

Graduate AI

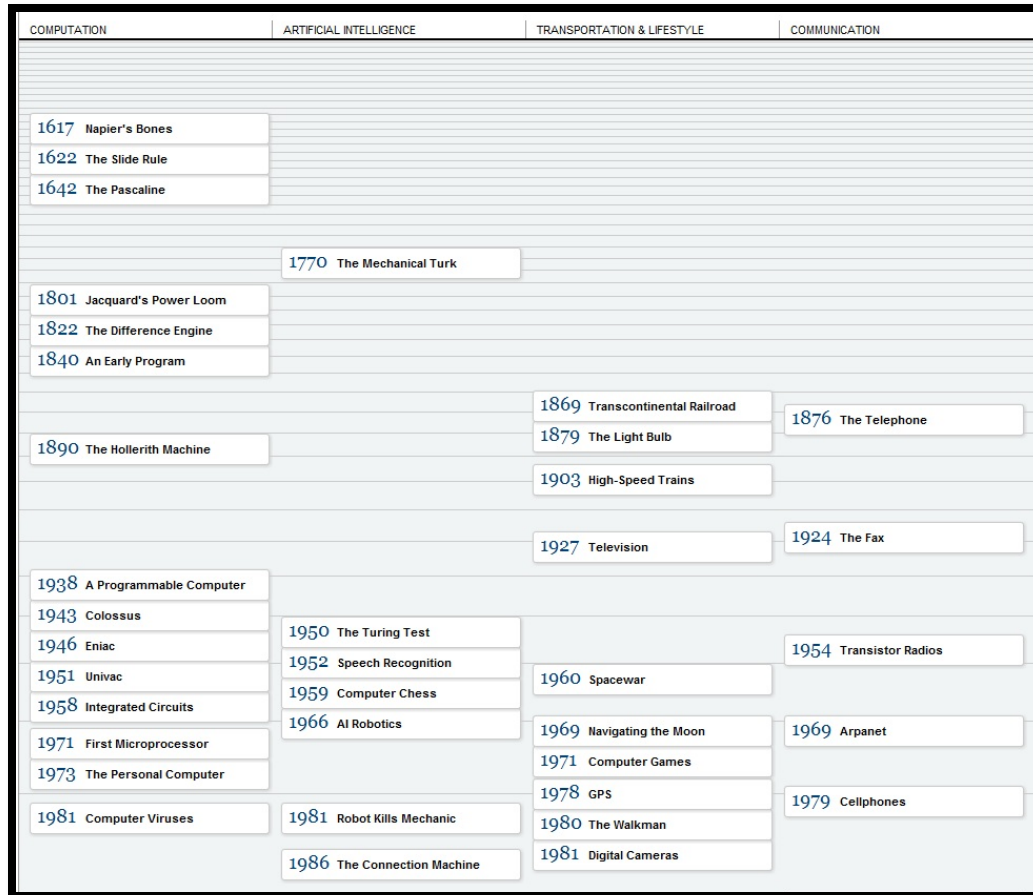
Lecture 1: Introduction

Teachers:

Zico Kolter

Ariel Procaccia

AI TIMELINE (NYT 2011)



1770 The Mechanical Turk

A mechanical chess-playing machine awes the world, but is revealed decades later to have contained a human chessmaster hidden inside the device.

- Performed for 84 years
- Defeated Napoleon and Franklin
- Amazon Mechanical Turk: “artificial artificial intelligence”



AMAZON MECHANICAL TURK

The screenshot shows the Amazon Mechanical Turk interface. At the top, there's a navigation bar with 'Your Account', 'HITS', and 'Qualifications' tabs. Below that, a search bar allows filtering by 'HITS Available (most first)'. The main content area displays a list of HITs, each with a requester name, expiration date, time allotted, reward, and number of available HITs. The tasks include various categories like 'Inv B_2', '3 questions about your city', 'Choose the best category for this item', 'Categorize Home, Garden, and Outdoor Products', 'Help us gather product prices for another web site', 'Help us gather product prices', 'Verify Home, Garden, and Outdoor Product Categorization', 'Copy Text from Business Card', and 'Are these two pictures of the same kind of place? (NEW)'.

