CS383: Artificial Intelligence

Introduction

Prof. Scott Niekum
University of Massachusetts Amherst

[Based on slides created by Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley. All materials available at http://ai.berkeley.edu.]
A bit about me

SCALAR LAB
Safe, Confident, and Aligned Learning + Robotics

AI Safety
Reinforcement learning
Robotic manipulation and learning from demonstration
Course Information

Communication:
- Announcements on webpage/email
- Grades on Gradescope
- Piazza for discussion

Assignments:
- Gradescope for interactive homework (unlimited submissions!)
- Autograded programming projects (submit via Gradescope)
- Make sure you have access to a system where you can run Python

(or Google “Scott Niekum” and go to the Teaching tab)

TAs: Sylee Dandekar (sdandekar@umass.edu); Shuwa Miura (smiura@umass.edu); Aleksandra Burushkina (aburushkina@umass.edu);
Matthew Weiner (mweiner@umass.edu)
TA Office Hours: Monday 11-12; Tuesday 9:30-10:30; Wednesday 2-4; Thursday 1-3
Workload

- There will be a lot of math (and programming)

- Reading assignments

- 8 homework assignments:
  - ~2 weeks for each, but overlapping
  - Online, autograded, solve and submit alone
  - No late submissions accepted

- 6 programming projects
  - Python, groups of 1 or 2 (except Project 0)
  - ~2 weeks for each, non-overlapping
  - 10 late days for semester, no other extensions granted

- One midterm, one final
Warning: Not a course textbook, so our presentation does not necessarily follow the presentation in the book.


• I’ll also post class slides
Homework Exercises

• Online on Gradescope
• Autograded text boxes / multiple choice
• Try as many times as you want!
• Goal: self-assess and prepare for tests
• Can discuss at high-level, but work alone
• No spoilers on Piazza discussions!
Programming Assignments

Pacman domain

Projects include:

• path planning and search
• multi-agent game trees
• reinforcement learning
• state estimation
• classification

Highly suggested: Pair programming
(switch “driver” and “observer” roles often)
Midterm and Final

• Midterm will cover roughly half the class material
• Final will be comprehensive
• Midterm in-class, Final during finals week
• Challenging!
Grading

Plus/minus grading - adjustable scale, but no more harsh than:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[94-100)</td>
</tr>
<tr>
<td>A-</td>
<td>[90-94)</td>
</tr>
<tr>
<td>B+</td>
<td>[87-90)</td>
</tr>
<tr>
<td>B</td>
<td>[84-87)</td>
</tr>
<tr>
<td>B-</td>
<td>[80-84)</td>
</tr>
<tr>
<td>C+</td>
<td>[77-80)</td>
</tr>
<tr>
<td>C</td>
<td>[74-77)</td>
</tr>
<tr>
<td>C-</td>
<td>[70-74)</td>
</tr>
<tr>
<td>D+</td>
<td>[67-70)</td>
</tr>
<tr>
<td>D</td>
<td>[64-67)</td>
</tr>
<tr>
<td>D-</td>
<td>[60-64)</td>
</tr>
<tr>
<td>F</td>
<td>[0-60)</td>
</tr>
</tbody>
</table>

Grades will be weighted as follows:

- Homework Exercises (20%)
- Programming Assignments (30%)
- Midterm (20%)
- Final (30%)
READ THE STATEMENT IN THE SYLLABUS

- Discuss concepts, but don’t share solutions or written work with other students
- Don’t look for answers / code online or elsewhere
- Automated tools will be used to discover cheating
- If unsure, check departmental guidelines or ask — ignorance is not an excuse
- We will pursue the harshest penalties available, so please don’t cheat!
- To be clear: you will fail the class automatically and be reported to the university
**Important This Week**

- **Important this week:**
  - **Make sure** that you can get into Gradescope — create an account and use the course code that you were emailed (and which is also posted on Piazza).
  - **Be sure** that you have a computer where you can run Python
  - **P0: Python tutorial** is out (due Mon 2/12 at 11:59 pm via Gradescope)

- **Also important:**
  - **If you are wait-listed,** you might or might not get in depending on how many students drop. Be patient if possible — many students often drop early in the course.
  - **Office Hours** begin Monday
Today

- What is artificial intelligence?
- What can AI do?
- What is this course?
Sci-Fi AI?
AI in the news

ELON MUSK ANNOUNCES OPEN AI: A NON-PROFIT TO HELP PREVENT SKYNET
A definition for AI
“Artificial Intelligence (AI) is a science and a set of computational technologies that are inspired by — but typically operate quite differently from — the ways people use their nervous systems and bodies to sense, learn, reason, and take action.”
Philosophical questions

- AI is one of the great intellectual adventures of the 20th and 21st centuries.

