CS 188: Artificial Intelligence

Introduction

Instructors: Stuart Russell and Dawn Song
# Course Staff

## Professors

<table>
<thead>
<tr>
<th>Dawn Song</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stuart Russell</td>
</tr>
</tbody>
</table>

## GSIs

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jeffrey Tao</td>
<td>Jocelyn Chen</td>
<td>Jonathan Yang</td>
<td>Jasmine Collins</td>
<td>Nitish Dashora</td>
<td>Rishi Parikh</td>
<td></td>
</tr>
<tr>
<td>Evgeny P.</td>
<td>Ayush Kamat</td>
<td>Raiyymbek A.</td>
<td>Anikait Singh</td>
<td>Tarun Amarnath</td>
<td>Ajay Sridhar</td>
<td>Arvind Rajaraman</td>
</tr>
</tbody>
</table>
Course Information

- **Communication:**
  - Announcements, questions on Piazza
  - Staff email: [cs188@berkeley.edu](mailto:cs188@berkeley.edu)
  - Office hours:
    - Stuart: Monday 1.30-3, Tuesday 3.30-5*
    - Dawn: TBD
- **Sections start this week**
- **Work:**
  - Projects (25%), homework (10% + 10%)
    - P0 (Python) due 1/21, HW0 (math) due 1/26
  - Midterm (20%), final (35%)
  - Participation up to 5% extra (be nice!)
  - Fixed grading scale (85% A, 80% A-, etc.)

[http://inst.cs.berkeley.edu/~cs188](http://inst.cs.berkeley.edu/~cs188)

---

**Course Information**

**CS 188 | Spring 2022**

Lectures: Tu/Th 2:00-3:30 pm, Online

**Description**

This course will introduce the basic ideas and techniques underlying the design of intelligent computer systems. A specific emphasis will be on the statistical and decision-theoretic modeling paradigm.

By the end of the course, you will have built autonomous agents that efficiently make decisions in fully informed, partially observable and adversarial settings. Your agents will draw inferences in uncertain environments and optimize actions for arbitrary reward structures. Your machine learning algorithms will classify handwritten digits and photographs. The techniques you learn in this course apply to a wide variety of artificial intelligence problems and will serve as the foundation for further study in any application area you choose to pursue.

See the syllabus for slides, deadlines, and the lecture schedule. Readings refer to fourth edition of AIMA unless otherwise specified.

**Lecture reference materials**

We make lecture recordings available as links to Kaltura, which you can find posted together with other materials on the Syllabus page of this website shortly after the lecture. These links will work only if you are signed into your UC Berkeley Courses (Canvas) account.
Some Historical Statistics

- Homework and projects: instruction (iterate/learn till you nailed it)

- Exams: assessment
Textbook

Policies (see website)

- For online lectures:
  - Camera on, mic off
  - Please do ask questions: “Hand Up” or write in Chat

- We (staff) are here to help
  - Please do observe academic integrity policies!
  - Please don’t exclude your fellow students!
Today

- What is artificial intelligence?
- Where are we and how did we get here?
- How do we think about the design of AI systems?
Movie AI
Movie AI
AI is the biggest risk we face as a civilisation, Elon Musk says

Billionaire burn: Musk says Zuckerberg’s understanding of AI threat ‘is limited’

'Sociopathic' robots could overrun the human race in a generation

Computers should be trained to serve humans to reduce their threat to the human race, says a leading expert on artificial intelligence
United Kingdom Plans $1.3 Billion Artificial Intelligence Push
France to spend $1.8 billion on AI to compete with U.S., China
EU wants to invest £18 billion on AI development
China’s Got a Huge Artificial Intelligence Plan
IBM's Watson Jeopardy Computer Shuts Down Humans in Final Game

DAILY NEWS 9 March 2016

Sili
‘I’m in shock!’ Human world’s best human

Blizzard will show off Google's Deepmind AI in StarCraft 2 later this week

By Andy Chalk 4 hours ago

Google and Blizzard launched the artificial intelligence project in 2016.
A (Short) History of AI

Demo: HISTORY – MT1950.wmv
A short prehistory of AI

- **Prehistory:**
  - **Philosophy** (reasoning, planning, learning, science, automation)
  - **Aristotle:** For if every instrument could accomplish its own work, obeying or anticipating the will of others . . . if, in like manner, the shuttle would weave and the plectrum touch the lyre without a hand to guide them, chief workmen would not want servants, nor masters slaves
  - **Psychology** (learning, cognitive models)
  - **Linguistics** (grammars, formal representation of meaning)

- **Near miss (1842):**
  - Babbage design for universal machine
  - Lovelace: “a thinking machine” for “all subjects in the universe.”
“An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made if we work on it together for a summer.”

