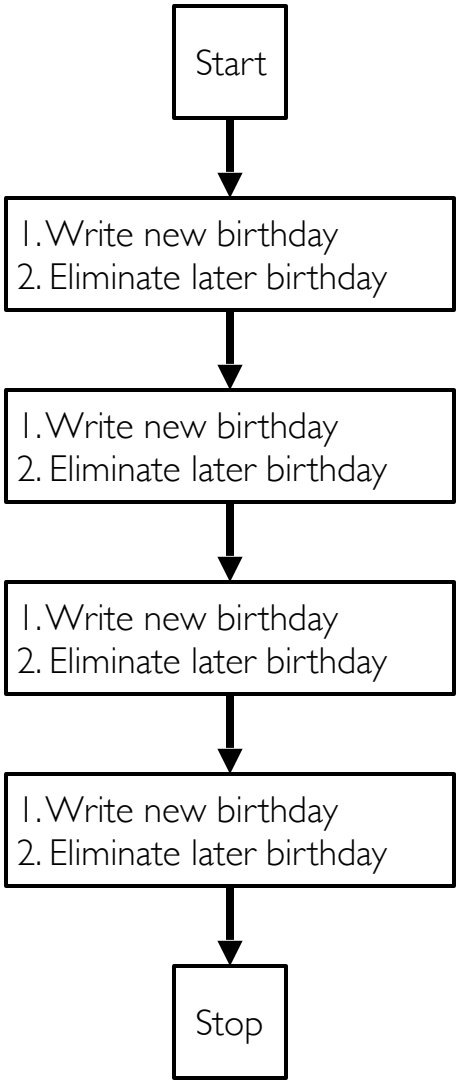
- Comprehensive, required for all students
- Consist of a report, a program and a presentation

:

Overall Score	Grade
100% -- 90%	A / A-
90% -- 80%	B+ / B / B-
80% -- 70%	C+ / C / C-
70% -- 60%	D+ / D / D-
60% -- 0%	E

Finding the earliest birthday - method 1

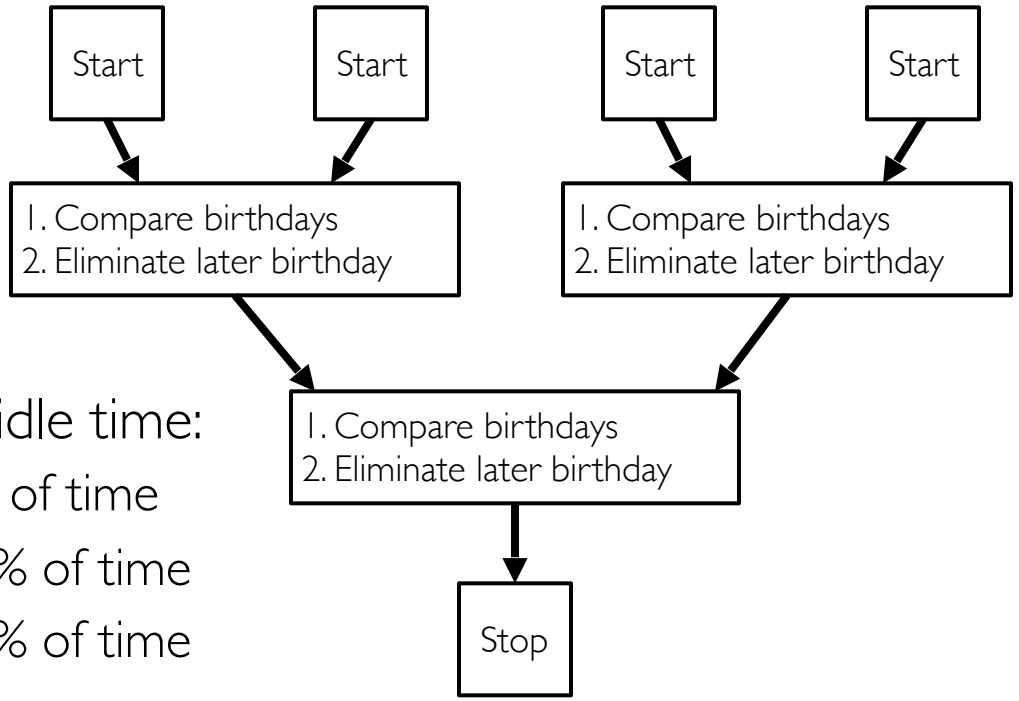
- Requires as many steps as people:
  - 4 people – 4 steps
  - 16 people – 16 steps
  - 32 people – 32 steps
- Each person spends most of their time sitting idle:
  - 4 people – Each person idle 75% of the time
  - 16 people – Each person idle 94% of the time
  - 32 people – Each person idle 97% of the time



