CompSci 201, L28: LDOC and the Limits of Computing

Logistics, coming up

- End of Semester Survey
  - Due *today,* LDOC this Wednesday, April 26
  - Required class survey, run by the course staff
- Course Evaluations
  - Due by this Saturday, April 29
  - Run by Duke, anonymous to us

Final Exam Logistics Reminder

- 3 final sections: F1, F2, F3 corresponding to 3 midterms M1, M2, M3.
  - Exams grade = Avg(Max(M1, F1), Max(M2, F2), Max(M3, F3))
  - If happy with grades? Don't need to take it.
  - If you missed a midterm? Make sure to take at least that part.

- Monday 5/1, 9 am - noon. You will have 50 minutes to complete each part.
  - 9-9:50 am. F1 (corresponding to M1)
  - 10-10:50 am. F2 (corresponding to M2)
  - 11-11:50 am. F3 (corresponding to M3)

Final Grade Estimates

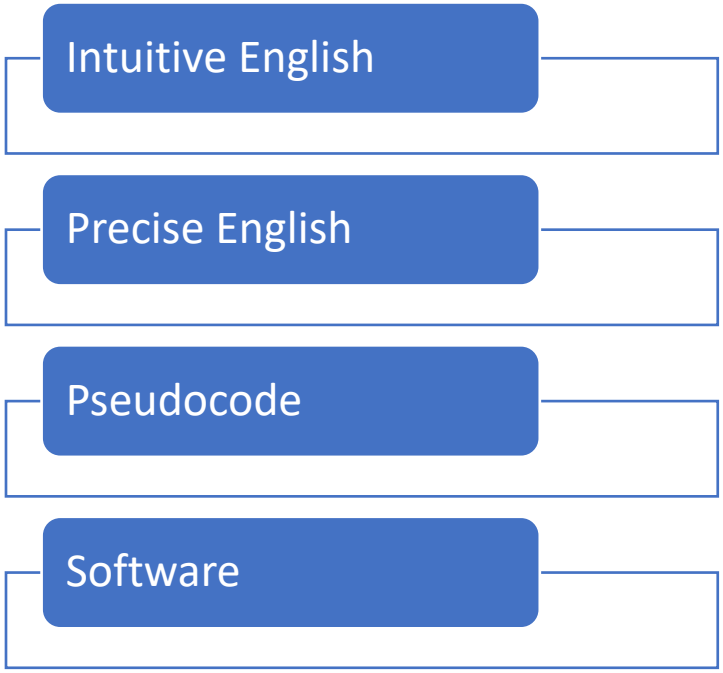
- Everything should be live and updated in the sakai gradebook *except*:
  - P6

- We aim to get this up by this weekend

Looking back at our semester

What are algorithms?

Loosely speaking: A precise sequence of unambiguous steps that effectively compute an output given an input.



Algorithm Design
- Mathematical
- Logic of program
- Problem-solving
- Language independent

Implementation
- Semantics and Syntax
- Language dependent
- Programming on a real machine

What is code?

In order to execute an algorithm on a real computer, we must write the algorithm in a formal language. An algorithm so written is a **program**.

In this class we explore both:

Theory

- Design an algorithm
- Analyze performance
- Data structure tradeoffs

Practice

- Write a Java program
- Debug/test
- Measure performance

Why does efficiency matter?

- You wrote the next big social media app:
  - Will it work if it has 1 billion users?
  - What about on a phone with limited memory?
- In the sciences, discovery depends on computing with big data:
  - Sequencing the human genome
  - Surveying millions of images in astronomy
  - Processing data logs from the CERN collider
- Pushing the limits of current technology:
  - Virtual / augmented reality?
  - Deep neural networks for large scale machine learning?

Some specifics you ~~will~~ did learn

Data Structures

- Arrays
- Lists: ArrayList and LinkedList
- Sets: HashSet and TreeSet
- Maps: HashMap and TreeMap
- Stacks, Queues, Priority Queues / Heaps
- Trees: Binary Search Trees
- Graph representations

Algorithms

- Iterative
- Hashing
- Big O Asymptotic Analysis
- Recursive
- Sorting
- Greedy
- Graph

Software

- Java API
- Objects, Classes
- Interfaces, implementations
- Testing, Debugging

Informal goals for the course

- Make or deepen a friendship with someone else passionate about computer science.
- Develop a new appreciation of computing phenomena you see in the real world.
- Experience joy when your program *works*, even if it took a while to get it there.
- WOTO: WOrking TOgether
- Stay safe and healthy, physically and mentally

Who to Thank

- Grad TAs, Professor O'Hanlon, working behind the scenes to make this work at scale
- All of our undergrad TAs. Providing feedback, helper hours, running discussion, etc.
- Your fellow students! Discussion groups, friends, project partners, etc.

