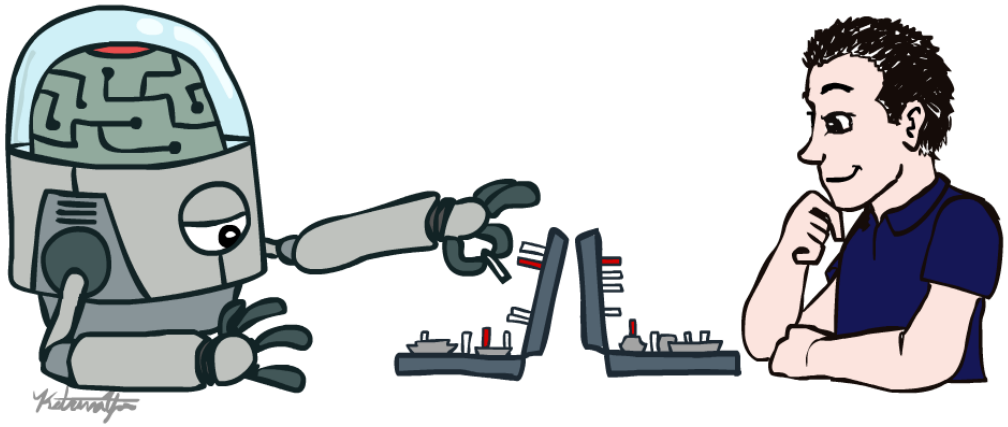
CSCE 580: Artificial Intelligence

Introduction



Instructor: Pooyan Jamshidi

University of South Carolina

Course Staff

Instructor

Pooyan Jamshidi

TA

Jianhai Su

Course Information

https://pooyanjamshidi.github.io/csce580/

- Communication:
  - Announcements on webpage/emails/dropbox
  - Questions? Discussion on piazza
  - Staff email: suj@email.sc.edu
- Course technology:
  - Website
  - Piazza
  - Gradescope
  - CSE Dropbox
  - Autograded projects
  - Regular homework
  - Help us make it awesome!

Course Information

- Course Website: https://pooyanjamshidi.github.io/csce580/
- Piazza: http://piazza.com/sc/spring2019/csce580
  - Discussion boards for each assignment and the course overall
    - PLEASE post questions on course material (don't be shy)
    - Answer others' questions - if you know the answer ;-)
    - Learn from others' questions and answers
    - Check it Often

Course Information

- Prerequisites:
  - Required: CSCE 350: Data Structures and Algorithms
  - There will be a lot of math (and programming)
  - Prior computer programming experience is required. Additional background in data structures and algorithms, linear algebra, and probability will all be helpful.
  - You should be prepared to review basic probability on your own if it is not fresh in your head.

Course Information

- 5 programming projects: Python, groups of 1 or 2
  - 7 late days for the entire semester (maximum 2 for a given project)
- ~5-10 homework assignments:
  - Written, solve together, write up alone, electronic submission through dropbox
- One midterms, one final
- Participation can help on margins
- Fixed scale
- Academic integrity policy

Exam Dates

- Midterm: March 19, Thursday, 4:25 p.m. (In Class)
- Final: May 5, Tuesday, 4:00 p.m.

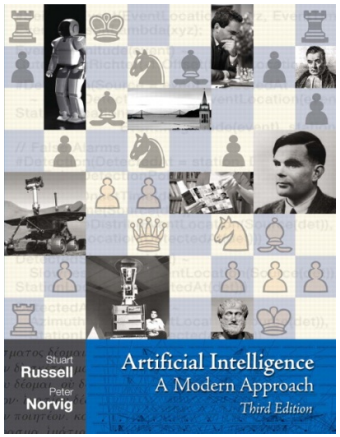
- There will be no alternative exams, put them in your calendar

Discussion Section

- There will be few discussion sections
- Topic: review / warm-up exercises
- Will be announced via Piazza

Textbook

- Not required, but for students who want to read more we recommend
  - Russell & Norvig, AI: A Modern Approach, 3rd Ed.



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

Laptops in Lecture

- Laptops can easily distract students behind you
- Please consider sitting towards the back if using your laptop in lecture

Course Information: Assignments

- There will be programming and written assignments

- You will have a total of 7 late days for these assignments, *up to two of which can be used for each assignment*.
  - To allow you the flexibility to manage unexpected issues
  - Additional late days will not be granted except under truly exceptional circumstances
  - If you've used up all your late days, you lose 20% per day (see details on the course website)

Course Information

- Programming Language: Course programming projects will be in Python.
- P0 is designed to teach you the basics of Python (Due: Jan 20)

Project 0

- Due on Monday 20, 11:59 pm
- To be done alone
- Details on the course website, announcements on Dropbox and Piazza
- Submission via Dropbox

- Python 3.6
- Autograder: We have provided a local autograder and a set of test cases for you to evaluate your code. The local autograder is a file called autograder.py.

Project 0

- **What to submit:**
  - The files that are required in the project's description (addition.py, buyLotsOfFruit.py, and shopSmart.py). Please use comments appropriately across your code.
  - A short README.txt file that specifies:
    - Your name and ID.
    - A brief description (i.e. a short paragraph) that includes the main ideas of your implementation.

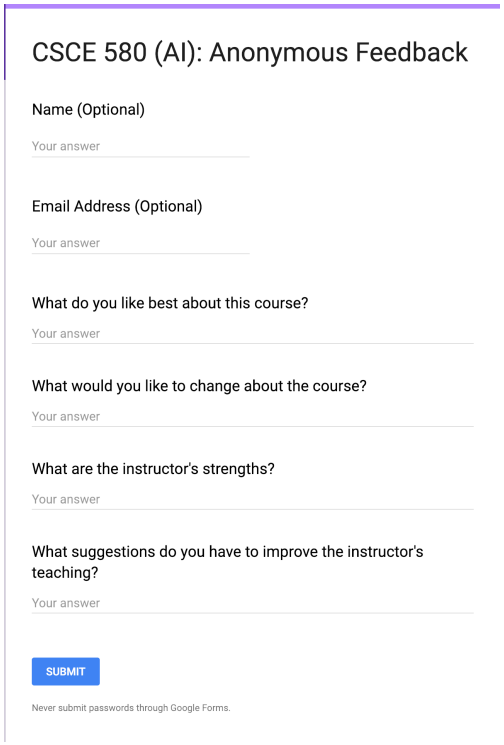
- **Place your files in a single folder inside the archive.**
  Submit your assignment on Dropbox as a single archive file (.zip), with the name csce580-p0-lastname-uscid

Homework 0

- Due on Monday 20, 11:59 pm
- To be done alone
- Details on the course website, announcements on Dropbox and Piazza
- Submission via Dropbox
- To assess whether you are mathematically prepared for the second half of the course.

Course Information: Feedback

- Please give feedback (positive or negative) as often as and as early as you can.



CSCE 580 (AI): Anonymous Feedback

Name (Optional)

Your answer

Email Address (Optional)

Your answer

What do you like best about this course?

Your answer

What would you like to change about the course?

Your answer

What are the instructor's strengths?

Your answer

What suggestions do you have to improve the instructor's teaching?

Your answer

SUBMIT

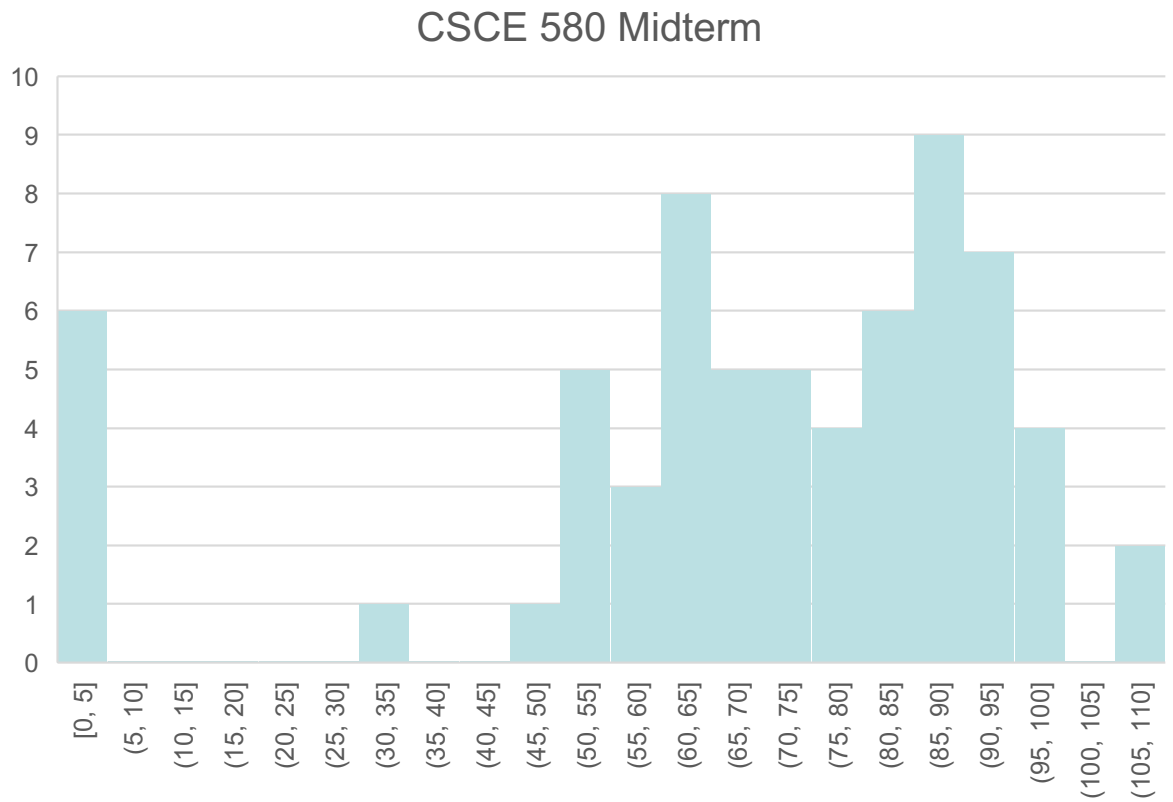
Never submit passwords through Google Forms.

