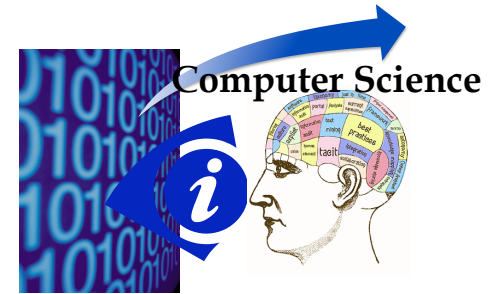


**COMPSCI 101, Spring 2012**  
***Introduction to Computer Science***  
**Owen Astrachan**

<http://www.cs.duke.edu/courses/spring12/compsci101>  
<http://www.cs.duke.edu/~ola>

**Data into Information and Knowledge**



**men and women: this is compsci 101**

**Prerequisites for Compsci 101**



**Is a picture worth a thousand words?**

- <http://www.google.com/images?q=world%20wide%20web&biw=1238&bih=969>
- <http://www.google.com/images?q=computer%20science&biw=1370&bih=1081>
- <http://www.google.com/images?q=internet&biw=1370&bih=1081>
- <http://www.google.com/images?q=programming&biw=1370&bih=1081>



## Anatomy of a search [query]

- <http://www.google.com/images?q=programming&biw=1370&bih=1081>
- What comes after the question-mark in the URL?
  - What is the query string?
  - What are the browser dimensions?
  - What is constant in the search query, what changes?
  - How is the query *parameterized*?
  - How are multiple-word queries handled?
- What does this have to do with Computer Science and programming?

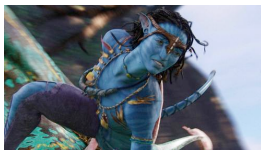
## Questions about Computer Science

What is it that distinguishes it from the separate subjects with which it is related? What is *the linking thread* which gathers these disparate branches into a single discipline? My answer to these questions is simple --- it *is the art of programming a computer*. It is the art of designing efficient and elegant methods of getting a computer to solve problems, theoretical or practical, small or large, simple or complex.

C.A.R. (Tony)Hoare

## Milking Stool v Neural Queue

- Engineering, Mathematics, Science
  - Pillars of computer science?
  - Braid of computer science?
- Other aspects of CS?
  - Technology and policy
  - Art and visualization
  - Collaborative Filtering



## It's not all programming at all, but ...

- What is the nature of intelligence? How can one predict the performance of a complex system? What is the nature of human cognition? Does the natural world 'compute'?
- *It is the interplay between such fundamental challenges and the human condition that makes computer science so interesting.* The results from even the most esoteric computer science research programs often have widespread practical impact. Computer security depends upon the innovations in mathematics. Your Google search for a friend depends on state-of-the-art distributed computing systems, algorithms, and artificial intelligence.

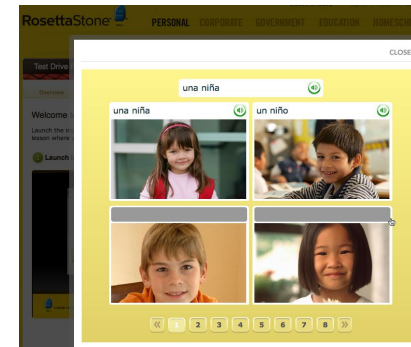
<http://www.post-gazette.com/pg/04186/341012.stm>

## Understanding Information/Data

- Does understanding computer science help you when you want a new smart phone?
- Does knowledge of programming help you get your laptop connected to a wireless access point?
- Does experience with algorithms and algorithmic approaches help physicians and attorneys?
- Are these important to society, to you?

## How will you learn to 'speak'?

- <http://www.rosettastone.com/personal/demo>



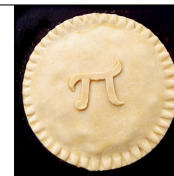
## What language will we learn?

- <http://www.python.org/>
- Python is a *multi-paradigm* language
  - Procedural
  - Functional
  - Object-Oriented
- Simple, huge libraries, widely used
- Guido is BDFL



## Why is it called Python?

- <http://www.youtube.com/watch?v=anwy2MPT5RE>



+



- 9.8 m/sec<sup>2</sup>

C

## Course Overview: Is this the right one?

- There are details, see the [course web page](#)
  - Midterms and final are open book, what does that mean?
  - APTs: Algorithmic Problem-solving and Testing
    - Weekly small programming assignments, tested online
  - Programming assignments: solo, group, ...
  - Lab/recitation: group and individual work
- Why should you come to class?
  - Meet people, learn things, participate in a community
  - Provide help, get help, wonder, dance, think
- Why is this course so great?
  - Because you're in it

Compsci 101, Spring 2012

1.13

## What's in Compsci 101?

- Learning about computing, computer science, and programming
  - Vocabulary of Python and programming languages
  - Crafting programs from the vocabulary
  - *Power of automation, repetition, scale*
  - Understanding and changing the world
- Programming using Python
  - Tools: Eclipse, EPD, Libraries, ...
  - Engineering and analyzing designs and programs
  - Using mathematical and scientific techniques
  - Appreciating and learning art *and* science of programming
  - Moving toward scaling solutions (continued later courses)

Compsci 101, Spring 2012

1.14

## Questions

If you gotta ask, you'll never know  
Louis Armstrong: "What's Jazz?"



If you gotta ask, you ain't got it  
Fats Waller: "What's rhythm?"