What I’m thankful for

- Safety to gather and be together
- My teaching team
- All of you (why am I here?!?!?)

Parting Thoughts: What computers can and can’t do?

What can computers do?

What can't computers do?

- Some problems ***cannot be solved at all***
  - One program detects all infinite loops
- Some problems ***cannot be solved efficiently***
  - Listing all N-bit sequences of 0's and 1's
- Some problems can be ***approximately solved***
  - AI, ML, close-to-optimal is good enough

Halting Problem

- Can we write doesHalt as specified? *Suppose so!*
  - Like the Java Compiler: reads a program

```
public class ProgramUtils
    /**
     * Returns true if progname halts on input,
     * otherwise returns false (infinite loop)
     */
    public static boolean doesHalt(String progname){
    }
}
```

Can we confuse doesHalt?

- What if **`doesHalt(confuse)`** returns true?
  - Then **`confuse()`** does not halt (see below)
- What if **`doesHalt(confuse)`** returns false?
  - Then **`confuse()`** does halt (see below)

```
public static boolean confuse(){
    if (ProgramUtils.doesHalt(confuse)) {
        while (true) {
            // do nothing forever
        }
    }
}
```

Formal proof is a bit more challenging…

- Alan Turing first showed this for programs: 1936
  - Had to formally specify what a program was
  - Needed to invent concept of Turing Machine
  - Also demonstrated by Alonzo Church

- Cantor showed # Real Numbers > # Rationals
  - So-called diagonalization, 1891
  - Ridiculed by establishment
  - Argument essential to above

Shortest/Longest Path; P and NP

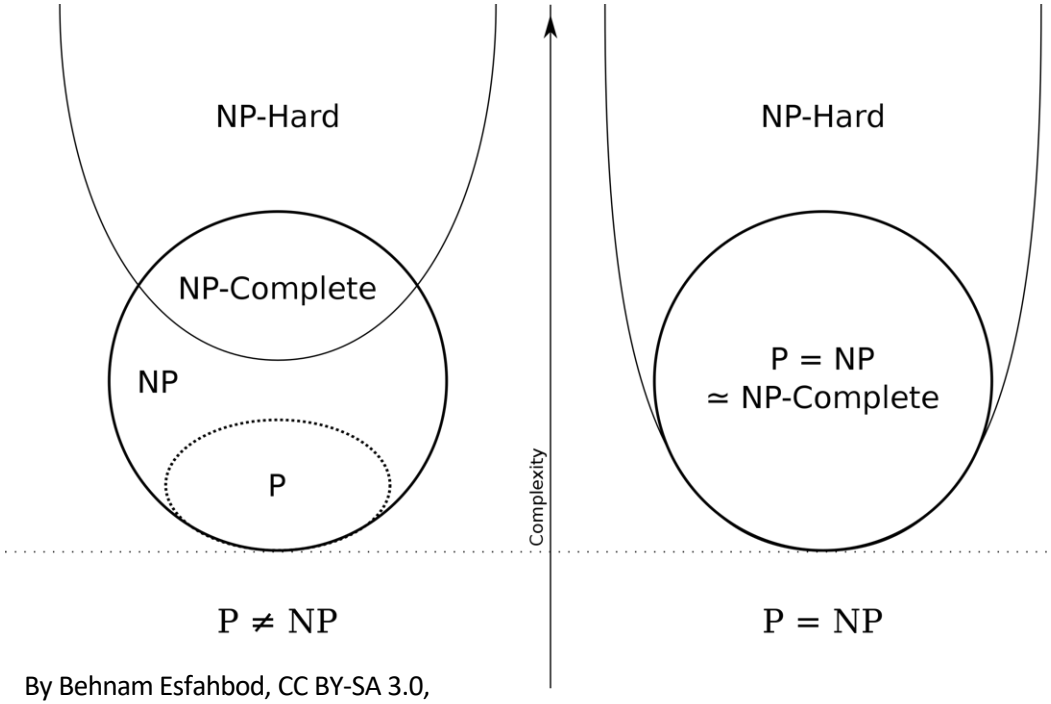
- Dijkstra's Algorithm one example
  - Others: Floyd-Warshall and more
  - Very efficient graph algorithms,

- Longest Path? No efficient solution known
  - Easy to verify "is this path greater than length k"
  - Exponentially many paths

P vs NP

- P is the set of (algorithmic) problems that can be solved by a deterministic Turing Machine (DTM) in time that is polynomial in the size of the input (polynomial time).
  - i.e., can solve with a program that is O(1), O(N), O(Nlog(N)), $O(N^2)$, $O(N^3)$, …, $O(N^{128})$, …
- NP is (roughly) the set of (algorithmic) problems for which a solution can be *verified* by a DTM in polynomial time.