https://tinyurl.com/yapwj7bp

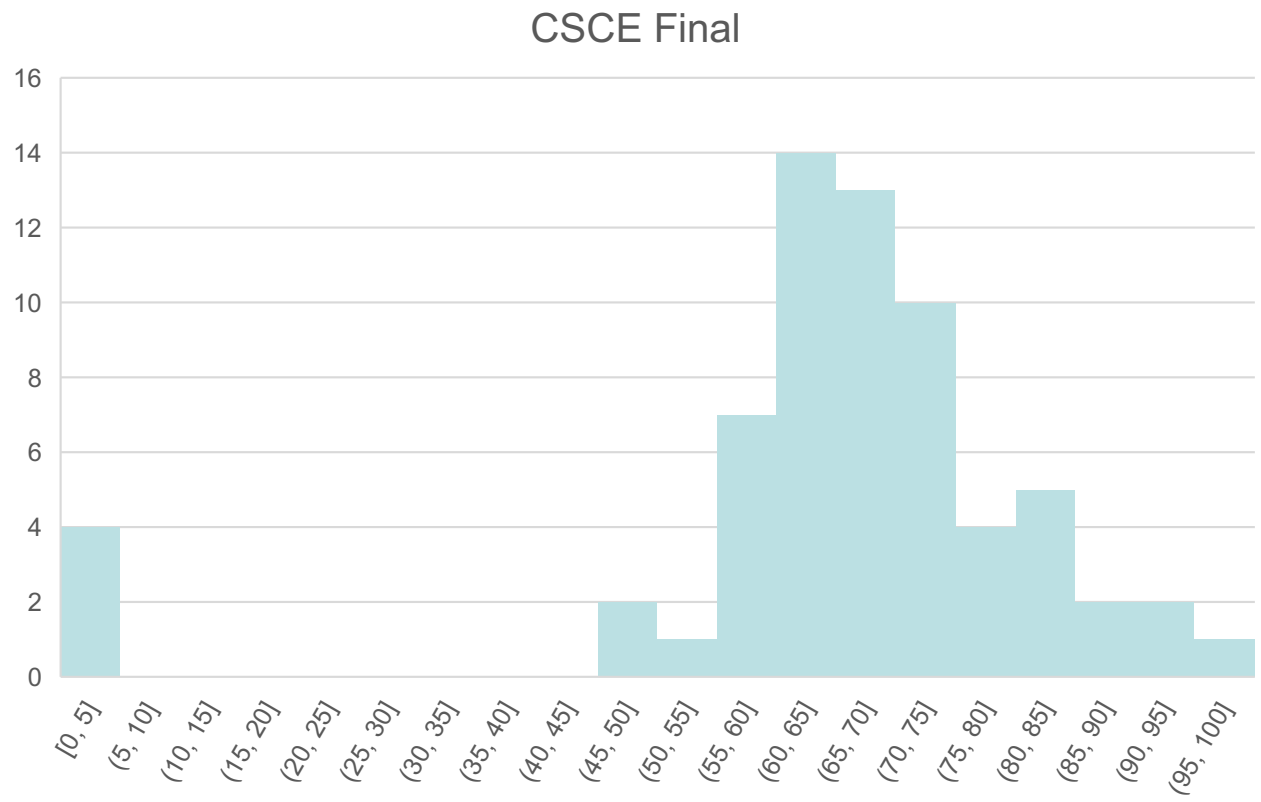
Acknowledgements

- Dan Klein
- Pieter Abbeel
- Stuart Russell
- Brad Miller
- Nick Hay
- John DeNero
- Pooyan Fazli

