CompSci 101
Introduction to Computer Science

January 19, 2016

Prof. Rodger
Announcements

• Reading for next time on calendar page
  – RQ 2 due Thursday

• Assignment 1 out today
  – Have fun with Blockly

• Install your environment

• Lab 1 this week!

• Plan for Today:
  – Problem Solving and Python
Review Scratch program

- "Hello World"
- Scratch Program
- Colors
  - Duke blue: motion
  - Mustard: control
  - Light blue: sensing
  - Orange: data
  - Purple: looks
```python
'''
Created on Jan 14, 2016
@author: Susan
'''
print "hello CompSci 101 students"
```
Python data reading code

```python
f = open("kjv10.txt")
st = f.read()
total = len(st)
zc = st.count('z')
print "total # chars = ",total
print "number of z's",zc
for ch in 'aeiou':
    print ch, st.count(ch)
```
Explaining Python code?

Duke Connection: Fred Brooks '53

• What Would FB Say?
"The most important single decision I ever made was to change the IBM 360 series from a 6-bit byte to an 8-bit byte, thereby enabling the use of lowercase letters. That change propagated everywhere."

"Fred Brooks" by Copyright owned by SD&M (www.sdm.de) - Request for picture sent by email to Fred Brooks by uploader (Mark Pellegrini; user:Raul654) Fred sent this photo back, along with contact information for Carola Lauber at SD&M, who gave copyright permission. Licensed under CC BY-SA 3.0 via Wikimedia Commons - https://commons.wikimedia.org/wiki/File:Fred_Brooks.jpg#/media/File:Fred_Brooks.jpg
Why is programming fun?

Fred Brooks

• First is the sheer joy of making things
• Second is the pleasure of making things that are useful
• Third is the fascination of fashioning complex puzzle-like objects of interlocking moving parts
• Fourth is the joy of always learning
• Finally, there is the delight of working in such a tractable medium. The programmer, like the poet, works only slightly removed from pure thought-stuff.
Lab 1 This Week

• Install Before attending if can
• Modify Python program
• Scratch program
Our Programming Environment

• Install 5 items

• Why Java? – not using
• Eclipse – platform for development
• Python – programming language
• Pydev – Python IDE for Eclipse
• Ambient – turnin/snarf files to/from Duke
How does one get help in CompSci 101?

- Consulting hours
  - Sunday-Thursday 7-11pm
- Office hours (prof, TAs)
- Collaborate with other students
- Piazza
  - Ask questions
  - Do not post your code and ask what is wrong!
  - Post error message and line of code for error message
  - If added class late, may need to add yourself
How to succeed in Compsci 101

• Start assignments early, they'll take longer than you think they will
• Read the book, but we'll cover the material in class, so …
• Collaborate well, but be sure you can do work on your own!
• Be curious, work hard at beginning, think carefully
Algorithm

- Recipe
- Sequence of steps that constitute instructions
- Step-by-step procedure for calculations

What does Nate Silver do?
http://53eig.ht/1tZy909

How do Netflix and Amazon know me?
- Compsci101 project: capable of implementation as a program, but much more basic
  http://moreintelligentlife.com/content/features/anonymous/slaves-algorithm
Google “algorithm”

- AI Algorithm Identifies Humorous Pictures
  MIT Technology Review - Jan 8, 2016
  Today, Arjun Chandrasekaran from Virginia Tech and pals say they’ve trained a machine-learning algorithm to recognize humorous scenes ...

- Tinder’s secret algorithm matches users based on their ‘des’
  That successful connection you made on Tinder was not fate; it was an algorithm deciding how desirable you truly are. Tinder’s CEO Sean Rad ...

- Tripdelta finds the cheapest fares using a complicated algorithm
  Daily Mail - Jan 13, 2016
  Using a complicated algorithm, Tripdelta.com searches the depths of the internet to find ‘secret’ flights that are as much as 80 per cent cheaper ...
Developing an Algorithm

- [http://www.youtube.com/watch?v=AEBbsZK39es](http://www.youtube.com/watch?v=AEBbsZK39es)

$193, $540, $820, $700, $749. Are these reasonable? Why?
I'm thinking of a number …

• You guess. I'll tell you *high, low, or correct*
  – Goal: guess quickly, minimal number of guesses
  – Number between 1 and 100…
  – Number between 1 and 1000…

• Can you describe an algorithm, instructions, that would allow someone to use your instructions to play this game correctly. Start with 1 and 100, but ideally your instructions work with 1 and N

Analyzing the *binary search* algorithm

- Is the algorithm correct?
  - Try it, again, and again and …
  - Reason about it: logically, informally, …

- How efficient is the algorithm?
  - How many guesses will it take (roughly, exactly)
  - Should we care about efficiency?

- When do we really care about efficiency?
  - Examples?
Find Narten
Looking for a Needle in a Haystack

- If a computer can examine 10 million names/numbers a second, suppose the list isn't sorted, or I say "yes/no", not "high/low"
  - How long to search a list of 10 million?
  - How long to search a list of a billion?
  - 14 billion pixels in a 2 hour blu-ray movie

- What about using binary search? How many guesses for $1000, 10^6, 10^9, 10^{12}$
  - One of the things to remember: $2^{10} = 1024$
Algorithms that scale: another example

• Human Genome Project
  – Multiple approaches, relying heavily on computational power and algorithms
  – Combine reads of DNA sequences, we'll look at an illustrative example

• These combine bio/chemistry techniques with computational techniques to recreate the sequencing, e.g., CGATTCCG… from "live data", actual DNA.
Eugene (Gene) Myers

- Lead computer scientist/software engineer at Celera Genomics, then at Berkeley, now at Janelia Farms Research Institute (HHMI)

"What really astounds me is the architecture of life. The system is extremely complex. It's like it was designed." ... "There's a huge intelligence there."

- BLAST and WG-Shotgun
Whole Genome Shotgun with words

- Creation algorithm
  - Take a phrase
  - Replicate it four times
  - Chop into "chunks"
    - 15-22 characters

- How to recreate original phrase?

From Algorithms to Code

• An algorithm that scales needs to run on a computer --- programming to the rescue!

• Extensive libraries help with programming
  – Brain or Neuroscience
  – Engineering and Mathematics
  – Genomics
  – Graphic User Interfaces, …

• We are using Python, extensible and simple
Understanding terminology: code

• Move from "Hello World" to "Hello Around the World"
  – Look at Python, code, libraries
  – Learning (reviewing) terminology about Python

```python
print "hello world"
```

```python
f = open("hello.txt")
for line in f:
    print line
```
Hello around world code

Running and Understanding Code

- Need Python compiler/interpreter
  - We're using Canopy, includes libraries
- Need an editor development environment
  - We use Eclipse and PyDev, open source and widely used, Ambient is Duke Plugin
- You need experience thinking and coding and debugging ideas and code:
  - Installing the suite of tools can be cumbersome
    - Persist, Perservere, Get Help, start over 😊