John McCarthy and Claude Shannon
Dartmouth Workshop Proposal
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: chess, checkers (RL), theorem proving
  - 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—2012: Statistical approaches + subfield expertise**
  - Resurgence of probability, focus on uncertainty
  - General increase in technical depth
  - Agents and learning systems... “AI Spring”?

- **2012—___: Excitement: Look, Ma, no hands again?**
  - Big data, big compute, deep learning
  - AI used in many industries
AI as Designing Rational Agents

- An **agent** is an entity that *perceives* and *acts*.
- A **rational agent** selects actions that maximize its expected **utility**.
- Characteristics of the **sensors, actuators, and environment** dictate techniques for selecting rational actions.

**This course** is about:
- General AI techniques for many problem types
- Learning to choose and apply the technique appropriate for each problem

Pac-Man is a registered trademark of Namco-Bandai Games, used here for educational purposes.
Agents and environments

- An agent *perceives* its environment through *sensors* and *acts* upon it through *actuators* (or *effectors*, depending on whom you ask)
- The *agent function* maps percept sequences to actions
- It is generated by an *agent program* running on a *machine*
A human agent in Pacman
The task environment - PEAS

- **Performance measure**
  - -1 per step; +10 food; +500 win; -500 die; +200 hit scared ghost

- **Environment**
  - Pacman dynamics (incl ghost behavior)

- **Actuators**
  - Left Right Up Down or NSEW

- **Sensors**
  - Entire state is visible (except power pellet duration)
PEAS: Automated taxi

- **Performance measure**
  - Income, happy customer, vehicle costs, fines, insurance premiums

- **Environment**
  - US streets, other drivers, customers, weather, police...

- **Actuators**
  - Steering, brake, gas, display/speaker

- **Sensors**
  - Camera, radar, accelerometer, engine sensors, microphone, GPS

PEAS: Medical diagnosis system

- **Performance measure**
  - Patient health, cost, reputation

- **Environment**
  - Patients, medical staff, insurers, courts

- **Actuators**
  - Screen display, email

- **Sensors**
  - Keyboard/mouse
<table>
<thead>
<tr>
<th>Environment types</th>
<th>Pacman</th>
<th>Backgammon</th>
<th>Diagnosis</th>
<th>Taxi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully or partially observable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-agent or multiagent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deterministic or stochastic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Static or dynamic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrete or continuous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Known physics?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Known perf. measure?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Agent design

- The environment type largely determines the agent design
  - **Partially observable** => agent requires *memory* (internal state)
  - **Stochastic** => agent may have to prepare for *contingencies*
  - **Multi-agent** => agent may need to behave *randomly*
  - **Static** => agent has time to compute a rational decision
  - **Continuous time** => continuously operating *controller*
  - **Unknown physics** => need for *exploration*
  - **Unknown perf. measure** => observe/interact with *human principal*
Simple reflex agents

- Agent
  - Sensors
  - What the world is like now
    - Condition-action rules
      - What action I should do now
    - Actuators
- Environment
class GoWestAgent(Agent):

def getAction(self, percept):
    if Directions.WEST in percept.getLegalPacmanActions():
        return Directions.WEST
    else:
        return Directions.STOP
Eat adjacent dot, if any
Eat adjacent dot, if any
Can we (in principle) extend this reflex agent to behave well in all standard Pacman environments?

- No – Pacman is not quite fully observable (power pellet duration)
- Otherwise, yes – we can *(in principle)* make a lookup table.....

*How large would it be?*
Reflex agents with state

Agent

- State
  - How the world evolves
  - What my actions do
  - Condition-action rules

Environment

- Sensors
  - What the world is like now
- Actuators
  - What action I should do now
Goal-based agents

Agent

Goals

State

How the world evolves

What my actions do

What the world is like now

What it will be like if I do action A

What action I should do now

Actuators

Environment

Sensors
Spectrum of representations

(a) Atomic

(b) Factored

(c) Structured
Outline of the course

- Deterministic
- Stochastic
- Known
- Unknown
- Atomic
- Factored
- Structured
- Search
- Logic
- RL
- Bayes nets
- MDPs
- First-order logic