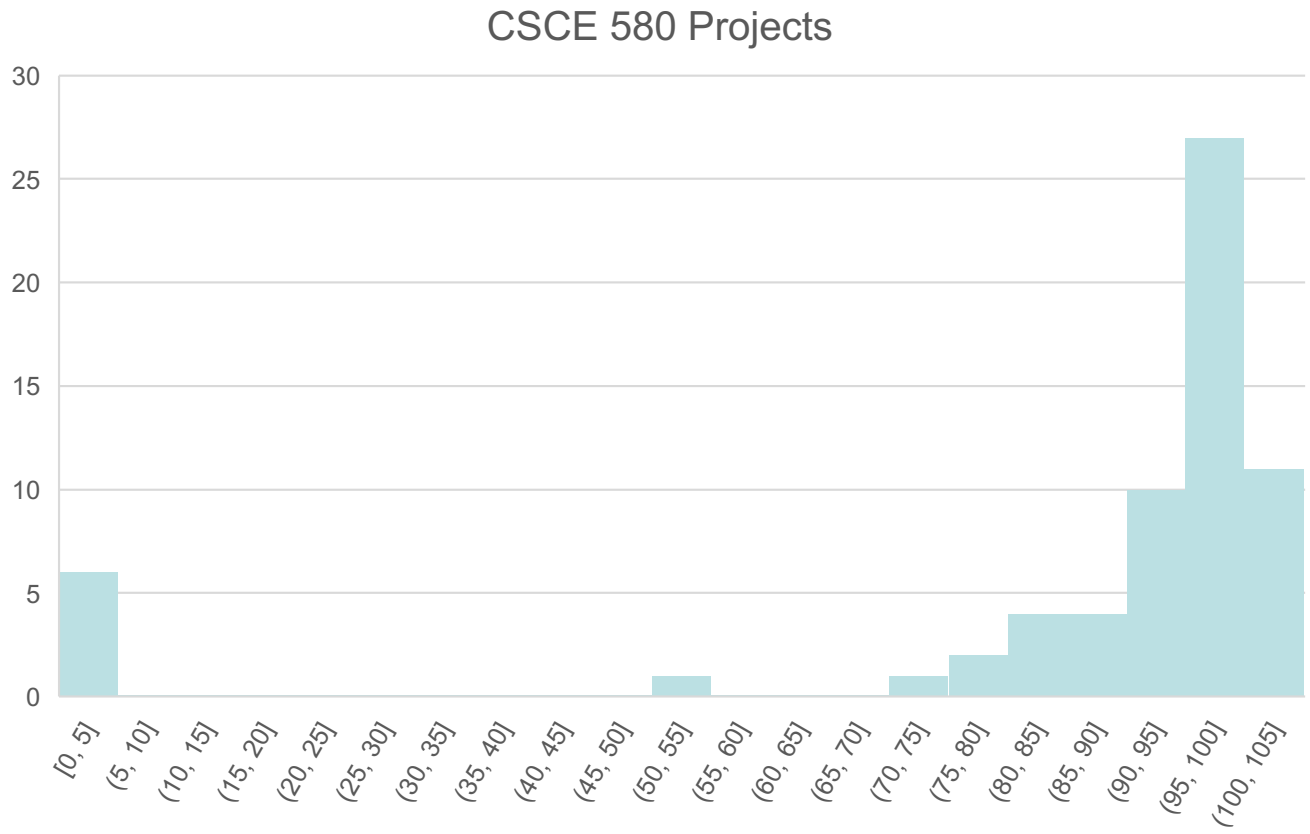
Historical Statistics



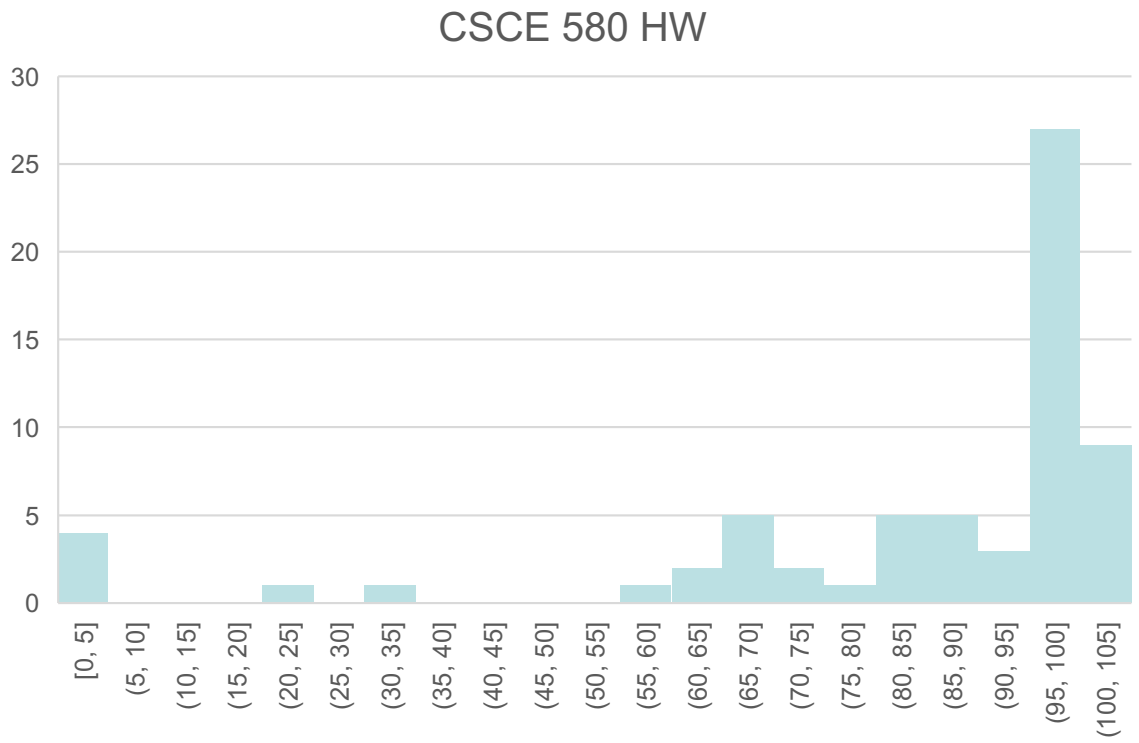
Historical Statistics



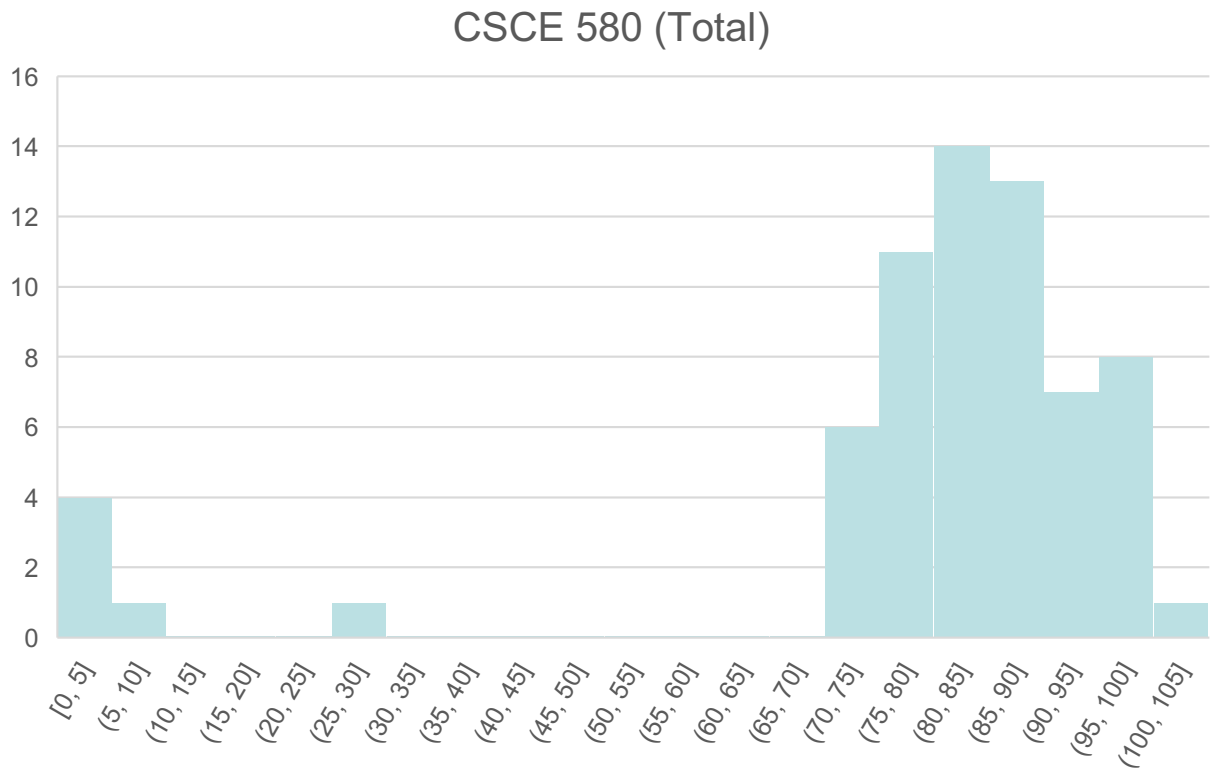
Historical Statistics



Historical Statistics



Historical Statistics

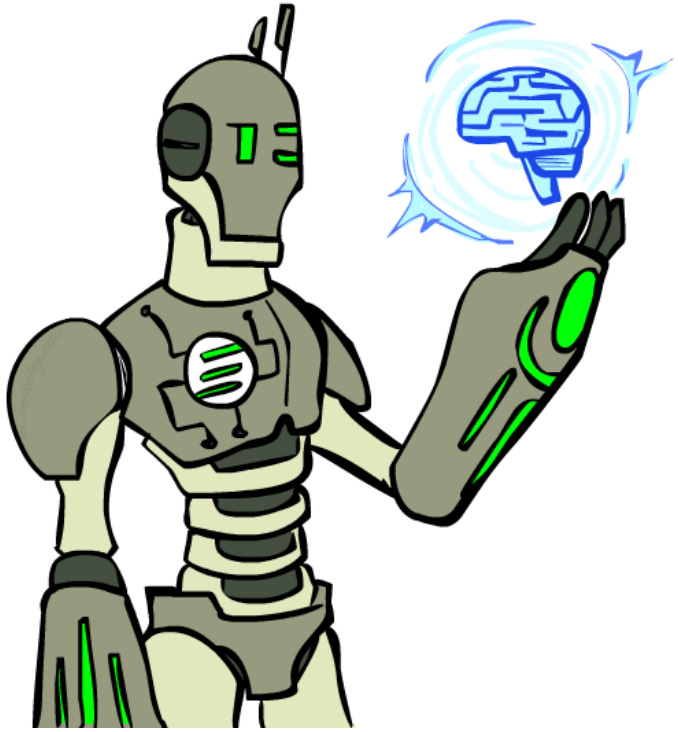


Important This Week

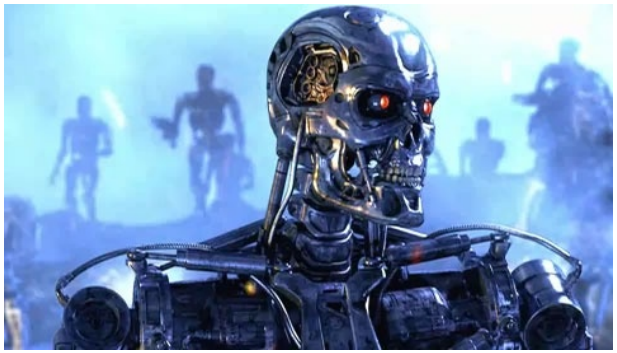
- Important this week:
  - **Checkout website:** https://pooyanjamshidi.github.io/csce580/
  - **Register** for the class on piazza --- our main resource for discussion and communication
  - **P0: Python tutorial** is out (due on Monday 1/20 at 11:59pm)
  - **HW0: Math diagnostic homework** is out (due on Monday 1/20 at 11:59pm)
  - **Office Hours** start next week, and you can catch the professor after lecture

Today

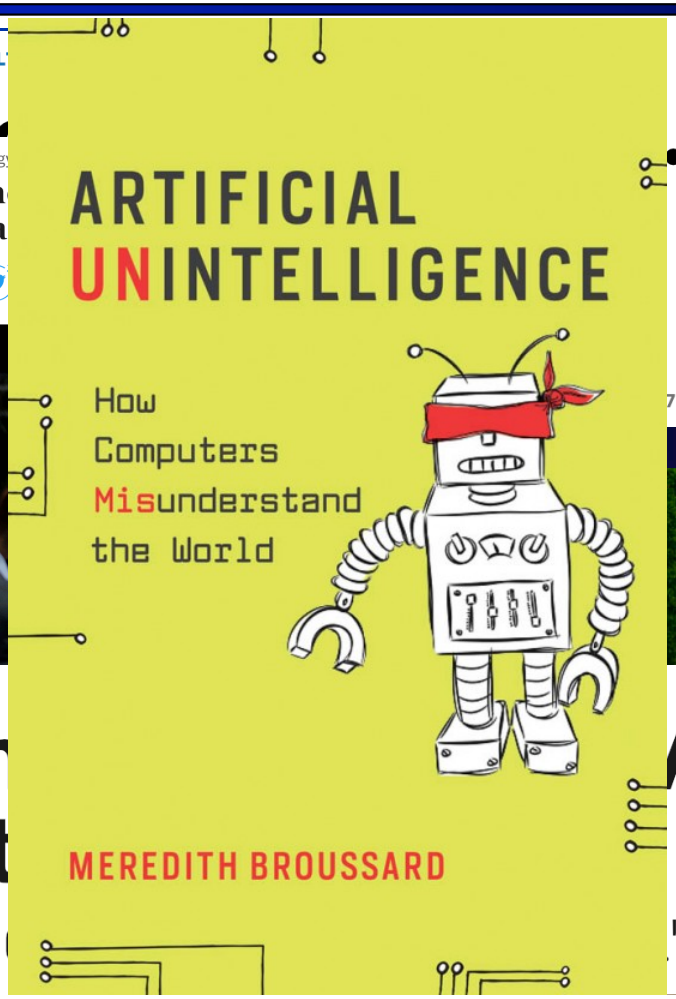
- What is artificial intelligence?
- What can AI do?
- What is this course?



Sci-Fi AI?

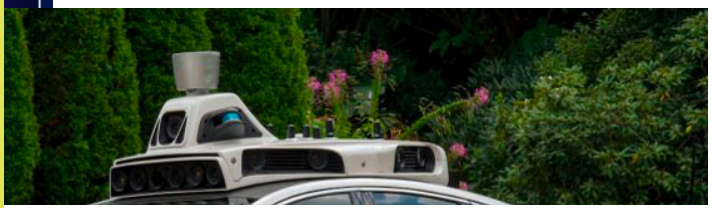


AI in News?



. can transform
know it. Doctors
slow down

7 2018 • 10:53 AM EDT



ARTIFICIAL INTELLIGENCE

Stephen Hawkin
Humans Altoget

AI could essentially be a "new form

AI May Replace

What is AI?

The science of making machines that:

Rational Decisions

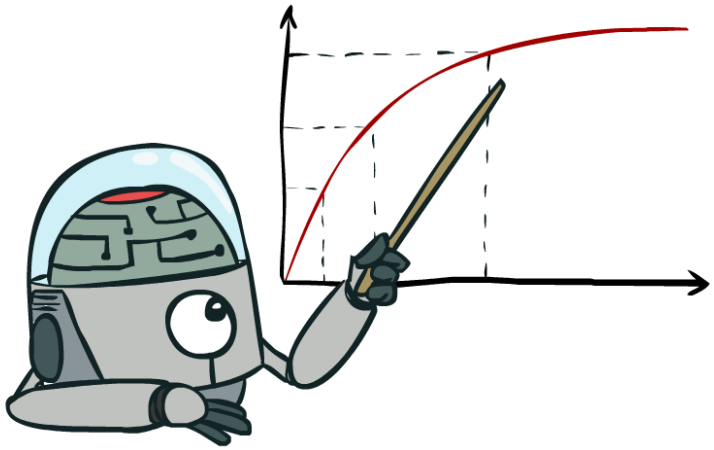
We'll use the term **rational** in a very specific, technical way:

- 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



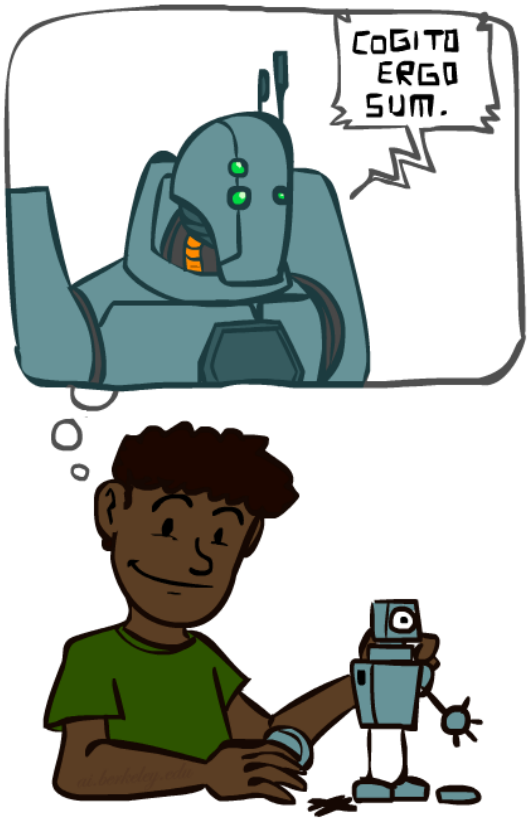
What About the Brain?

