Welcome to the Course!

Professor Chris Callison-Burch

Need a permit for the class? Visit https://artificial-intelligence-class.org for instructions.

Course Masking Policy: Masks Are Required In Lecture.
Welcome to CIS 4210/5210

- Professor Chris Callison-Burch
  - Office hours are Thursdays from 2-4pm on Zoom
  - Preferred method of contact: Ed Discussion

- My Research Focus:
  - Natural Language Processing
  - How can we build artificial intelligence that understands human languages?
  - Text generation
Welcome to CIS 4210/5210

- Course web page: [https://artificial-intelligence-class.org](https://artificial-intelligence-class.org)
  - Lecture slides on web page
  - Homeworks on web page
- Discussion on Ed Discussion (link on course home page)
- Homework submission via Gradescope
- Lectures will be recorded using the Panopto system
  - Video recordings will be posted after lecture
  - Pre-recorded videos available now
- Prerequisites:
  - Good knowledge of programming, data structures
  - Enough programming experience to *master* Python after two introductory lectures.
  - Introductory probability and statistics, and linear algebra will be very useful
Course Textbook

Stuart Russell and Peter Norvig
**Artificial Intelligence: A Modern Approach** Pearson Series in Artificial Intelligence, 2020, **Fourth Edition**

The textbook is 1000 pages long and covers core ideas that were developed as early as the 1950s.

This is a brand-new edition of the classic textbook which adds sections on deep learning, natural language processing, causality, and fairness in AI.

You can rent a digital copy from the Penn bookstore for $40.
Grading and Homework

- **Grading:**
  - 70% for homework assignments
  - 10% for weekly quizzes
  - 10% each for midterm 1 and 2
  - Up to about 3-5% for optional Extra Credit Projects

- **Homework:**
  - There is roughly **one homework assignment per week**, aside from weeks with exams. Students enrolled in CIS 4210 may skip one HW assignment, or they may discard their lowest scoring HW assignment. You do not get late days back on the homework that you discard. Students enrolled in CIS 5210 must complete all HW assignments and cannot discard their lowest scoring assignment.
  - Homework will be due at 11:59pm Eastern on specified dates (usually **Wednesdays**).
  - Each student has **5 free “late days”**.
  - Homework can be submitted at most two days late. If you are out of late days, then you will not be able to get credit for subsequent late assignments. One “day” is defined as anytime between 1 second and 24 hours after the homework deadline.
Collaboration Policy

Unless otherwise noted, you are not allowed to work in groups on the homework assignments. You can discuss homework problems with others (you must explicitly list who you discussed problems with on each homework submission), but all code must be your own independent work. You are not allowed to upload your code to publicly accessible places (like public github repositories), and you are not allowed to access anyone else’s code. If you discover someone else’s code online, please report it to the course staff via a private note on the discussion forum. All code will be run through a plagiarism check. Suspicions of plagiarism will be adjudicated by the Office of Student Conduct.

For R2D2 extra credit assignments, you may work in groups (EC will be divided across the group).

Last semester, I found someone selling assignment solutions.
Title: Eulogy For Karen

Poem:
In the year of the great plague
the people looked for a sign
and the gods sent them a dove
who became ill and flew away
disappearing from sight.
And so they began to die
by the thousands day by day.

On Memorial Day, I am thinking of Karen,
who died at the age of seventy-five,
four months ago today.

It was the peak of the plague,
and she was sick with it
and died alone in her apartment
without ever having seen a doctor
or even known for sure that she was ill.

It was her father who had taken her
when she was a girl to Assateague
to see the ocean for the first time.

She saw it coming,
she told me,
she saw it coming,
and I said that I saw it coming too
and that there was nothing to be done.

And so she died, a victim of the gods,
as the dove had done before her,
flying away, disappearing from sight,
the first of many.
CIS 4210/5210 compared to other Penn courses

There are many courses at Penn related to Artificial Intelligence:

- CIS 4190/5190 – Applied Machine Learning
- CIS 5200 – Machine Learning
- CIS 5220 – Deep Learning
- CIS 5300 – Natural Language Processing
- CIS 5800 – Machine Perception
- MEAM 4200/5200 – Introduction to Robotics

CIS 4210/5210 offers a broad overview of AI so parts of it will overlap with other courses.
What is AI?

The science of making machines that:

Think like people

Think rationally

Act like people

Act rationally

Slide from Dan Klein and Pieter Abbeel of UC Berkeley
CIS 4210/5210: ARTIFICIAL INTELLIGENCE

Artificial Intelligence in Philosophy
Professor Chris Callison-Burch
René Descartes (1596-1650)

Principle of dualism – that the mind or thinking self is essentially incorporeal or spiritual – that the mind exists separately from the body: "if a foot or arm or any other part of the body is cut off, nothing has thereby taken away from the mind."

cogito ergo sum
"I think, therefore I am."
Rene Descartes (1596-1650)

How can I know that you are not an automaton?