I. Introduction to computation

Finding the earliest birthday - method 2

- Simultaneous events mean fewer steps:
  - 4 people – 2 steps
  - 16 people – 4 steps
  - 32 people – 5 steps



- Fewer steps mean less idle time:
  - 4 people – idle ≤ 50% of time
  - 16 people – idle ≤ 75% of time
  - 32 people – idle ≤ 80% of time

Conclusion #1: Computers can't see the "big picture" – only the immediate task at hand.
Conclusion #2: Not all programs are equal – some are faster or more flexible than others.

II. Problem-Solving

A. Understand the Problem
- Do you understand all the words & terms that are being used?
- What are you being asked to find or show?
- Is there enough information to solve the problem?
- Can you draw a picture that might help?

B. Come Up With a Plan
- Guess and check, make a list, or draw a picture.
- Look for a pattern, or find a key equation.
- Try solving a simplified version of the problem.
- Work backwards.

C. Carry Out the Plan
- Be aware that you may run into roadblocks or dead-ends!
- Check to see if your results make sense.
- Don't be afraid to start over!

D. Make Your Solution Computer-Friendly
- Imagine you are writing to a student not in this class.
- Keep things brief… but make sure that you don't leave anything out.
- Write a step-by-step list of instructions… like writing a recipe.

II. Problem-Solving

Some Practice Questions

Here are a few problems to think about. Use the strategies from the previous slide, and write down at least three facts or observations that you think are important when it comes to solving the problem. We'll discuss the pros and cons of each fact/observation before trying to solve the problems.

1. Same birthday. You and your classmates want to know if there are students sharing the same birthday. You have everyone's birthday date (Month and Day), how do you quickly find it out?

2. Pizza Prices. You're trying to decide what size pizza to order, and have the choice of a 12" pizza for $13 or a 14" pizza for $16. Which one gives you the most pizza per dollar?

3. Finding the Day of the Week. What day of the week is 23 December 2017? What about 23 December 2087?

III. Environment setup

Python

- Go to https://www.python.org/downloads/
- Download Python 2.7.11
- Select the option "add python.exe to Path"
- Select "Will be installed in local hard drive"
- Double click on the file and click next/yes for all questions
- Wait a moment……
- Click Finish
- Done!

III. Environment setup

Atom

- Go to https://atom.io/
- Click on Download installer
- Open the file and follow the instructions

- You are ready to go!

III. Environment setup

Anaconda

- I am aware that the book talks about Anaconda. Here is my opinion about it:

- It is really good!
- If you are interested in installing Anaconda instead of Python and pip… Go for it!
- In class we will use regular Python and Atom, but I don't want to stop you in any way from exploring other software

III. Environment setup

R

- Go to https://www.r-project.org/
- Click on Download CRAN
- Select the area that it is closest to you (Fred Hutchinson Cancer Research Center, Seattle, WA)
- Select your operating system
- Select "base"
- Download file
- Click on the file and click next on all the prompts. (leave the default values)
- Click Finish!

Resources

Before you leave today…

apply for a curly account

https://www.cs.plu.edu/hub/accounts/requests/new

Finish survey about your background (available on course website):

https://www.surveymonkey.com/r/LFFGBQ2

For next time…

- Check the book if you would like to buy one
- Check out class website and Sakai regularly:

https://www.cs.plu.edu/~caora/cs133/

Resources

Account for the first few days (active through Feb 22, 2017):

user name: firstday
password: Spring17

Resources

Create an Account…

– Open the Firefox browser

– Go to https://www.cs.plu.edu/hub/

– Click on *Request link*

– Review PLU Policies

– Click on *I agree link*