- Brains (human minds) are very good at making rational decisions, but not perfect
- Brains aren't as modular as software, so hard to reverse engineer!
- "Brains are to intelligence as wings are to flight"
- Lessons learned from the brain: memory and simulation are key to decision making

Course Topics

- Part I: Making Decisions
  - Fast search / planning
  - Constraint satisfaction
  - Adversarial and uncertain search
- Part II: Reasoning under Uncertainty
  - Bayes' nets
  - Decision theory
  - Hidden Markov Models
- Part III: Machine learning
  - Naïve Bayes
  - Perceptrons and Logistic Regression
  - Neural Networks
  - Decision Trees and Support Vector Machines

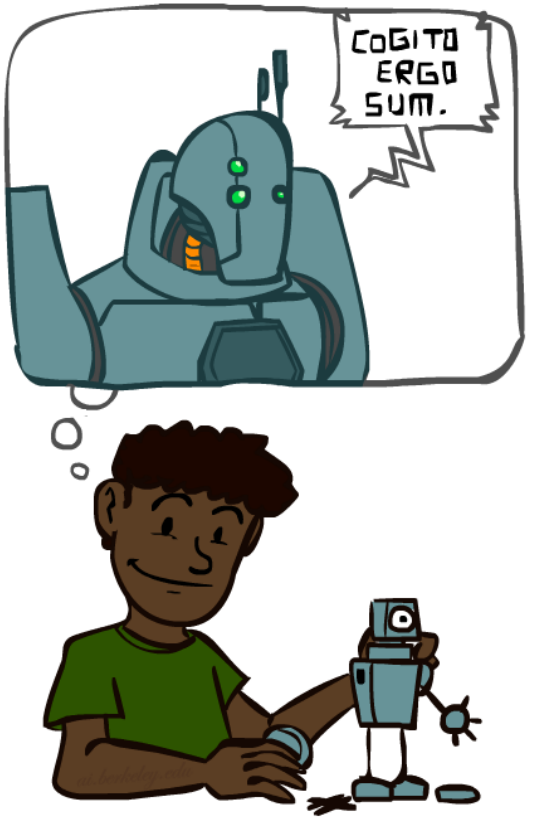
A (Short) History of AI



Demo: HISTORY – MT1950.wmv

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
  - 1988—93: Expert systems industry busts: “AI Winter”
- 1990—: Statistical approaches
  - Resurgence of probability, focus on uncertainty
  - General increase in technical depth
  - Agents and learning systems… “AI Spring”?
- 2012—: Where are we now?
  - Big data, big compute, neural networks
  - Some re-unification of sub-fields
  - AI used in many industries



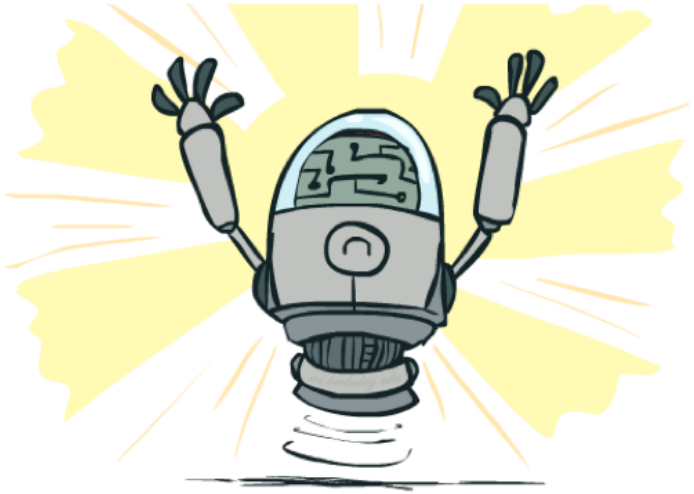
What Can AI Do?

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?

- ✓ Buy a week's worth of groceries on the web?

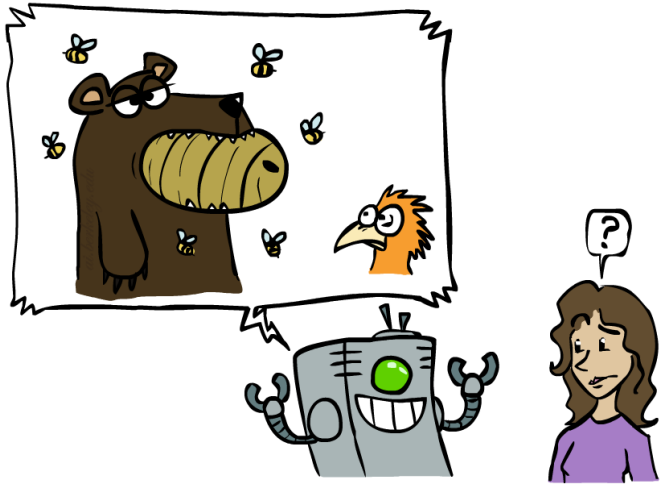
- ? Discover and prove a new mathematical theorem?
- ✗ Converse successfully with another person for an hour?
- ? Perform a surgical operation?
- ✓ Put away the dishes and fold the laundry?
- ✓ Translate spoken Chinese into spoken English in real time?
- ✗ Write an intentionally funny story?



Unintentionally Funny Stories

- One day Joe Bear was hungry. He asked his friend Irving Bird where some honey was. Irving told him there was a beehive in the oak tree. Joe walked to the oak tree. He ate the beehive. The End.

- Henry Squirrel was thirsty. He walked over to the river bank where his good friend Bill Bird was sitting. Henry slipped and fell in the river. Gravity drowned. The End.



- Once upon a time there was a dishonest fox and a vain crow. One day the crow was sitting in his tree, holding a piece of cheese in his mouth. He noticed that he was holding the piece of cheese. He became hungry, and swallowed the cheese. The fox walked over to the crow. The End.

[Shank, Tale-Spin System, 1984]

Natural Language

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

- Language processing technologies
  - Question answering
  - Machine translation

"Il est impossible aux journalistes de rentrer dans les régions tibétaines"

Bruno Philip, correspondant du "Monde" en Chine, estime que les journalistes de l'AFP qui ont été expulsés de la province tibétaine du Qinghai "n'étaient pas dans l'illégalité".

Les faits Le dalaï-lama dénonce l'"enfer" imposé au Tibet depuis sa fuite, en 1959

Vidéo Anniversaire de la rébellion

"It is impossible for journalists to enter Tibetan areas"

Philip Bruno, correspondent for "World" in China, said that journalists of the AFP who have been deported from the Tibetan province of Qinghai "were not illegal."

Facts The Dalai Lama denounces the "hell" imposed since he fled Tibet in 1959

Video Anniversary of the Tibetan rebellion: China on guard

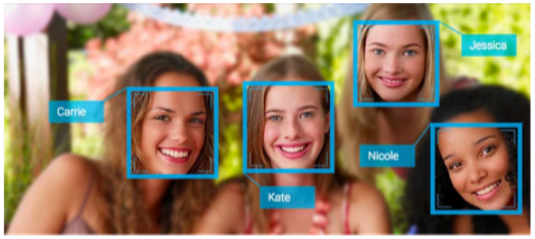
  - Web search
  - Text classification, spam filtering, etc…

Vision (Perception)

PIXELS -> INFO/DECISION

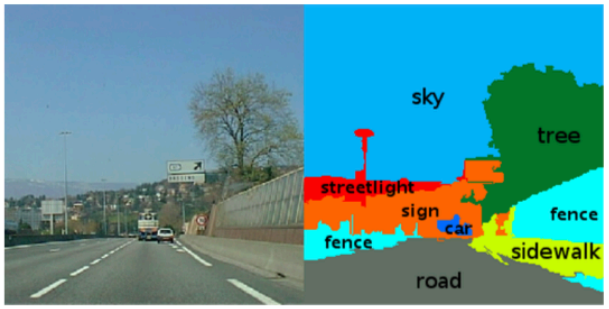
E.g.:

Face detection and recognition



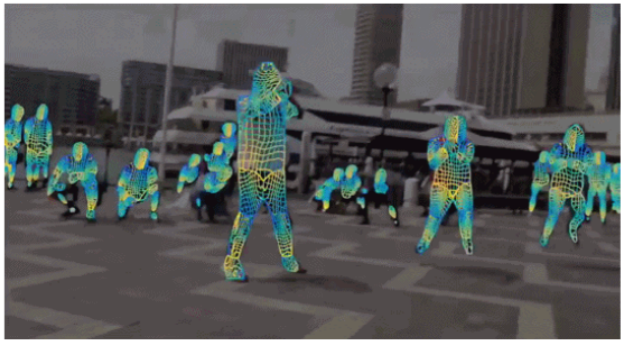
Source: TechCrunch

Semantic Scene Segmentation



[Caesar et al, ECCV 2017]

3-D Understanding



[DensePose]

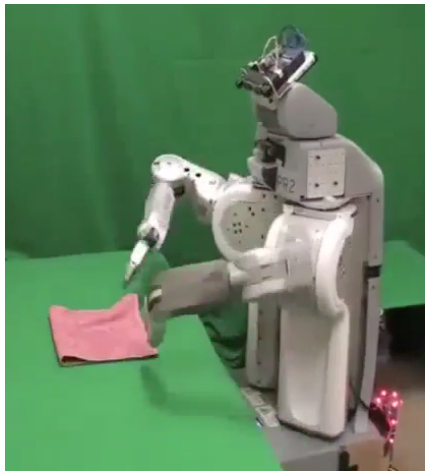

TRACKING:
N
NW
NE
W
E

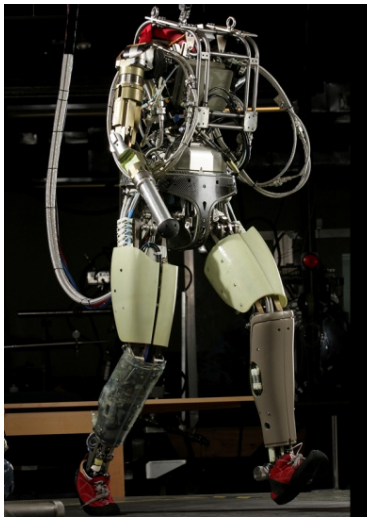

car 1.000
car 1.000
LOUIS VUITTON

Robotics

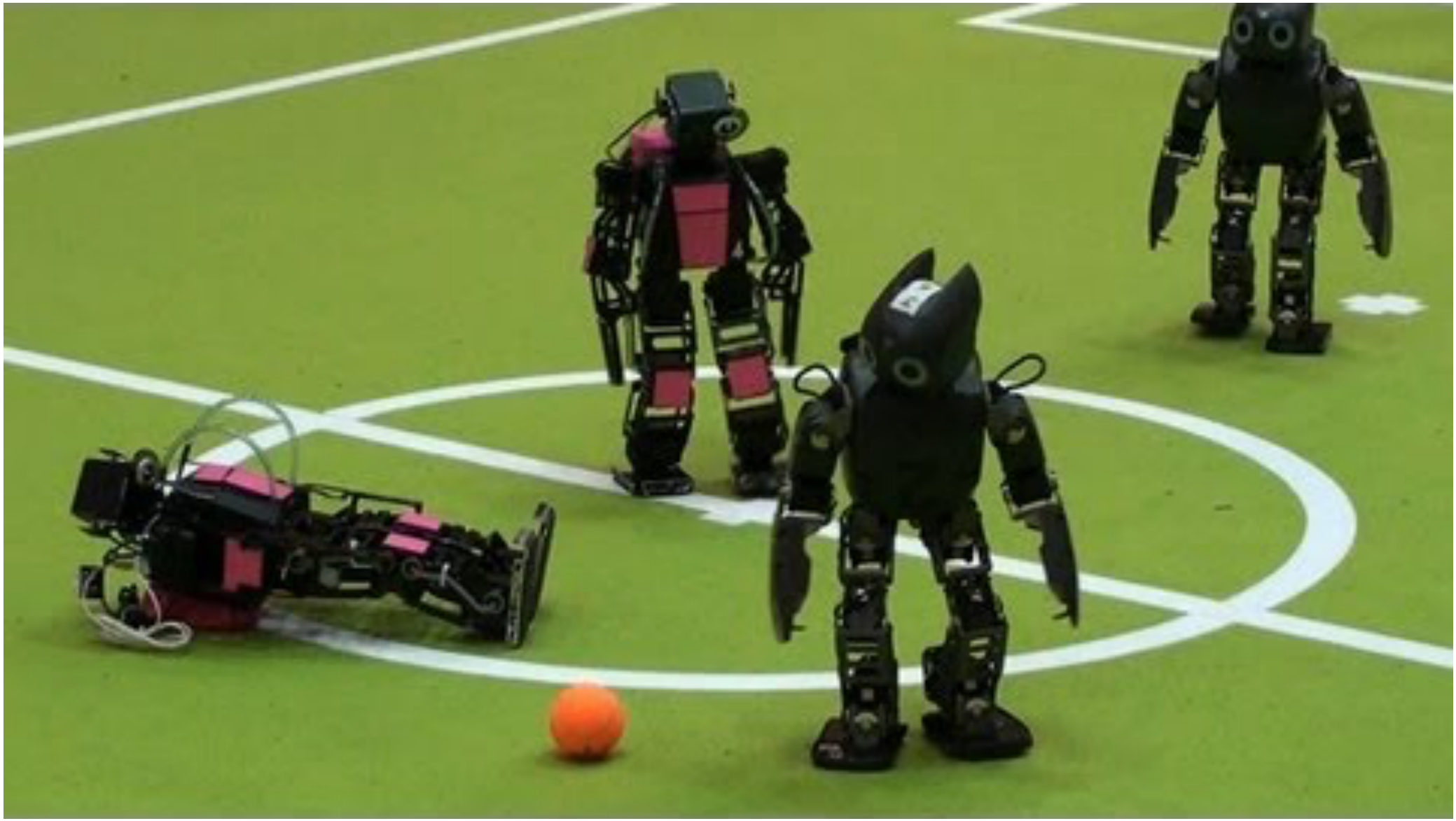
Demo 1: ROBOTICS – soccer.avi
Demo 2: ROBOTICS – soccer2.avi
Demo 3: ROBOTICS – gcar.avi
Demo 4: ROBOTICS – laundry.avi
Demo 5: ROBOTICS – petman.avi