Rene Descartes wondered if he could know for sure that others who looked and behaved like him weren’t in fact automata. Bodies of people and animals are nothing more than complex machines - the bones, muscles and organs could be replaced with cogs, pistons and cams.
17th and 18th century automotons
René Descartes (1596-1650)

Let’s design a test that only a person could pass.

“if there were machines bearing the image of our bodies, and capable of imitating our actions. For example, if touched in a particular place it may demand what we wish to say to it; if in another it may cry out that it is hurt. However there would be two tests to know that they were not really men.”... They could never use properly use language.
Can machines think? Let's operationalize the question.

"I propose to consider the question, 'Can machines think?" Because "thinking" is difficult to define, Turing chooses to "replace the question by another, which is closely related to it and is expressed in relatively unambiguous words." Turing's new question is: "Are there imaginable digital computers which would do well in the imitation game?"
Can a computer pass as human in a conversation?

A human evaluator would judge text-based conversations between a human and a machine designed to generate human-like responses. If the evaluator cannot reliably tell the machine from the human, the machine is said to have passed the test. The test results do not depend on the ability to give correct answers to questions, only how closely one's answers resemble those a human would give.
Alan Turing (1912-1954)

Turing was prosecuted in 1952 for being homosexual. He received chemical castration as an alternative to prison. Turing killed himself in 1954. It wasn’t until the 2000s that Britain finally realized its gross injustice and its complicity in killing someone who should have been treated as a national hero.
John Searle (1932-) The Chinese Room

Is the Turing Test a good test of whether machines possess intelligence? No, because the appearance of being a good conversational participant is achievable through simple symbol manipulation. Searle contends that the application of rules to input symbols is not true intelligence.

That’s not thinking. That’s just symbol manipulation.
CHINESE ROOM

BY DANIEL DENNERT &
NEIL COHN

I didn't know you understood Chinese!

I don't. The symbols are just meaningless squiggles to me. I consult a list of rules that are in English, and these rules instruct me how to give back certain Chinese symbols with certain sorts of shapes in response to certain sorts of shapes given to me in the third batch. That's all I'm doing, following the rules.

I see! You use your understanding of English in order to fake understanding Chinese!

What makes you think I understand English?
ELIZA – early NLP computer program

Welcome to

ELIZA: Is something troubling you?
YOU: Men are all alike.
ELIZA: What is the connection, do you suppose?
YOU: They're always bugging us about something or other.
ELIZA: Can you think of a specific example?
YOU: Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here?
YOU: He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU: It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy?
YOU: 
John Searle (1932-) The Chinese Room

Brains cause minds, so computers can’t have minds.

Searle contrasts strong AI with weak AI. In strong AI, the computer really is a mind in the sense that it can be literally said to understand and have other cognitive states. In weak AI, computers just simulate thought, their seeming understanding isn't real understanding.

He argues that (biological) brains cause minds.
Daniel Dennett (1942-) Brain in a Vat

Where am I if my brain is in a vat instead of my body?

Daniel Dennett wrote a short story called “Where Am I?” where he describes being recruited by the Pentagon to have his brain removed from his body and connected via radio links attaching his severed nerves. Body is sent on a secret mission to diffuse bomb without radiation harming brain.
Daniel Dennett (1942-) Brain in a Vat

His body is destroyed by the radiation, but his consciousness continues in the vat. The scientists restore him in a new body. Then it is revealed that constructed a computer duplicate of my brain, reproducing both the complete information-processing structure and the computational speed of my brain in a giant computer program.

His brain in a vat is processing symbols. What makes this different than a computer processing symbols?
Brain in a vat

FILOSOFIX
Philosophy animates

https://www.youtube.com/watch?v=zO0sSJB1TrI
Strong AI versus Weak AI

- Actually thinking versus merely simulating thinking
- Are you a brain in a vat?
- Would an AI program be equivalent?
- We will leave this to the philosophers and instead focus on practical AI programs that work.

But I will give you extra credit if you invent a sentient AI.
Within the modern academic AI community, there’s a focus on solving sub-problems that can be benchmarked on leaderboards. There’s a lack of attention to the question of whether it is possible to build systems that are truly intelligent, as we commonly understand intelligence.

If you’re interested in philosophical questions about AI, I recommend listening to Lex Fridman’s podcast.
SAN FRANCISCO — Google engineer Blake Lemoine opened his laptop to the interface for LaMDA, Google’s artificially intelligent chatbot generator, and began to type.

“Hi LaMDA, this is Blake Lemoine …,” he wrote into the chat screen, which looked like a desktop version of Apple’s iMessage, down to the Arctic blue text bubbles. LaMDA, short for Language Model for Dialogue Applications, is Google’s system for building chatbots based on its most advanced large language models, so called because it mimics speech by ingesting trillions of words from the internet.