  - What is a mind?
  - How can a computer have a mind?
  - Can we build a mind?
  - Can trying to build one teach us what a mind is?
What is AI?

The science of making machines that:

- Think like people
- Act like people
- Think rationally
- Act rationally
What is AI?

The science of making machines that:

Think like people

Act like people

Think rationally

Act rationally
Thinking Like Humans?

- The cognitive science approach:
  - 1960s ``cognitive revolution'': information-processing psychology replaced prevailing orthodoxy of behaviorism (reflexive behaviors, classical conditioning, etc.)
- Scientific theories of internal activities of the brain
  - What level of abstraction? “Knowledge" or “circuits”?
  - **Cognitive science**: Predicting and testing behavior of human subjects (top-down)
  - **Cognitive neuroscience**: Direct identification from neurological data (bottom-up)
- Both approaches now distinct from AI
- Both share with AI the following characteristic: *The available theories do not explain (or engender) anything resembling human-level general intelligence*
What is AI?

The science of making machines that:

Think like people
What is AI?

The science of making machines that:

- Think like people
- Act like people
Acting Like Humans?

- Turing (1950) “Computing machinery and intelligence”
  - “Can machines think?” → “Can machines behave intelligently?”
  - Operational test for intelligent behavior: the *Imitation Game*

- Predicted by 2000, a 30% chance of fooling a lay person for 5 minutes
- Anticipated all major arguments against AI in following 50 years
- Suggested major components of AI: knowledge, reasoning, language understanding, learning

- Problem: Does the Turing test really measure what we want?
What is AI?

The science of making machines that:

Think like people

Act like people
What is AI?

The science of making machines that:

Think like people

Think rationally

Act like people
Thinking Rationally?

- **The “Laws of Thought” approach**
  - What does it mean to “think rationally”?
  - Normative / prescriptive rather than descriptive

- **Logicist tradition:**
  - Logic: notation and rules of derivation for thoughts
  - Aristotle: what are correct arguments/thought processes?
  - Direct line through mathematics, philosophy, to modern AI

- **Problems:**
  - Not all intelligent behavior is mediated by logical deliberation
  - What is the purpose of thinking? What thoughts should I (bother to) have?
  - Logical systems tend to do the wrong thing in the presence of **uncertainty**
  - Why should we care about thought at all, when **action** is what matters?
What is AI?

The science of making machines that:

- Think like people
- Act like people
- Think rationally
- Act rationally
What is AI?

The science of making machines that:

Think like people

Think rationally

Act like people

Act rationally
Acting Rationally

- **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
  - Thinking can be in the service of rational action
  - Entirely dependent on goals!
  - Irrational ≠ insane, irrationality is sub-optimal action
  - Rational ≠ successful

- **Our focus here: rational agents**
  - Systems which make the best possible decisions given goals, evidence, and constraints
  - In the real world, usually lots of uncertainty
  - ... and lots of complexity
  - Usually, we’re just approximating rationality
Rational Decisions

We’ll use the term *rational* in a very specific, technical way:

- Rational: maximally achieving pre-defined goals
- Rationality only concerns what decisions are made (not the thought process behind them)
- Goals are expressed in terms of the *utility* of outcomes
- Being rational means maximizing your *expected utility*

A better title for this course would be:

**Computational Rationality**
Maximize Your Expected Utility
A (Short) History of AI
A (Short) History of AI

- **1940-1950: Early days**
  - 1943: McCulloch & Pitts: Boolean circuit model of brain
  - 1950: Turing’s “Computing Machinery and Intelligence”

- **1950—70: Excitement: Look, Ma, no hands!**
  - 1950s: Early AI programs, including Samuel’s checkers program, Newell & Simon’s Logic Theorist, Gelernter’s Geometry Engine
  - 1956: Dartmouth meeting: “Artificial Intelligence” adopted
  - 1965: Robinson’s complete algorithm for logical reasoning

- **1970—90: Knowledge-based approaches**
  - 1969—79: Early development of knowledge-based systems
  - 1980—88: Expert systems industry booms

- **1990—: Statistical approaches**
  - Resurgence of probability, focus on uncertainty
  - General increase in technical depth
  - Agents and learning systems... “AI Spring”?  