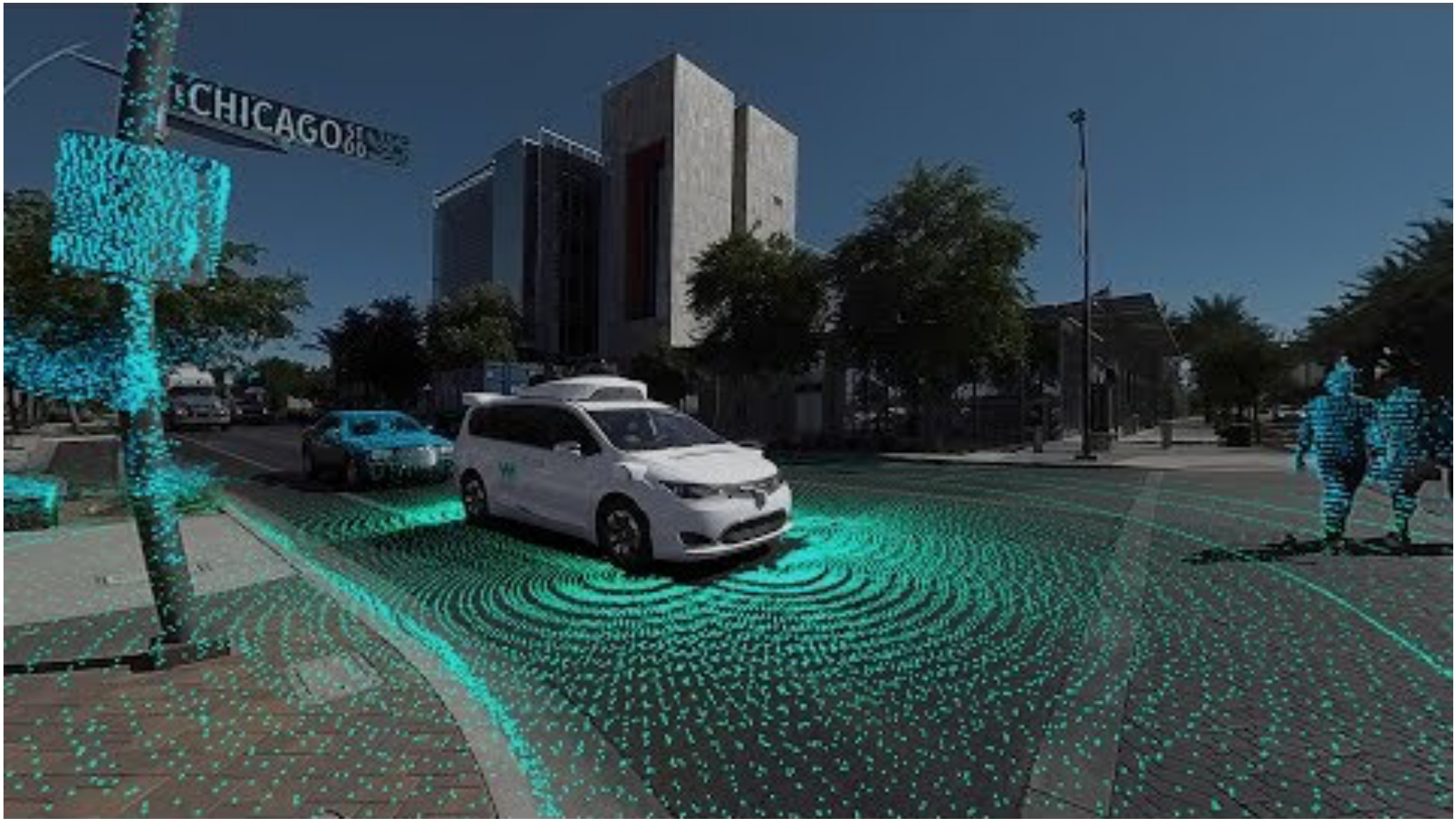
- Robotics
  - Part mech. eng.
  - Part AI
  - Reality much harder than simulations!
- Technologies
  - Vehicles
  - Rescue
  - Soccer!
  - Lots of automation…
- In this class:
  - We ignore mechanical aspects
  - Methods for planning
  - Methods for control