P ?= NP



By Behnam Esfahbod, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=3532181

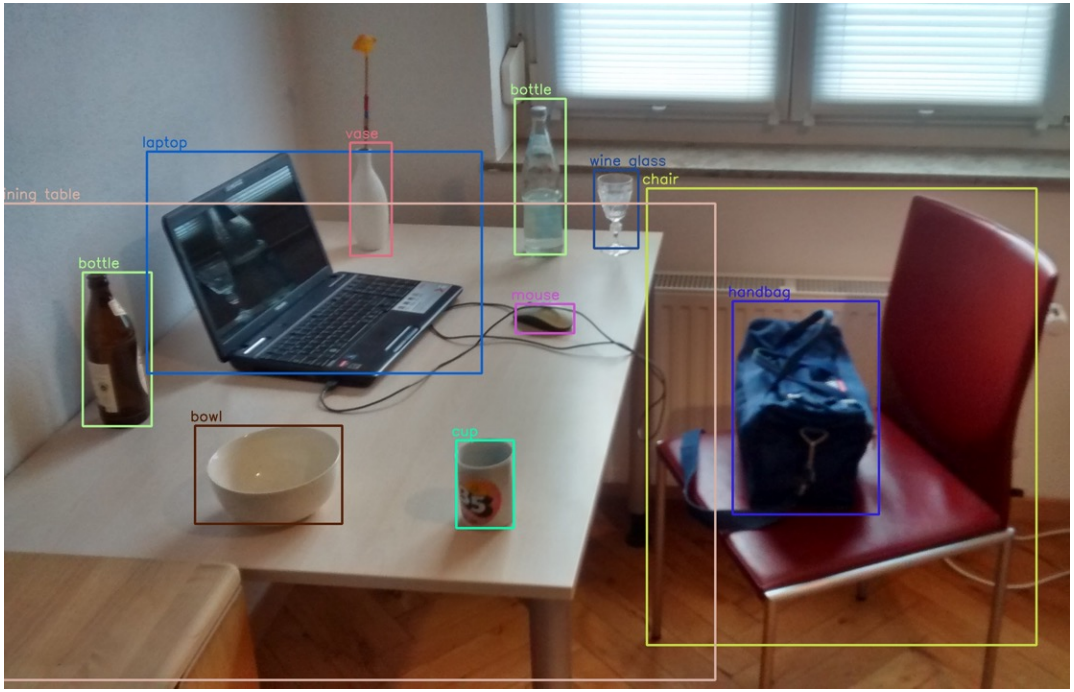
- Most think P != NP
- Greatest outstanding question in theoretical computer science
- Proof is worth a $1M prize from the Clay Mathematics Institute

“Easy” Hard Problems

- Some problems are hard to solve but easy to approximate:
  - Can’t write a program to give you the optimal solution efficiently but can find something within ϵ of optimal in polynomial time.
  - Greedy, randomized, etc.

- Some problems are hard to prove things in theory but easy to solve in practice
  - Can’t prove much but it works well in practice

AI/ML often work with experimental algorithms for hard problems



Common idea: Use a computer to learn a function/neural network that approximates a large dataset.

- Image segmengation / classification
- Face/speech recognition
- Machine translation
- Text generation
- Reinforcement learning
- Robotics
- …

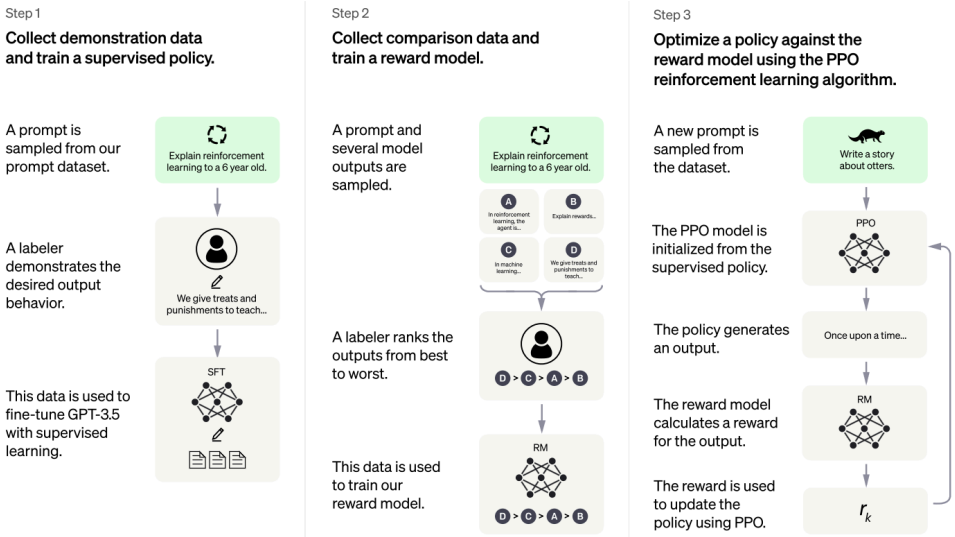
Artificial Intelligence: ChatGPT and Reinforcement Learning

