Welcome to CIS 521

- Lecturer: Mitch Marcus, mitch@cis.upenn.edu
  - Course home page: http://www.cis.upenn.edu/~cis521
- Discussion via Piazza (link on course home page)
- Prerequisites: Good knowledge of programming, data structures. Introductory probability and statistics, logic. Ability to master Python after a couple of class lectures

Grading & Homework

Grading:
- 50% Homeworks
- 25% Midterm 1
- 25% Midterm 2

Homework:
- Homework will be due at 11:59 on specified dates. Late homework only be accepted if an extension has been granted in advance.
- There will be about 50% more homeworks this year – 9 or 10 in all. Your lowest homework grade will be dropped, so extensions will be granted sparingly.
- We insist that students follow Penn's Code of Academic Integrity. Academic dishonesty, as defined in the Code of Integrity, will not be tolerated.

Singularity

- “One conversation centered on the ever accelerating progress of technology ... which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue – Stanislaw Ulam, about John von Neumann, May 1958

“I want to design a machine that will be proud of me – Danny Hillis”
With artificial intelligence we’re summoning the demon – Elon Musk

Full artificial intelligence could spell the end of the human race – Steven Hawking

A REAL Achievement : DARPA Grand Challenge 2005

Older Real Accomplishments of AI

Recent Significant Advances In NLP

- IBM’s Watson
  - Web-scale information extraction & question answering

- Apple’s Siri
  - Interactive Dialogue Systems

- Google Translate
  - Automatic Machine Translation

- 1991: AI Logistics Planning for Gulf War
- 1997: Deep Blue defeated the reigning world chess champion Garry Kasparov
- 1998: Deep Space 1 (launched) – Remote Agent Experiment

“Invisible” AI

- Computer Algebra Systems (Maple, Mathematica)
- Credit Evaluation, Fraud Detection
- Internet Search, Spam Filtering
- Handwritten character recognition (checks, US mail)
**What is AI?**

Views of AI fall into four categories:

<table>
<thead>
<tr>
<th>Thinking humanly</th>
<th>Thinking rationally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acting humanly</td>
<td>Acting rationally</td>
</tr>
</tbody>
</table>

*We will focus on "acting rationally"*

**Acting humanly:** Social robots

- Cynthia Breazeal: MIT

**Acting humanly:** Social robots

**Thinking humanly:**

- Cognitive modeling
  - 1960s "cognitive revolution": information-processing psychology, a.k.a. cognitive psychology
    - Requires scientific theories of internal activities of the brain
    - How to validate? Requires
      1. Predicting and testing behavior of human subjects
      2. Direct identification from neurological data (bottom-up)
        - Cognitive Neuroscience
    - Both approaches are now distinct from AI
    - Caveat: ACT-R & SOAR communities do computational modeling of high level mental functions

**Acting rationally:** Turing Test

- Turing (1950) "Computing machinery and intelligence": "Can machines think?"
- Operational test for intelligent behavior: the Imitation Game
  - Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
  - Anticipated most major arguments against AI
  - Suggested major components of AI: knowledge, reasoning, language understanding, learning
Thinking rationally: "laws of thought"
- Aristotle: what are correct arguments/thought processes?
- Several Greek schools developed various forms of logic: notation and rules of derivation for thoughts; may or may not have proceeded to the idea of mechanization
- Direct line through mathematics and philosophy to modern AI
- Problems:
  1. Not all intelligent behavior is mediated by logical deliberation
  2. What is the purpose of thinking? What thoughts should I have?
  3. Ignores the hard problem of perception
  4. Most logical inference is intractable

Acting rationally: rational agents
- Rational behavior: doing the right thing
- The right thing: that which is expected to maximize goal achievement, given the available information
- Doesn't necessarily involve thinking – e.g., blinking reflex – but thinking should be in the service of rational action

Rational agents
- An agent is an entity that perceives and acts
- This course is about effective programming techniques for designing rational agents
- Abstractly, an agent is a function from percept histories to actions:
  \( f: \mathcal{S} \rightarrow \mathcal{A} \)
- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- Caveat: computational limitations make perfect rationality unachievable
  \( \Rightarrow \) design best program for given machine resources

Two Approaches to AI
- Symbolic representations of the world
  - Relations between entities
    - "Mitch’s bicycle is red"
    - (isa B3241 bicycle) (color B3231 red) (owns B3241 P119)
    - (isa P119 person) (name P119 “Mitch”)
  - Explicit logical models
  - Logical inference, Search
  - Chess, Sudoku, computer games, …
- Statistical models
  - Prediction by look-up or by weighted combinations
    - \( P(y=bicycle) = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_3 + \ldots \)
  - Machine Learning, Machine vision, speech recognition, …

Course Overview – First Half
Module 0: Introduction
- Intelligent Agents
- Python Programming
Module 1: Search Strategies
- Uninformed & Informed Search (Homework: Robot search)
- Constraint Satisfaction (Homework: Sudoku Solver)
- Adversarial Search (Homework: Game Playing Bot)

Course Overview – Second Half
Module 2: Machine Learning and Natural Language Processing
- Review of Probability
- Naive Bayes (Spam Filtering) & Bayesian Networks
  (Homework: Build a spam filter)
- Perceptrons and Support Vector Machines
- Hidden Markov Models & Part of Speech Tagging
  (Homework: Generate fake Frankenstein text, Build Part of Speech Tagger)
Module 3: Knowledge Representation and Logic
- Logical Agents (Homework if time: Solve Logic Puzzles)
- Topics in Knowledge Representation
- The Singularity: A critique