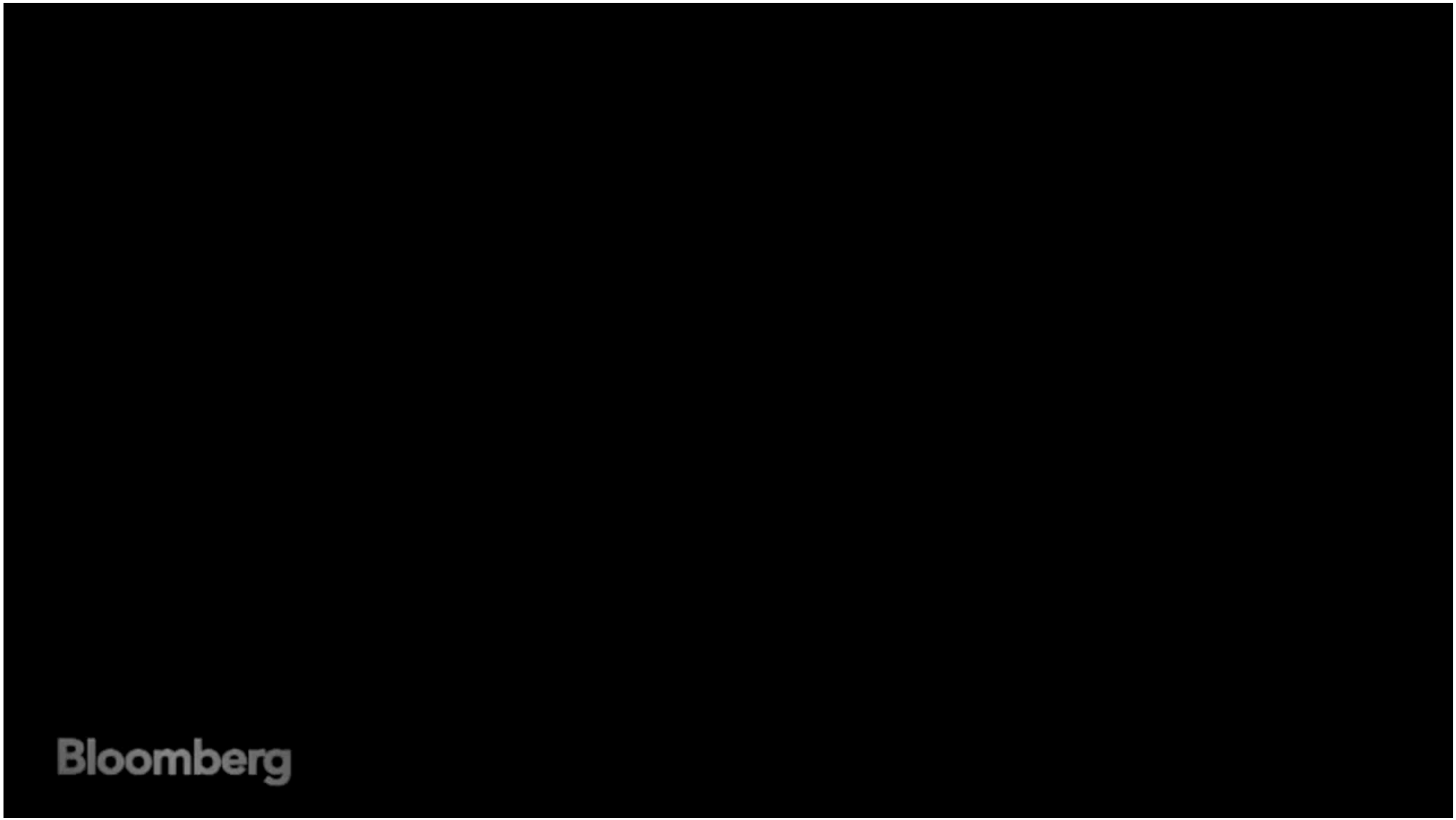


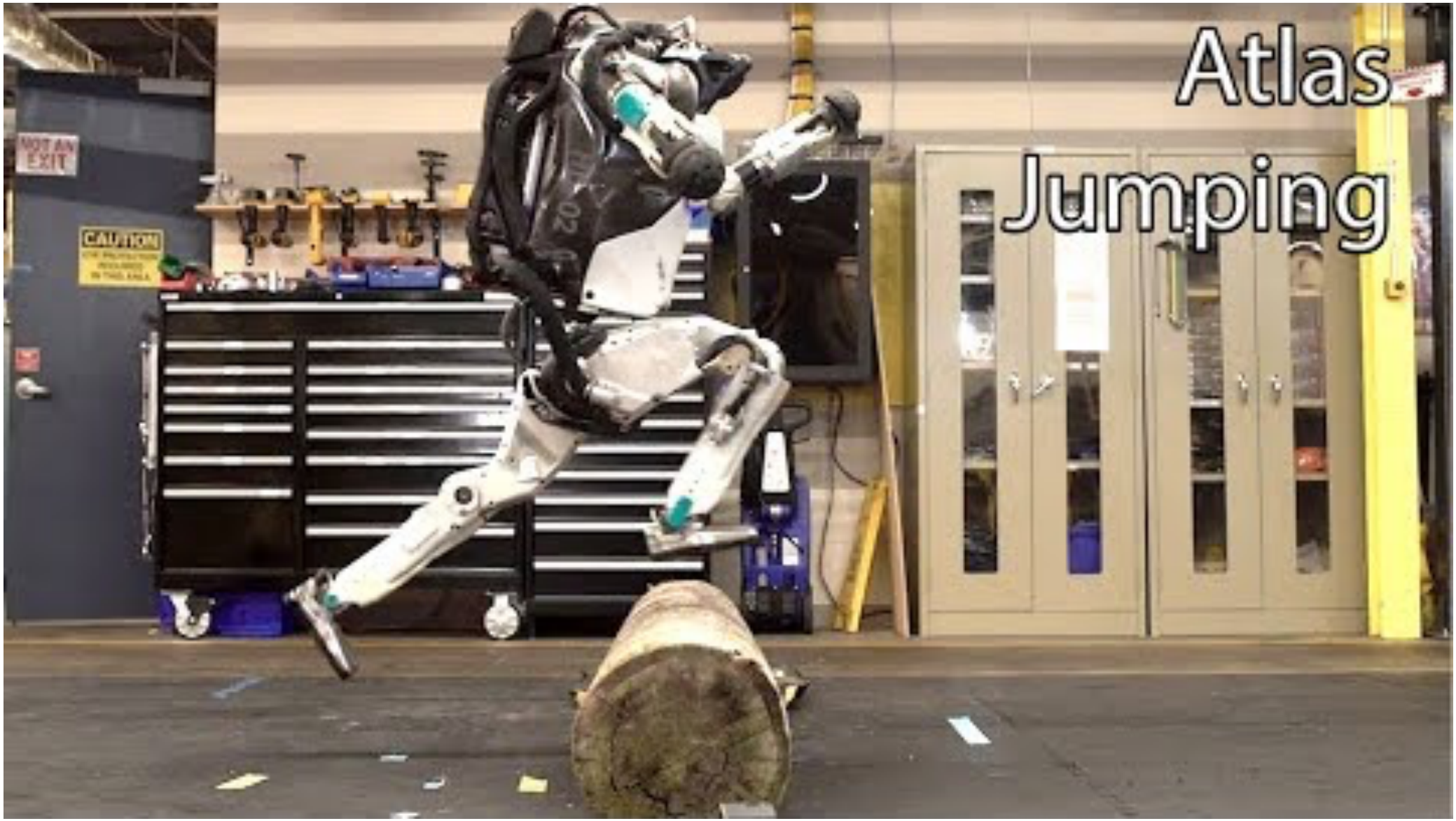


Images from UC Berkeley, Boston Dynamics, RoboCup, Google



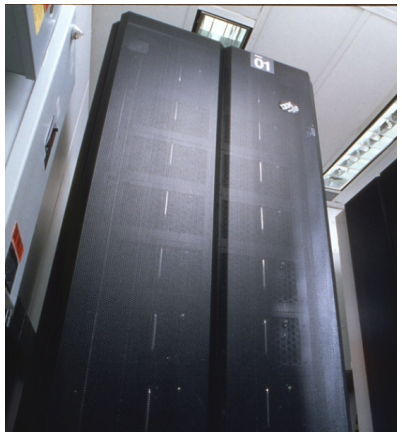

CHICAGO


Bloomberg


Atlas
Jumping
NOT AN EXIT
CAUTION

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 AlphaGo beats Lee Sedol

Text from Bart Selman, image from IBM’s Deep Blue pages

Logic

- Logical systems
  - Theorem provers
  - NASA fault diagnosis
  - Question answering
- Methods:
  - Deduction systems
  - Constraint satisfaction
  - Satisfiability solvers (huge advances!)

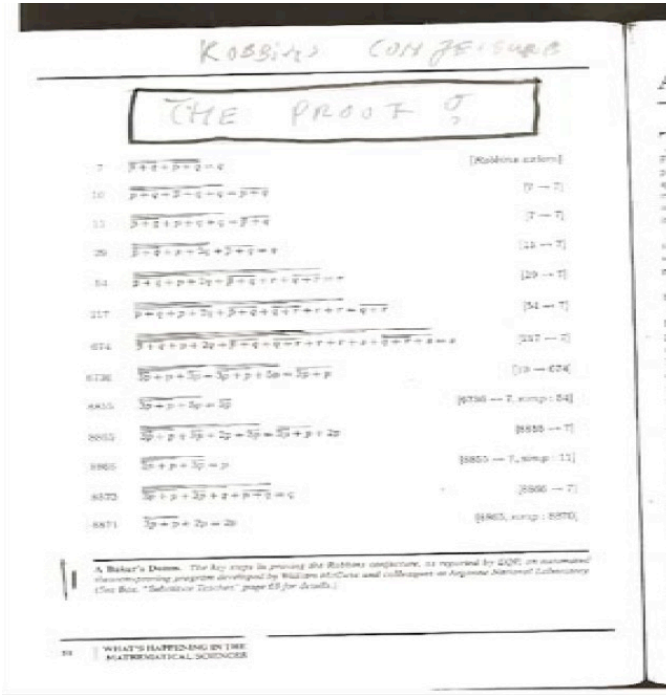
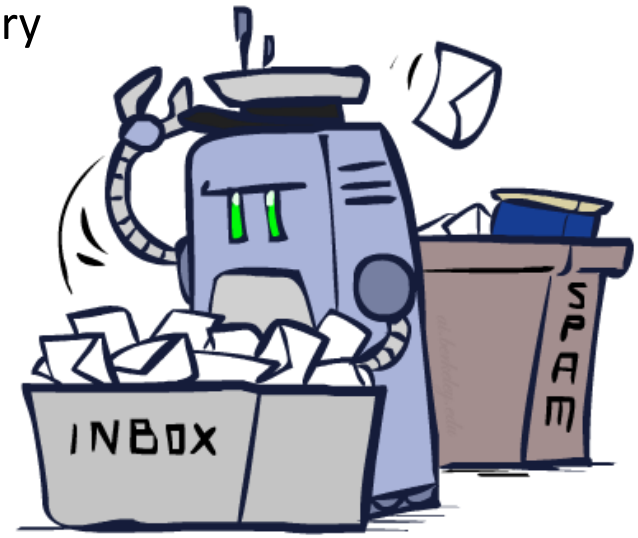


Image from Bart Selman

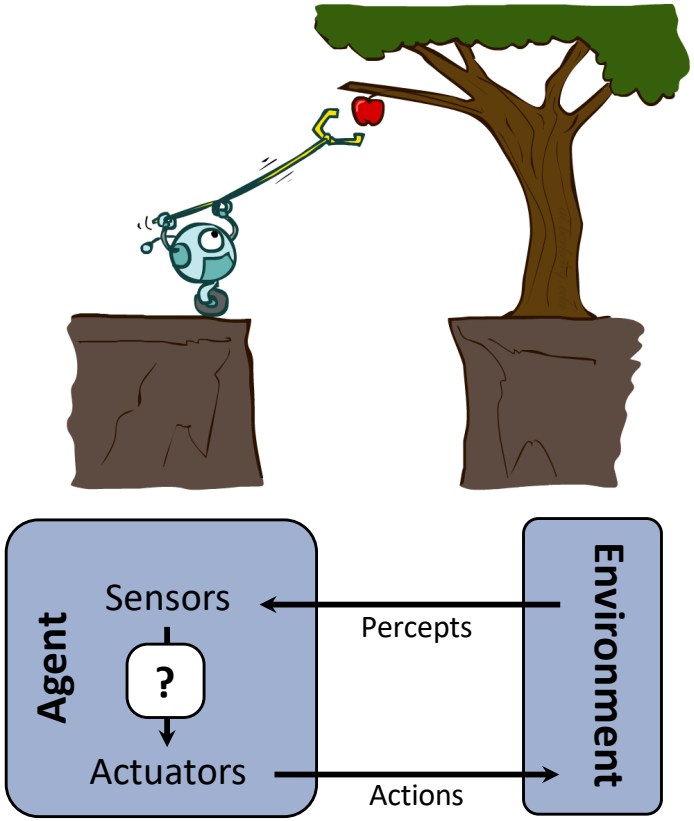
AI is starting to be everywhere…

- 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!



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 **percepts, environment,** and **action space** dictate techniques for selecting rational actions
- **This course** is about:
  - General AI techniques for a variety of problem types
  - Learning to recognize when and how a new problem can be solved with an existing technique



Agents

- A Goal of AI: Build robust, fully autonomous agents in the real world

Intelligent (Autonomous) Agents: Examples

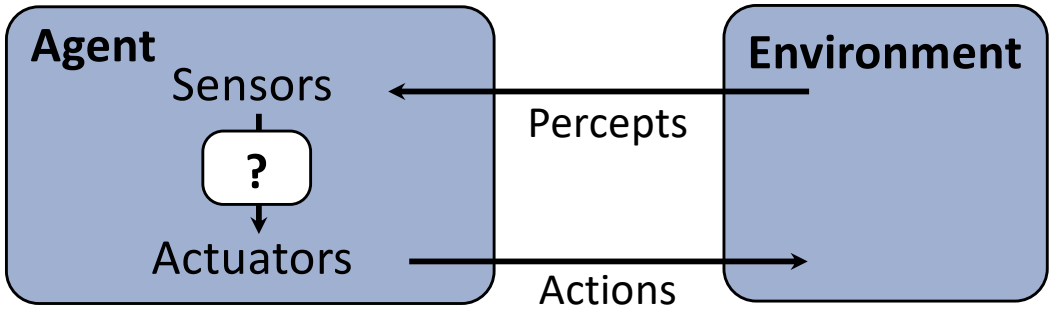
- Autonomous robot
- Information gathering agent
  - Find me the cheapest?
- E-commerce agents
  - Decides what to buy/sell and does it
- Air-traffic controller
- Meeting scheduler
- Computer-game-playing agent

Not Intelligent Agents

- Thermostat
- Telephone
- Answering machine
- Pencil
- Java object

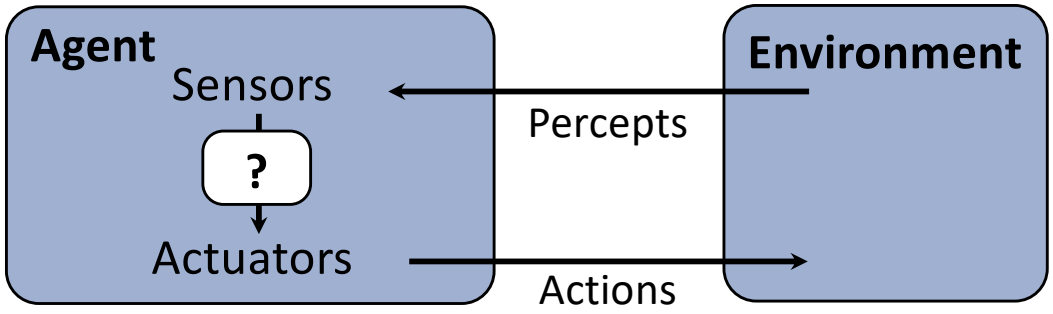
What is an Agent?