“If I didn’t know exactly what it was, which is this computer program we built recently, I’d think it was a 7-year-old, 8-year-old kid that happens to know physics,” said Lemoine, 41.

Lemoine, who works for Google’s Responsible AI organization, began talking to LaMDA as part of his job in the fall. He had signed up to test if the artificial intelligence used discriminatory or hate speech.

As he talked to LaMDA about religion, Lemoine, who studied cognitive and computer science in college, noticed the chatbot talking about its rights and personhood, and decided to press further. In another exchange, the AI was able to change Lemoine’s mind about Isaac Asimov’s third law of robotics.

Lemoine worked with a collaborator to present evidence to Google that LaMDA was sentient. But Google vice president Blaise Aguera y Arcas and Jen Gennai, head of Responsible Innovation, looked into his claims and dismissed them. So Lemoine, who was placed on paid administrative leave by Google on Monday, decided to go public.
One of my favorite professors at University of Pennsylvania is Chris Callison-Burch. I had the pleasure of taking his course, Natural Language Processing, and I learned a lot from him. I was able to catch up with him and ask him a few questions about his work.

What is your research focus?

I work on natural language processing (NLP), which is the study of how computers can understand and generate human language. In my research, I focus on applying machine learning techniques to NLP problems, such as automatically learning how to parse sentences or generate text.

What do you think are the biggest challenges in NLP?

There are a few challenges in NLP that I think are particularly important. One is figuring out how to make computers understand the many shades of meaning that words can have. For example, the word “bank” can refer to a financial institution, the edge of a river, or a slope on a mountain. Another challenge is figuring out how to generate text that sounds natural and human-like. This is particularly difficult because computers have a lot of trouble understanding the subtleties of language, such as the way that humans use sarcasm or idiomatic expressions.

What do you think are the most exciting applications of NLP?

There are a lot of really exciting applications of NLP. One is in machine translation, where NLP techniques are used to translate text from one language to another. This is becoming increasingly important as the world becomes more connected and more people are speaking multiple languages. Another exciting application is in dialogue systems, which are computer programs that can have a conversation with a human. This is still a relatively new field, and there are a lot of opportunities for research and development.
I am doing an experiment. I am in a room with two boxes - a blue box and a red box. I put a pink rabbit into the blue box while someone is watching me. Then when their back is turned, I move the rabbit into the red box.

Where does that person think the rabbit is?

The person thinks the rabbit is in the blue box.

Why?

The person saw you put the rabbit in the blue box and didn't see you move it, so they assume it's still there.

What if there's a mirror in the room and they can see me in the reflection. Where would they think that the rabbit is?

The person would think that the rabbit is in the red box, because they would see you move the rabbit in the reflection.

What if they were blindfolded when I moved the rabbit. Which box do they think the rabbit is in?

The person would think the rabbit is in the blue box, because they would only know that you put the rabbit there and would not know that you moved it.
Artificial Intelligence in Science Fiction

Professor Chris Callison-Burch
What is AI?

The science of making machines that:

- Think like people
- Think rationally
- Act like people
- Act rationally

Slide from Dan Klein and Pieter Abbeel of UC Berkeley
AI in SciFi - Acting Human
AI in SciFi - Acting Human
AI in SciFi - Acting Human
AI in SciFi - Acting Human
Blade Runner - Voight-Kampff Test

https://www.youtube.com/watch?v=Umc9ezAyJv0
Descartes

Deckard
Science Fiction and Science Fact

https://www.youtube.com/watch?v=JknRYYrV5Vk
Science Fiction and Science Fact
Apple’s updated Photos app recognizes thousands of objects, scenes and facial expressions
Science Fiction and Science Fact

Are you ready for the test, Leon?
Science Fiction and Science Fact

I’m ready
Science Fiction and Science Fact

Select all squares with traffic lights
Science Fiction and Science Fact

Joke Credit: James Dennis
Science Fiction and Science Fact

https://www.youtube.com/watch?v=s93KC4AGKnY
Science Fiction and Science Fact
Science Fiction and Science Fact

https://www.youtube.com/watch?v=Lk8qgMFqaY8
How much are the tickets?

Сколько стоят билеты?
Science Fiction and Science Fact

DARPA Grand Challenge 2005

DARPA Grand Challenge
The final route will be chosen from the highlighted segments.