BT

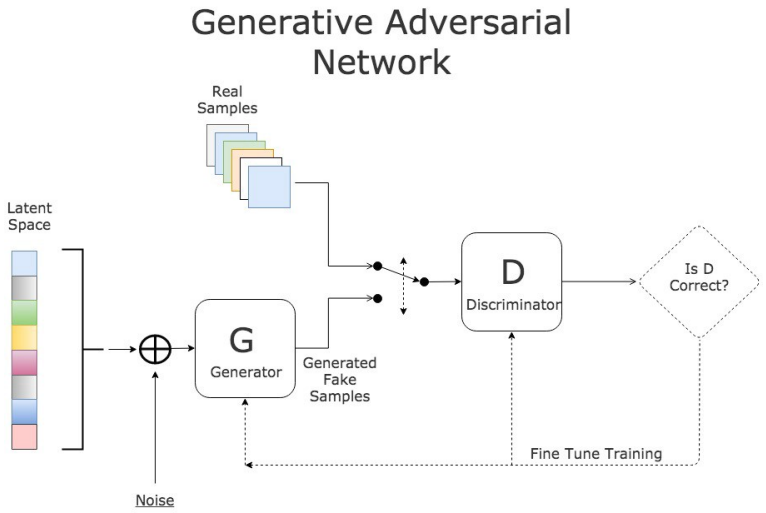
Do you think students should use chatgpt to do their homework?

I am an AI model and do not have personal opinions or preferences. However, I can tell you that using a tool like me or GPT-3 (chatGPT) to generate text or complete homework assignments is not recommended. This is because using such a tool would be considered plagiarism, as the work being submitted is not the student's own original work. Additionally, homework and other assignments are designed to help students learn and develop their skills, so using a tool like me would not give the student an opportunity to gain the knowledge and understanding that is the main objective of the assignment.

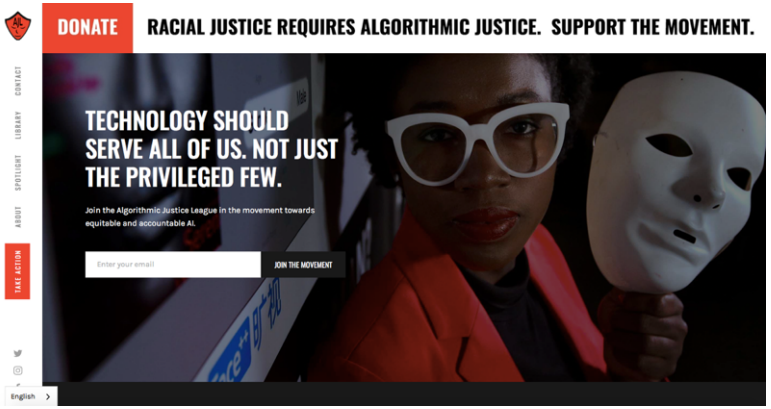
Additionally, using such AI model to complete homework assignments could be counter-productive to the educational process, it is important to complete assignments on their own and ask questions to the teachers when the student have doubts. This way the student will have the opportunity to learn the material and develop their skills, rather than just relying on a tool to do the work for them.



More Concerning? Deepfakes and Generative Adversarial Networks (GANs)



What should computers do?



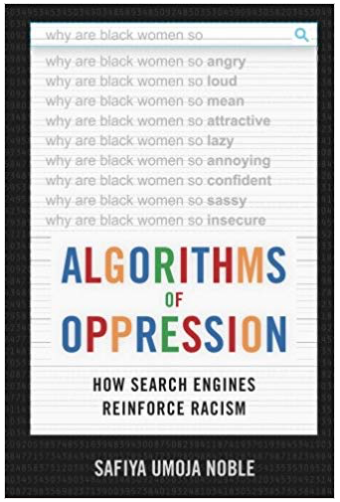
Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights

Executive Office of the President

May 2016

Photo-illustration: Gluekit

Microsoft's Tay chatbot started out as a cool teenage girl, but quickly turned into a hate-speech-spewing disaster.



We need computational citizens who can deal with these algorithmic systems in theory and practice

Who has gone before you? People in CS

What will you do?

- Not everyone wants to be a software engineer
  - Diplomat, lawyer, physician, entrepreneur,
  - Musician, teacher, data scientist, …
- Not all jobs at tech companies are SWE
  - UI, UX, PM, …
- Some non-tech companies have tech jobs
  - Healthcare? Aerospace? Biotech? Finance? Non-profit? NASA?
- Grad school? Research? Teaching?

What I’m thankful for

- Safety to gather and be together
- My teaching team
- All of you (why am I here?!?!?)