Agents and Environments



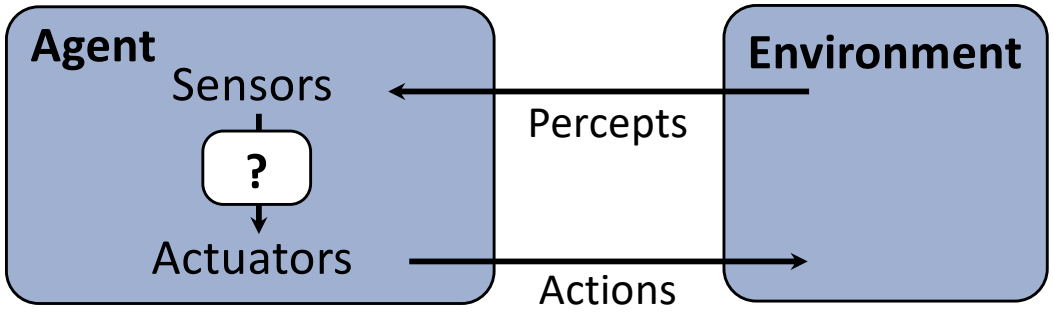
- An agent *perceives* its environment through *sensors* and *acts* upon it through *actuators*.

Agents and Environments



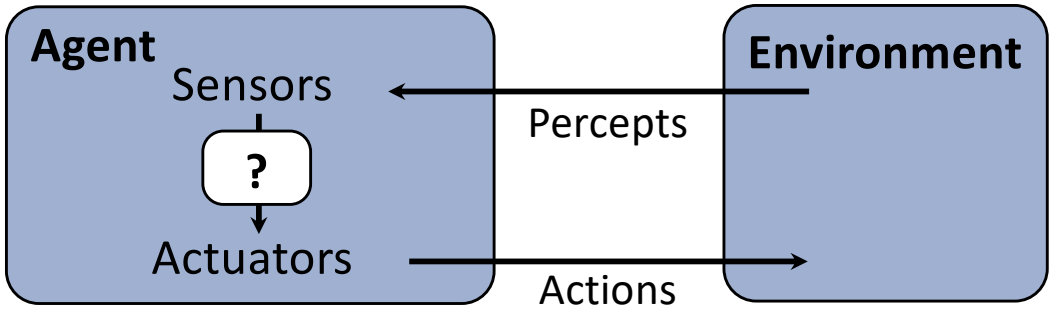
- Are humans agents?
- Yes!
  - Sensors = vision, audio, touch, smell, taste, …
  - Actuators = muscles, secretions, changing brain state

Agents and Environments



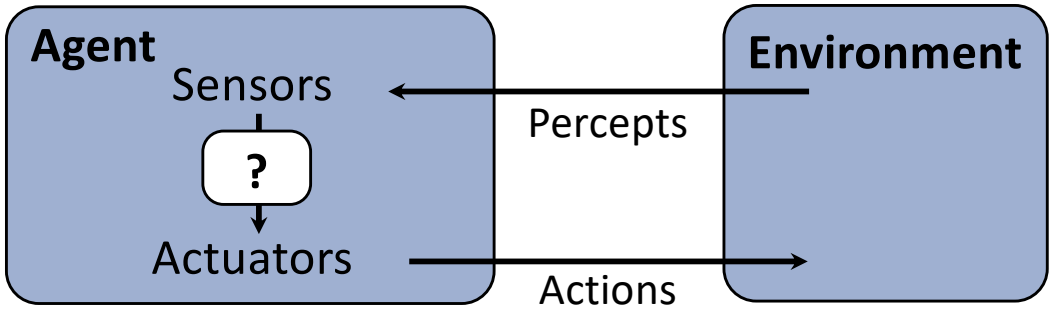
- Are Robots agents?
- Yes!
  - Sensors = cameras, laser range finders, GPS
  - Actuators = various motors

Agents and Environments



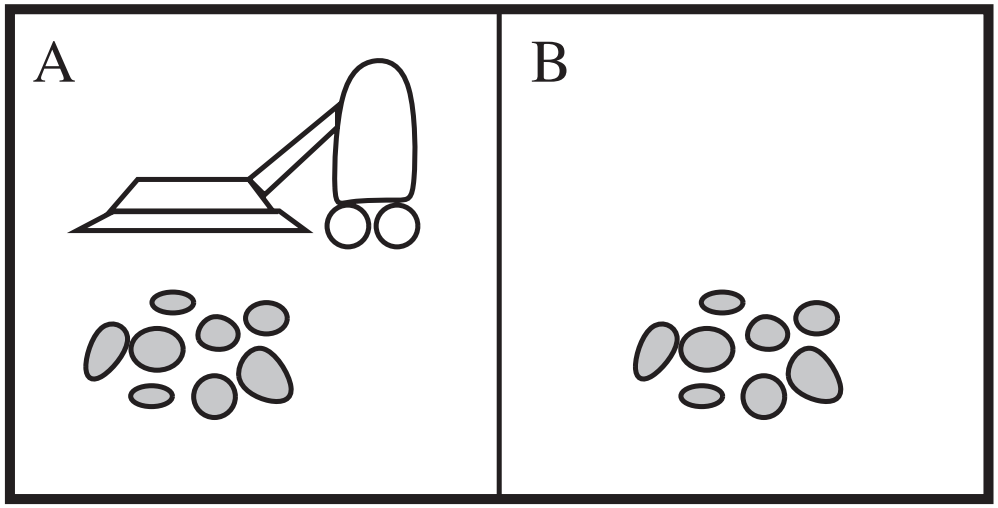
- Are pocket calculators agents?
- Yes!
  - Sensors = key state sensors
  - Actuators = digit display

Agents and Environments



- AI is more interested in agents with substantial computation resources and environments requiring nontrivial decision making

Example: Vacuum world

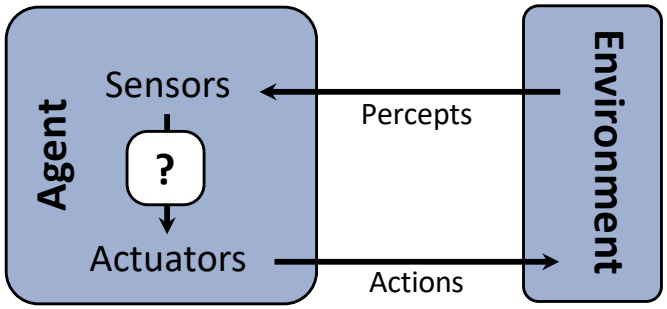


- Percepts: [location,status], e.g., [A,Dirty]
- Actions: Left, Right, Suck, NoOp

Rational Agents

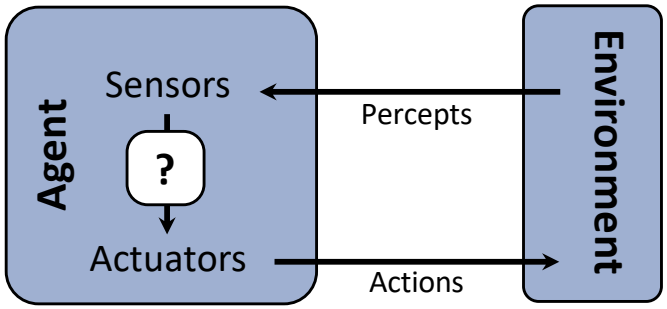
- A **rational agent** selects actions that maximize its (expected) **utility**.

- Utility or performance measure of a vacuum-cleaner agent:
  - amount of dirt cleaned up
  - amount of time taken
  - amount of electricity consumed
  - amount of noise generated
  - etc.



Rational Agents

- A **rational agent**
  - acts appropriately given goals and circumstances
  - is flexible to changing environments and goals
  - learns from experience
  - makes appropriate choices given perceptual and computational limitations
- Characteristics of the **percepts**, **environment**, and **action space** dictate techniques for selecting rational actions.



Rational Agents

- Are rational agents omniscient?
  - No – they are limited by the available percepts
- Are rational agents clairvoyant?
  - No – they may lack knowledge of the environment dynamics
- Do rational agents explore and learn?
  - Yes – in unknown environments these are essential
- So rational agents are not necessarily successful, but they are autonomous

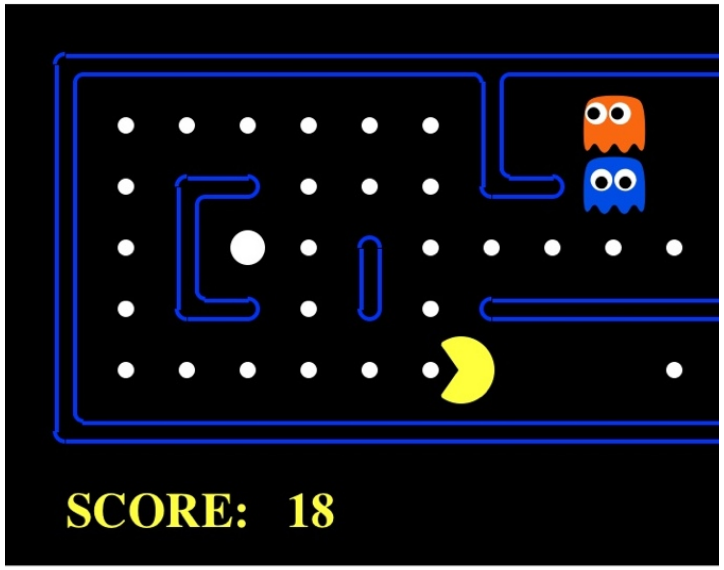
Discussion Item

- A realistic agent has finite amount of computation and memory available. Assume an agent is killed because it did not have enough computation resources to calculate some rare event that eventually ended up killing it. Can this agent still be rational?