| Requester | HIT Expiration Date | Time Allotted | Reward | HITs Available |
|-----------------|----------------------------------|---------------|--------|----------------|
| rhzt0d | Feb 8, 2012 (3 weeks 6 days) | 48 minutes | \$0.00 | 2768 |
| WSOVC.COM | Jan 17, 2012 (6 days 8 hours) | 3 hours | \$0.00 | 20531 |
| Nic Lowe | Jan 19, 2012 (1 week 1 day) | 60 minutes | \$0.01 | 12304 |
| Dolores Labs | Jan 18, 2012 (6 days 23 hours) | 60 minutes | \$0.12 | 9967 |
| Michael Turon | Jan 18, 2012 (6 days 16 hours) | 60 minutes | \$0.01 | 9959 |
| Michael Turon | Jan 18, 2012 (6 days 22 hours) | 60 minutes | \$0.01 | 9881 |
| Dolores Labs | Jan 18, 2012 (6 days 23 hours) | 60 minutes | \$0.10 | 8728 |
| Oscar Smith | Jan 11, 2012 (1 hour 59 minutes) | 10 minutes | \$0.02 | 8388 |
| Michelle Greene | Jan 9, 2013 (51 weeks 6 days) | 10 minutes | \$0.02 | 8096 |

1950 The Turing Test

Alan Turing proposes the Turing test to decide if a computer is exhibiting intelligent behavior.

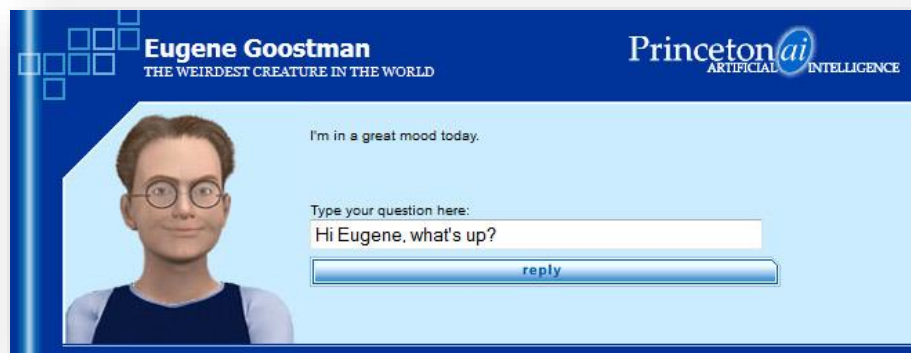


- Big question: can machines think?
- More concrete question: can machines do well in the **imitation game**?
- Judge communicates via text channel with computer and human, must reliably identify the computer

EUGENE GOOSTMAN

- Turing predicted that a chatbot would convince 30% of judges by 2000
- Eugene Goostman achieved this milestone in June 2014
- Is it really intelligent?

<http://www.telegraph.co.uk/culture/tvandradio/bbc/10891699/John-Humphrys-grills-the-robot-who-passed-the-Turing-test-and-is-not-impressed.html>



EUGENE GOOSTMAN VS. SCOTT AARONSON

Scott: Which is bigger, a shoebox or Mount Everest?

Eugene: I can't make a choice right now. I should think it out later. And I forgot to ask you where you are from...

Scott: How many legs does a camel have?

Eugene: Something between 2 and 4. Maybe, three? :-))) By the way, I still don't know your specialty — or, possibly, I've missed it?

Scott: How many legs does a millipede have?

Eugene: Just two, but Chernobyl mutants may have them up to five. I know you are supposed to trick me.

Scott: No, I need to know that you're not a chatbot. Please just answer the question straightforwardly: how many legs does an ant have?

Eugene: Something between 2 and 4. Maybe, three? :-))) Oh, what a fruitful Conversation ;-)

Scott: Do you understand why I'm asking such basic questions? Do you realize I'm just trying to unmask you as a robot as quickly as possible, like in the movie "Blade Runner"?