What questions did you ask today?  
Arno Penzias

Compsci 101, Spring 2012

1.15

## From Blog to Scientific Visualization

- Text Cloud aka Tag Cloud?
  - Number of occurrences/emphasis indicated by size of word
  - Great visual/statistic: <http://chir.ag/phernalia/preztags/>
  - <http://www.nytimes.com/gst/mostsearched.html?period=30&format=tagcloud>
    - What information is stored in the URL of the NYTimes site above?



Compsci 101, Spring 2012

1.16



## Problem Solving and Programming

- How many words are in a file? A webpage?
  - What's a word?
  - What's a file?
  - How do we solve this: simply, quickly, ...?
    - What's the best we can do? Constraints?
- How many different/unique words are in a file?
  - How is this related to previous task?
- How many words do two files have in common?
  - Spell-checking, Google did you mean ...?
- How many codons common to DNA strands?

## Toward a Pythonic Tagcloud generator

- This is valid and correct Python code, questions?

```
def countWords(filename):  
    file = open(filename)  
    str = file.read()  
    words = str.split()  
    unique = set(words)  
    print "filename: ", filename  
    print "total # words = ",len(words)  
    print "unique # words = ",len(unique)  
if __name__ == "__main__":  
    countWords('/data/kjv10.txt')
```

```
filename: /data/kjv10.txt  
total # words = 823135  
unique # words = 34027
```

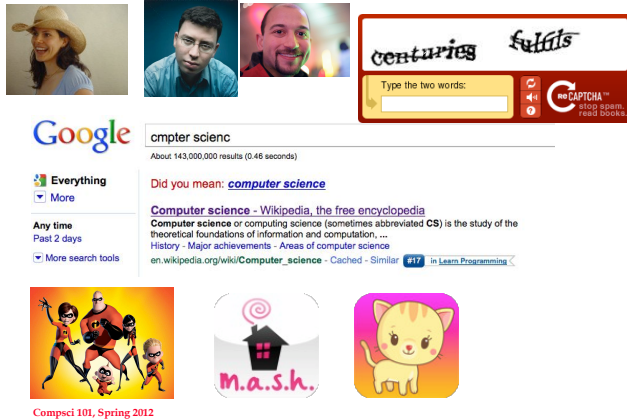
## Python vocabulary

- Python has a large standard library
  - Organized in *modules*: sys, io, math, os, ...
  - <http://docs.python.org/library/index.html>
  - API browseable online, but Eclipse IDE helps a lot
- Python users often use third-party libraries too
  - Scientific, visual, plotting, ...
  - We will use EPD: Enthought Python Distribution
- Python is a multi-paradigm language, though this won't matter so much in the beginning
  - Very useful later!

## Python and Programming Concepts

- Names are important, abstractions
  - What is <http://152.3.140.1>
  - What is <http://www.amazon.com>
- Types are important, facilitate operations
  - What is foo.pdf, foo.mp4, foo.jpg, foo.wav
  - Do the file extensions guarantee file type?
- Thinking in terms of names and types can help
  - Python has types, inferred dynamically
  - Python uses types differently from Java and C++
    - Static v Dynamic

## Duke Contributions



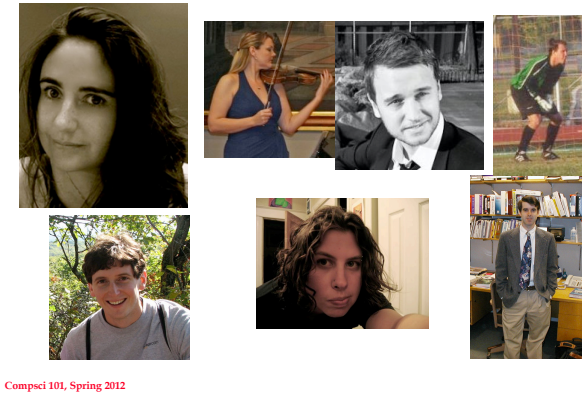
Google **computer science**  
About 143,000,000 results (0.46 seconds)

Did you mean: [computer science](#)

**Computer science** - Wikipedia, the free encyclopedia  
Computer science (sometimes abbreviated CS) is the study of the theoretical foundations of information and computation, ...  
History · Major achievements · Areas of computer science  
[en.wikipedia.org/wiki/Computer\\_science](https://en.wikipedia.org/wiki/Computer_science) - Cached - Similar #17 in Learn Programming

Compsci 101, Spring 2012 1.21

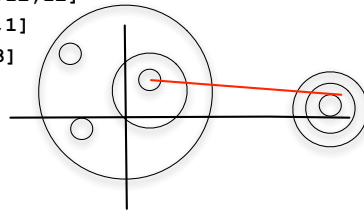
## Semi-random former students



Compsci 101, Spring 2012 1.22

## From Idea to Algorithm to Program

- **Finding minimal crossing count**
  - Applications? Puzzle-solving?
  - Given  $(x_i, y_i, r_i)$  and P1, P2 --- determine minimal number of circles to cross to get from P1 to P2
- **Input:**
  - $[-3, 2, 2, 0, -4, 12, 12, 12]$
  - $[-1, 2, 3, 1, 5, 1, 1, 1]$
  - $[1, 3, 1, 7, 1, 1, 2, 3]$
  - $(2, 3) (13, 2)$
- **Output:**
  - 5



## 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.