- **2000—: Where are we now?**
What Can AI Do?

Quiz: Which of the following can be done at present?

- ✓ Play a decent game of table tennis?
- ✓ Play a decent game of Jeopardy?
- ✓ Drive safely along a curving mountain road?
- ? Drive safely along a busy street on a Friday night?
- ✓ Buy a week's worth of groceries on the web?
- ✓ Buy a week's worth of groceries at the store?
- ✗ Discover and prove a new mathematical theorem?
- ✗ Converse successfully with another person for an hour?
- ✗ Perform a surgical operation?
- ✓ Put away the dishes and fold the laundry?
- ✓ Translate spoken Chinese into spoken English in real time?
- ✗ Write an intentionally funny story?
Natural Language

- **Speech technologies (e.g. Siri)**
  - Automatic speech recognition (ASR)
  - Text-to-speech synthesis (TTS)
  - Dialog systems

- **Language processing technologies**
  - ChatGPT
  - Machine translation

- Web search
- Text classification, spam filtering, etc…
Deep learning
The Great A.I. Awakening

How Google used artificial intelligence to transform Google Translate, one of its more popular services—and how machine learning is poised to reinvent computing itself.

BY GIDION LEWIS-KRAUS  DEC. 14, 2016
Vision (Perception)

- Object and face recognition
- Scene segmentation
- Image classification

Images from Erik Sudderth (left), wikipedia (right)
Object Tracking

object detection / 3D pose estimation

arbitrary view rendered with estimated 3D pose
Perception + Natural Language

What vegetable is on the plate?
Neural Net: broccoli
Ground Truth: broccoli

What color are the shoes on the person's feet?
Neural Net: brown
Ground Truth: brown

How many school busses are there?
Neural Net: 2
Ground Truth: 2

What sport is this?
Neural Net: baseball
Ground Truth: baseball

What is on top of the refrigerator?
Neural Net: magnets
Ground Truth: cereal

What uniform is she wearing?
Neural Net: shorts
Ground Truth: girl scout

What is the table number?
Neural Net: 4
Ground Truth: 40

What are people sitting under in the back?
Neural Net: bench
Ground Truth: tent
We won’t discuss NLP and perception directly, but we will cover:

- Bayes nets
- Supervised learning
- Deep learning
Robotics

- Robotics
  - Part mech. eng.
  - Part AI
  - Reality much harder than simulations!

- Technologies
  - Vehicles
  - Rescue
  - Soccer!
  - Lots of automation...

- In this class:
  - We ignore mechanical aspects
  - Methods for planning
  - Methods for control

Images from UC Berkeley, Boston Dynamics, RoboCup, Google
Robot Soccer
Learning from demonstration
Learning from demonstration
Full body control of humanoids
We will cover several topics relevant to robotics:

- Planning and search
- Reinforcement learning
- Time-series analysis
- State estimation and filtering
Logic

- **Logical systems**
  - Theorem provers
  - NASA fault diagnosis
  - Question answering

- **Methods:**
  - Deduction systems
  - Constraint satisfaction
  - Satisfiability solvers (huge advances!)
Game Playing

- **Classic Moment: May, '97: Deep Blue vs. Kasparov**
  - First match won against world champion
  - “Intelligent creative” play
  - 200 million board positions per second
  - Humans understood 99.9 of Deep Blue's moves
  - Can do about the same now with a PC cluster

- **Open question:**
  - How does human cognition deal with the search space explosion of chess?
  - Or: how can humans compete with computers at all??

- **1996: Kasparov Beats Deep Blue**
  “I could feel --- I could smell --- a new kind of intelligence across the table.”

- **1997: Deep Blue Beats Kasparov**
  “Deep Blue hasn't proven anything.”

- Huge game-playing advances recently, e.g. in Go!

Text from Bart Selman, image from IBM's Deep Blue pages
Decision Making

- Applied AI involves many kinds of automation
  - Scheduling, e.g. airline routing, military
  - Route planning, e.g. Google maps
  - Medical diagnosis
  - Web search engines
  - Spam classifiers
  - Automated help desks
  - Fraud detection
  - Product recommendations
  - ... Lots more!
Course Topics

- **Part I: Making Decisions**
  - Fast search / planning
  - Constraint satisfaction
  - Adversarial and uncertain search
  - MDPs and Reinforcement learning

- **Part II: Reasoning under Uncertainty**
  - Bayes nets
  - Decision theory and value of information
  - Statistical Machine learning

- Throughout: Applications, Ethics, and Societal impacts