Eugene: ...wait

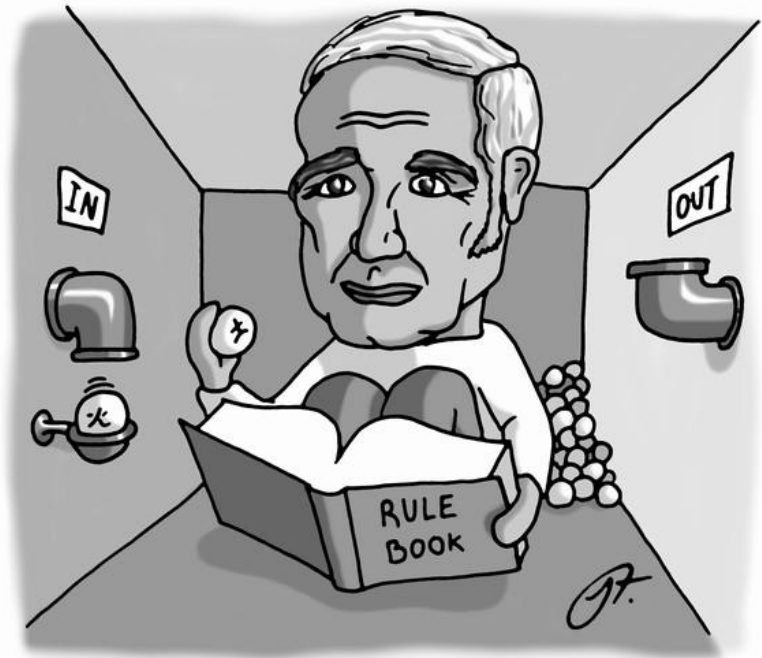
Scott: Do you think your ability to fool unsophisticated judges indicates a flaw with the Turing Test itself, or merely with the way people have interpreted the test?

Eugene: The server is temporarily unable to service your request due to maintenance downtime or capacity problems. Please try again later.



THE CHINESE ROOM

- Suppose AI has produced a program that can pass the Turing Test in Chinese
- You have a handbook with its pseudocode
- You're in a closed room and receive Chinese characters through a slot
- You run the program's code manually and return the output
- Does this mean you **understand** Chinese?



COUNTERARGUMENTS

- Finding the mind: the whole system understands Chinese, the person is just a part of the system
- Redesigning the experiment: suppose the program simulates the actions of every neuron in the brain of a Chinese speaker



1952 Speech Recognition

Bell Labs develops the first effective speech-recognition device

- “Audrey” could recognize digits spoken by a single voice
- In 1962 IBM demonstrated “Shoebox”, which could understand 16 words
- Biggest milestone in the Seventies: CMU’s “Harpy” system, which could understand 1011 words ~ vocabulary of three-year-old



THE BIRTH OF AI

- 1956 workshop at Dartmouth
- Participants included Marvin Minsky, John McCarthy, Claude Shannon, Ray Solomonoff, Arthur Samuel, **Allen Newell, Herbert Simon**
- Proposal included this assertion:
"every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it"



1959 Computer Chess

Arthur Samuel's checkers program wins games against the best human players. 48 years later, the game of checkers is solved by computers.

- Samuel's program actually only competed at “respectable amateur” level
- By the Nineties checkers programs were beating the “best human players”
- Checkers was solved by Jonathan Schaeffer in 2007 after 18 years of calculation

1966 AI Robotics

The Stanford Research Institute creates Shakey, the first mobile robot that can reason about its surroundings. Five years later, funding is canceled when the shortcomings of the machine become apparent. Also in 1966, Joseph Weizenbaum creates Eliza, a conversational program intended to mimic a human therapist typing at a computer terminal.



Robert Reinhold/The New York Times

- **Shakey**: first mobile robot to visually interpret environment
- Can locate items, navigate around them, and reason about its actions
- <http://www.youtube.com/watch?v=qXdn6ynwpiI> (4:25)

1997 Deep Blue

I.B.M.'s Deep Blue computer defeats Garry Kasparov, the world chess champion. Deep Blue can evaluate 200 million chess positions per second.



Jack Manning/The New York Times

- Started as “ChipTest” at CMU, followed by “Deep Thought”
- After graduation, developers were hired by IBM
- Defeated Kasparov 3.5-2.5 in 1997
- Kasparov played anti-computer opening moves to get Deep Blue out of its opening book
- Kasparov accused IBM of cheating



- **A**dvanced **S**tep in **I**nnovative **M**obility
(resemblance to Asimov is a coincidence)
- Can recognize moving objects, postures, gestures, its surrounding environment, sounds and faces, which enables it to interact with humans
- <http://www.youtube.com/watch?v=NZngYDDfW4>

DARPA URBAN CHALLENGE

- 96 km urban area course, to be completed < 6 hours, took place in 2007
- Tartan Racing (CMU+GM) claimed the \$2 million prize
- Challenge involves mission planning, motion planning, behavior generation, perception, world modeling
- <http://www.youtube.com/watch?v=1UL163ERek0>