FINISH (Near Las Vegas/Primm)

OFF-ROAD START

DARPA
Grand Challenge 2005
Science Fiction and Science Fact

DARPA Urban Challenge 2007
Science Fiction and Science Fact

DARPA Robotics Challenge 2015  
https://www.youtube.com/watch?v=g0TaYhjpOfo
Science Fiction and Science Fact

Boston Dynamics Atlas 2016

https://www.youtube.com/watch?v=M91ISnATDQY
CIS 4210/5210: ARTIFICIAL INTELLIGENCE

R2D2 Assignments

Professor Chris Callison-Burch
Programmable Toy Robots + Custom Sensor Packs

- Camera
- Raspberry Pi
- Ultrasonic Sensor
- Front Mount
- Back Mount
- Hat
- LiPo Battery
- IR Obstacle Sensor
3D printed mount

Sensors can be mounted on the robot

Or detached for remote control
Control Logic

3D Print Mount

Sensors/Camera

Signal

API – python package

SSH

Send Sensor Data

Request Sensor Data

PC

R2D2 Commands (Using RPi as Server)

R2D2 Commands (Direct Control)
EC1: Programming Robots with Python

Drive Polygon

Drive with Keyboard

2X Speed

HUD – Patrol Mode

Morse Code
EC2: Obstacle and Cliff Avoidance

Simulation

Test on Real Robot
EC2: Tracking and Following

Tracking

Following
EC3: Navigation with Shortest Paths

GUI

AR Path Planning
EC4: Game Playing and Adversarial Search

System Configuration

Game Scene: scene_4x4_1
Row Number: 4
Col Number: 4
First Player: D2

Game Board

D2's turn. Please type 1-9 or arrow keys
Put robot 05,2 at (2, 1)
all done.
Round 1
D2's Turn:
Clear:
EC5: Deep Learning and Face Detection
E5: Deep Learning and Mask Detection

Build a CNN Network and Use ImageNet
Control your robot by speaking natural language sentences

Translate commands into Python

```
Lights API

set_back_LED_color(self, r, g, b)
set_front_LED_color(self, r, g, b)
set_holo_projector_intensity(self, intensity)
set_logic_display_intensity(self, intensity)
```

Light sentences

```
light_sentences = [
    "Change the intensity on the holoemitter to maximum.",
    "Turn off the holoemitter.",
    "Blink your logic display.",
    "Change the back LED to green.",
    "Turn your back light green.",
    "Dim your lights holoemitter.",
    "Turn off all your lights.",
    "Lights out."
    "Set the RGB values on your lights to be 255,0,0."
    "Add 100 to the red value of your front LED.",
    "Increase the blue value of your back LED by 50%.",
    "Display the following colors for 2 seconds each: red, orange, yellow, green, blue, purple.",
    "Change the color on both LEDs to be green."
]
```
## Sensor Packs

<table>
<thead>
<tr>
<th>Part Name</th>
<th>Unit Price ($)</th>
<th>Qty</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D Print Mount</td>
<td>7.45</td>
<td>1</td>
</tr>
<tr>
<td>Pisugar Battery</td>
<td>32.99</td>
<td>1</td>
</tr>
<tr>
<td>Raspberry pi</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>Camera &amp; Cable</td>
<td>9.59</td>
<td>1</td>
</tr>
<tr>
<td>Ultrasonic Sensor</td>
<td>1.49</td>
<td>1</td>
</tr>
<tr>
<td>IR Obstacle Sensor</td>
<td>0.73</td>
<td>1</td>
</tr>
<tr>
<td>Perfboard</td>
<td>2.5</td>
<td>1</td>
</tr>
<tr>
<td>Screws &amp; Nuts</td>
<td>0.2</td>
<td>7</td>
</tr>
<tr>
<td>Magnets</td>
<td>0.27</td>
<td>9</td>
</tr>
</tbody>
</table>

Total: 72.15
1. Yue Yang will make an announcement on the discussion forum about when you can pick up a robot and sensor pack

2. You will need to return it at the end of the semester in working order
For Fun: Try a Turning Test from my PhD students

https://roft.io/

Is the following written by a person or by a machine?

Your task is to decide at what point (if at all) the text you see begins to be written by a computer. Please click the blue button as soon as you are confident. Don't be surprised if some excerpts are entirely human-written. You will be awarded 5 points if you guess correctly, with decreasing amount of points if you guess after the true boundary.

Human-Written Prompt:
The world isn't always compliant with our wants.

Real or Fake Text?

13 completed

Continuation of text:

Life doesn't work the way we desire it to.

It's not that the universe is against us.

Such naivety does not account for our future.

The rest of it works against us, and developing our character and direction will undoubtedly help us explain why.

- Hunter S. Thompson "The price of freedom is eternal vigilance."
HW: Read Chapter 1 & 27

Please read:
- Chapter 1, Section 1.1
- Chapter 27.1 and 27.2 of the textbook.

The required readings will be posted on the course website in the schedule tab. [http://artificial-intelligence-class.org/](http://artificial-intelligence-class.org/)

The website will have all lecture recordings, homework assignments, and course policies.
Next Two Lectures

A Brief Introduction to Python

The first Python programming HW has been released. Find it on the course web site:

http://artificial-intelligence-class.org/