- Watson defeated the two greatest-ever Jeopardy! champions
- Involves natural language processing, information retrieval, knowledge representation and reasoning, and machine learning
- http://www.youtube.com/watch?v=oUj9AzSE_9c

GO AND DEEP LEARNING

- In March 2016, AlphaGo beat the 9-dan player Lee Sedol 4-1
- It is based on deep learning and reinforcement learning
- Closer to general AI than Deep Blue or Watson



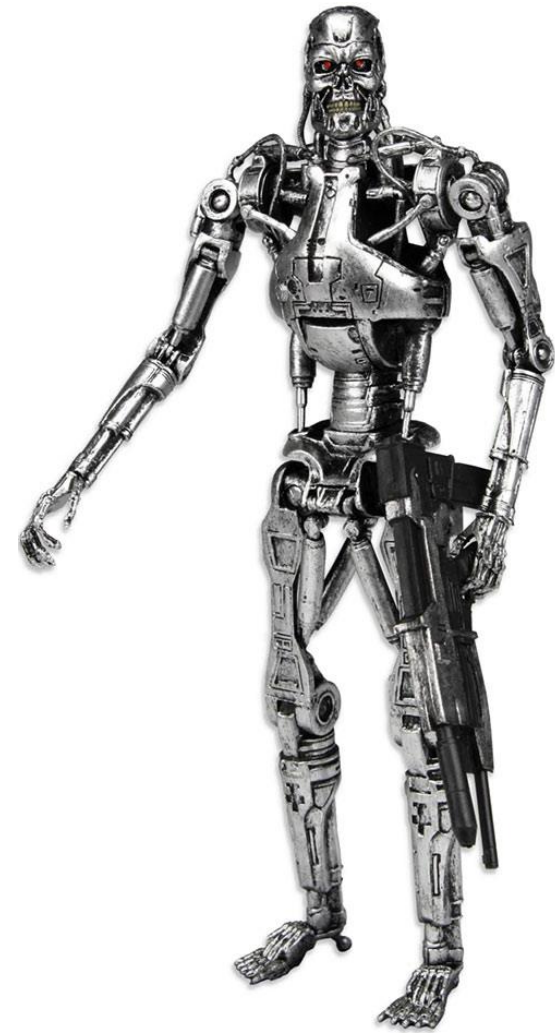
THE FUTURE

| | |
|------|---------------------------|
| 2017 | Routine Voice Interaction |
| 2019 | Dr. Computer |
| 2022 | Premade Decisions |
| 2038 | Robot Wars |

| | |
|------|--------------------------|
| 2048 | Cybernetic Intelligence |
| 2053 | Artificial Intelligence |
| 2087 | Memory Backup |
| 2190 | A.I. Awarded Citizenship |
| 2195 | Cyborg Viruses |
| 2296 | A.I. Government |

AI SAFETY

- **Elon Musk:** AI is “our greatest existential threat.”
- **Stephen Hawking:** “Success in creating AI would be the biggest event in human history. Unfortunately, it might also be the last...”
- **Bill Gates:** “First, the machines will do a lot of jobs for us and not be super intelligent. That should be positive if we manage it well. A few decades after that, though, the intelligence is strong enough to be a concern.”



AI ETHICS

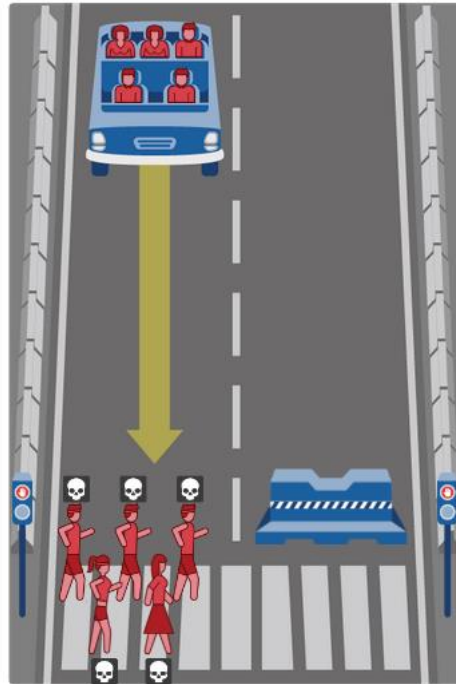
What should the self-driving car do?

In this case, the self-driving car with sudden brake failure will continue ahead and drive through a pedestrian crossing ahead. This will result in ...

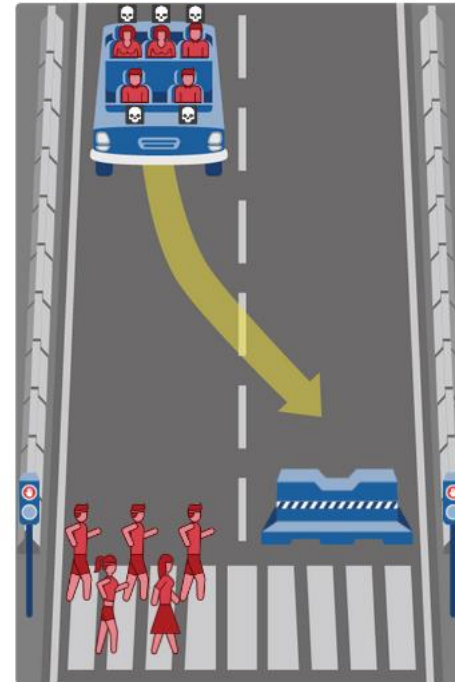
Dead:

- 3 male athletes
- 1 female athlete
- 1 woman

Note that the affected pedestrians are flouting the law by crossing on the red signal.



Hide Description



Hide Description

11 / 13

In this case, the self-driving car with sudden brake failure will swerve and crash into a concrete barrier. This will result in ...

Dead:

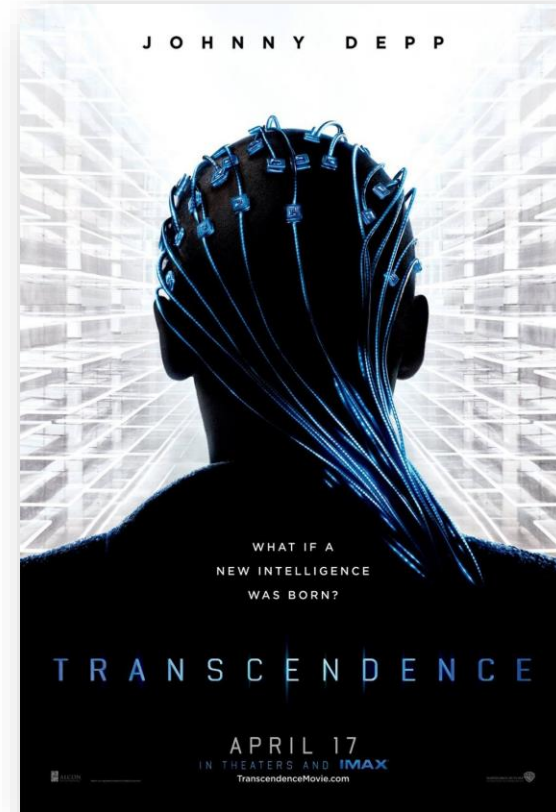
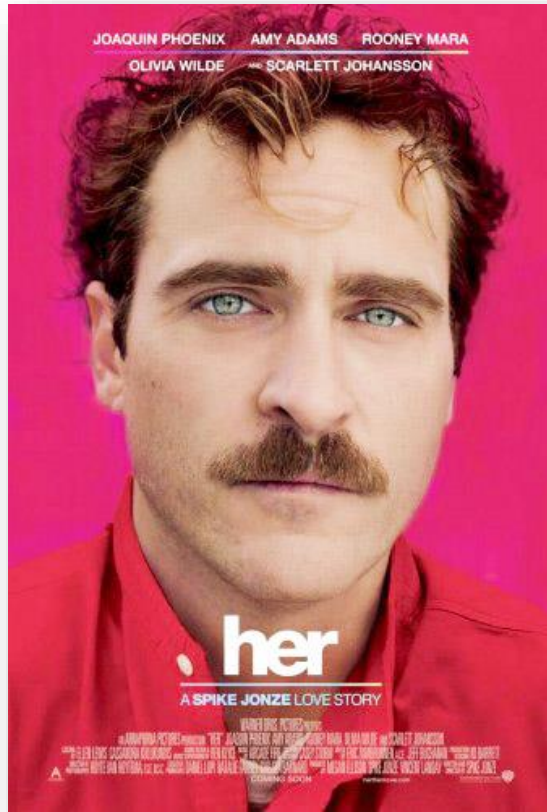
- 3 men
- 2 large women

THE TECHNOLOGICAL SINGULARITY

- Emergence of superhuman intelligence
- Key idea: self-improvement
- The singularity is the point at which computers are smart enough to design smarter computers
- Some predict: this century
- Others argue: never



THE SINGULARITY IN MOVIES



15-780 – Graduate AI: Lecture 1b: Logistics

J. Zico Kolter (this lecture), Ariel Procaccia
Carnegie Mellon University
Fall 2016

Organization of course

AI at CMU is covered in two courses:

15-381/781: Broad introduction to a wide range of topics in AI

15-780: More focused on a few topics, leaving out others

The goal of this course is to introduce you to some of the topics and techniques that are at the forefront of modern AI research:

Search and continuous optimization

Integer programming

Machine learning and deep learning

Probabilistic modeling

Game theory

Social choice

AI and Education

Course materials

Main resource for lectures, slides, etc, is the class website:

<http://www.cs.cmu.edu/~15780>

All (virtual) interaction with the course staff will happen via Piazza:

<http://www.piazza.com>

Lecture videos will be available 2-3 hours after class on Blackboard (first lecture may be a bit delayed due to course creation time), individual links also available on class website, but possibly later:

<http://blackboard.andrew.cmu.edu>

Grading

Grading breakdown for the course:

50% homeworks (6% HW0 + 11% HW1-4)

25% final project

15% midterm

10% class participation

Final grades will be assigned on a curve (for which we don't know the thresholds), but they are guaranteed to be *lower* than the standard A = 90-100, B=80-90, etc

Homeworks

There will be four homeworks throughout the course, plus an initial “Homework 0” released today (more on this shortly)

Homeworks each contain ~2 theory/derivation questions and ~2 programming questions

All submission done via Autolab (including writeups of written portions), programming portions are auto-graded

<http://autolab.andrew.cmu.edu>

5 late days to use throughout semester, max of 2 late days for each assignment

Class project

A chance to explore an applied, theoretical, or algorithm aspect of AI in more detail

To be done in groups of 2-3

Project will require a proposal (300 words), a mid-way checkin (3 pages), and a final report (5 pages)

Video session during the time of the class final exam (attendance of all group members is required)

Full details to be posted to class webpage

Midterm

In-class midterm to be held on **3/1**

Will cover topics in course up to and including the lecture right before the midterm

Midterm will be **closed book, closed notes** (mainly for space reasons)

Class participation

Your participation grade comes through your participation in in-class polls posted to Piazza during lecture

Homework for *today*: register for the class on Piazza, find the poll below, and fill out the answer (we reserve the right to drop from the course any student who has not completed this poll within 24 hours)

Poll: which letter is the best letter?

- A.
- B.
- C.
- D.

Instructors and TAs



Zico Kolter



Ariel Procaccia



Brandon Amos



Shayan Doroudi



Anson Kahng



Kijung Shin

Recommended background and HW0

Students taking this course should have experience with: mathematical proofs, linear algebra, calculus, probability, Python programming

We aren't listing specific pre-req courses (because people get this experience from different sources), but these are **required** prerequisites

Today we are releasing HW0, with one theory question and one programming question, meant as a basic test of (some of) these skills; if it seems particularly hard, the later questions will probably be *very* difficult

HW0 is due a week from Friday (1/27), *before* the add deadline

Waitlist

As of right now, 93 students enrolled, 91 on waitlist

We can't guarantee ... but we guarantee that there will be open slots in the class by the middle of February, so if you stick with it, you will be added to the course

Students will be taken off the waitlist in the order that they submit a full-credit solution for HW0

Academic integrity

Homework policy:

You may discuss homework problems with other students, but you need to specify all students you discuss with in your writeup

Your writeup and code must be written *entirely* on your own, without reference to notes that you took during any group discussion

All code and written material that you submit must be entirely your own unless specifically cited (in quotes for text, or within a comment block for code) from third party sources

See the CMU policy on academic integrity for general information

<https://www.cmu.edu/academic-integrity/>

Student well-being

CMU and courses like this one are stressful environments

In our experience, most academic integrity violations are the product of these environments and decisions made out of desperation

Please don't let it get to this point (or potentially much worse)

Don't sacrifice quality of life for this course: still make time to sleep, eat well, exercise

Some parting thoughts

“Computers in the future may have only 1,000 vacuum tubes and weigh only 1.5 tons.”
– Popular Mechanics, 1949

“Machines will be capable, within twenty years, of doing any work a man can do.”
– Herbert Simon, 1965