PEAS: Performance measure, Environment, Actuators, Sensors

PEAS: Pacman

- Performance measure
  - -1 per step; + 10 food; +500 win; -500 die;
- Environment
  - Maze, food, ghosts, …
- Actuators
  - Pacman's body and mouth
- Sensors
  - Some sort of Vision (Entire state is visible)



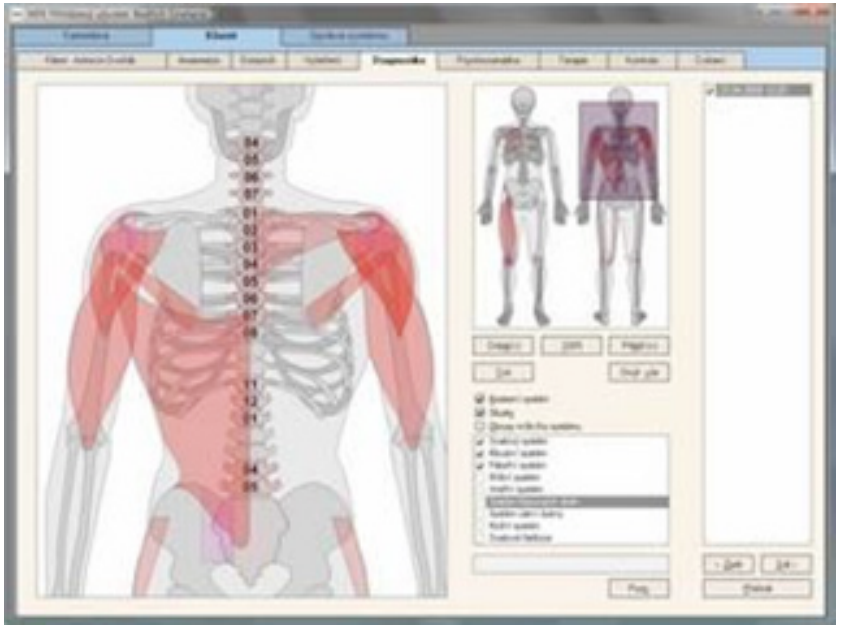
PEAS: Automated Taxi

- Performance measure
  - Income, happy customer, vehicle costs, fines, insurance premiums
- Environment
  - US streets, other drivers, customers
- Actuators
  - Steering, brake, gas, display/speaker
- Sensors
  - Camera, radar, accelerometer, engine sensors, microphone

Image: http://nypost.com/2014/06/21/how-google-might-put-taxi-drivers-out-of-business/

PEAS: Medical Diagnosis System

- Performance measure
  - Patient health, cost, reputation
- Environment
  - Patients, medical staff, insurers, courts
- Actuators
  - Screen display, email
- Sensors
  - Keyboard/mouse



Environment Types

Environment Types

- Fully Observable (vs. Partially Observable)
- Deterministic (vs. Stochastic)
- Episodic (vs. Sequential)
- Static (vs. Dynamic)
- Discrete (vs. Continuous)
- Single-Agent (vs. Multi-Agent):

Fully Observable vs. Partially-Observable Domains

- Fully-observable: The agent has access to all information in the environment relevant to its task.
- Partially-observable: Parts of the environment are inaccessible

Pacman	Crossword	Backgammon	Pick&Place Robot	Diagnosis	Taxi
Fully	Fully	Fully	Partially	Partially	Partially

Deterministic vs. Stochastic Domains

If an agent knew the initial state and its action, could it predict the resulting state? The dynamics can be:

- Deterministic: the resulting state is determined from the action and the state
- Stochastic: there is uncertainty about the resulting state

Pacman	Crossword	Backgammon	Pick&Place Robot	Diagnosis	Taxi
Deterministic	Deterministic	Stochastic	Stochastic	Stochastic	Stochastic

Episodic vs Sequential Domains

- Episodic: Current action is independent of previous actions.
- Sequential: Current choice of action will affect future actions

Pacman	Crossword	Backgammon	Pick&Place Robot	Diagnosis	Taxi
Sequential	Sequential	Sequential	Episodic	Sequential	Sequential

Static vs Dynamic Domains

- Static: Environment does not change while the agent is deliberating over what to do

- Dynamic: Environments does change

Pacman	Crossword	Backgammon	Pick&Place Robot	Diagnosis	Taxi
Static	Static	Static	Dynamic	Dynamic	Dynamic

Discrete vs Continuous Domains

- Discrete: A limited number of distinct, clearly defined states, percepts, actions, and time steps (otherwise continuous)

Pacman	Crossword	Backgammon	Pick&Place Robot	Diagnosis	Taxi
Discrete	Discrete	Discrete	Continuous	Continuous	Continuous

Single-agent vs. Multi-agent Domains

- Does the environment include other agents?
- If there are other agents whose actions affect us
  - It can be useful to explicitly model their goals and beliefs, and how they react to our actions
- Other agents can be: cooperative, competitive, or a bit of both

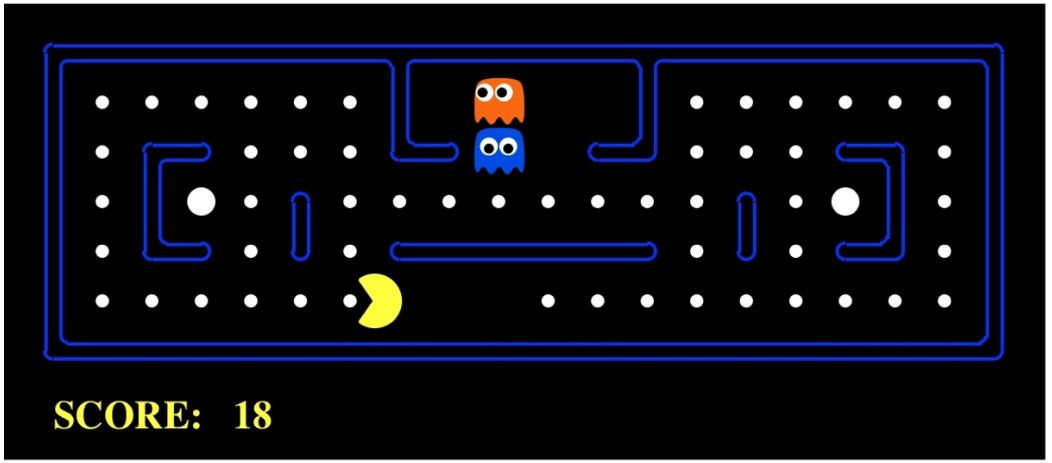
Pacman	Crossword	Backgammon	Pick&Place Robot	Diagnosis	Taxi
Multi	Single	Multi	Single	Single	Multi

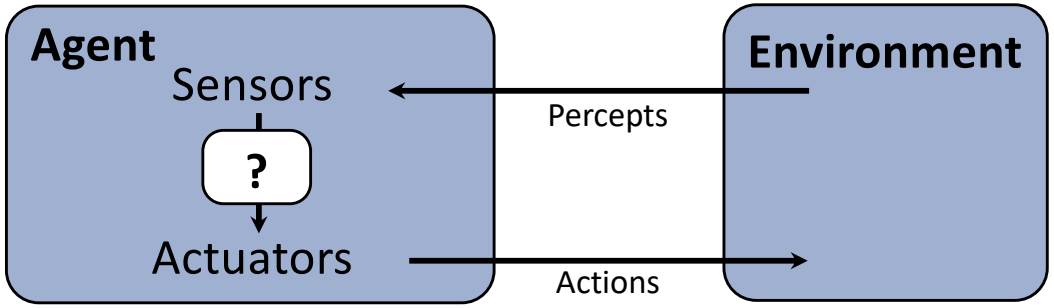
Environment Types: Summary

	Pacman	Crossword	Backgammon	Pick&Place Robot	Diagnosis	Taxi
Fully or Partially Observable	Fully	Fully	Fully	Partially	Partially	Partially
Deterministic or Stochastic	Deterministic	Deterministic	Stochastic	Stochastic	Stochastic	Stochastic
Episodic or Sequential	Sequential	Sequential	Sequential	Episodic	Sequential	Sequential
Static or Dynamic	Static	Static	Static	Dynamic	Dynamic	Dynamic
Discrete or Continuous	Discrete	Discrete	Discrete	Continuous	Continuous	Continuous
Single-agent or Multiagent	Multi	Single	Multi	Single	Single	Multi

Agent Types

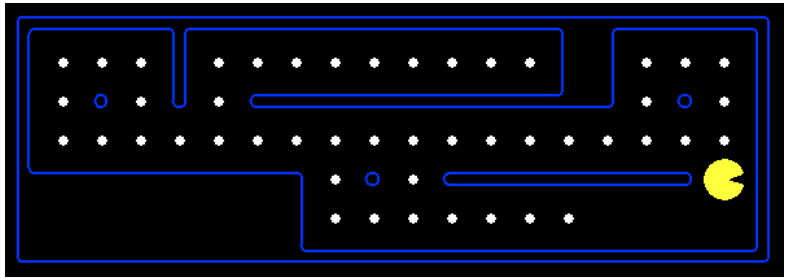
Pac-Man as an Agent

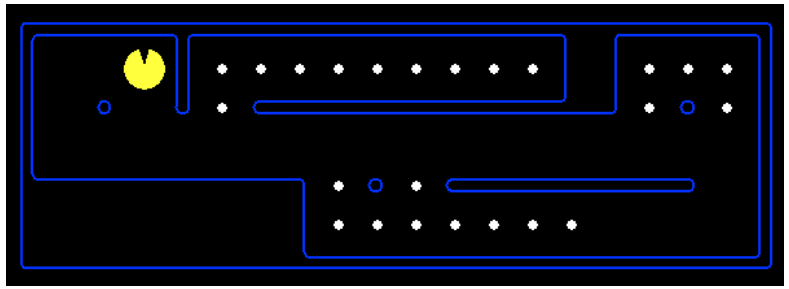




Reflex Agents

- Reflex agents:
  - Choose action based on current percept (and maybe memory)
  - May have memory or a model of the world's current state
  - Do not consider the future consequences of their actions
  - Consider how the world IS
- Can a reflex agent be rational?



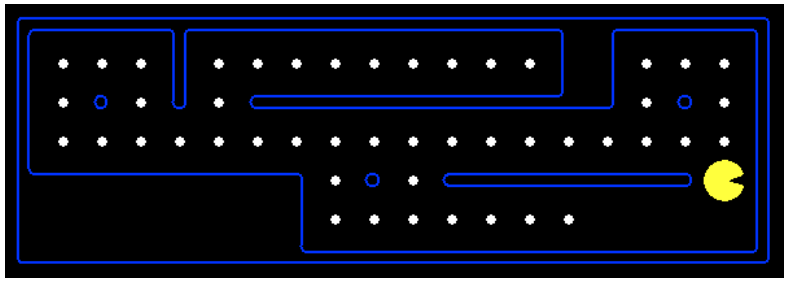


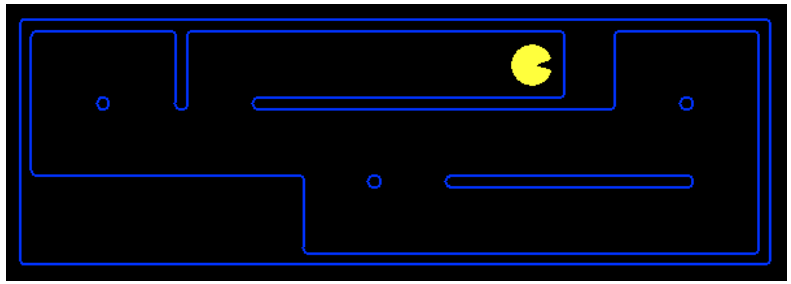
Video of Demo Reflex Optimal

Video of Demo Reflex Odd

Planning Agents

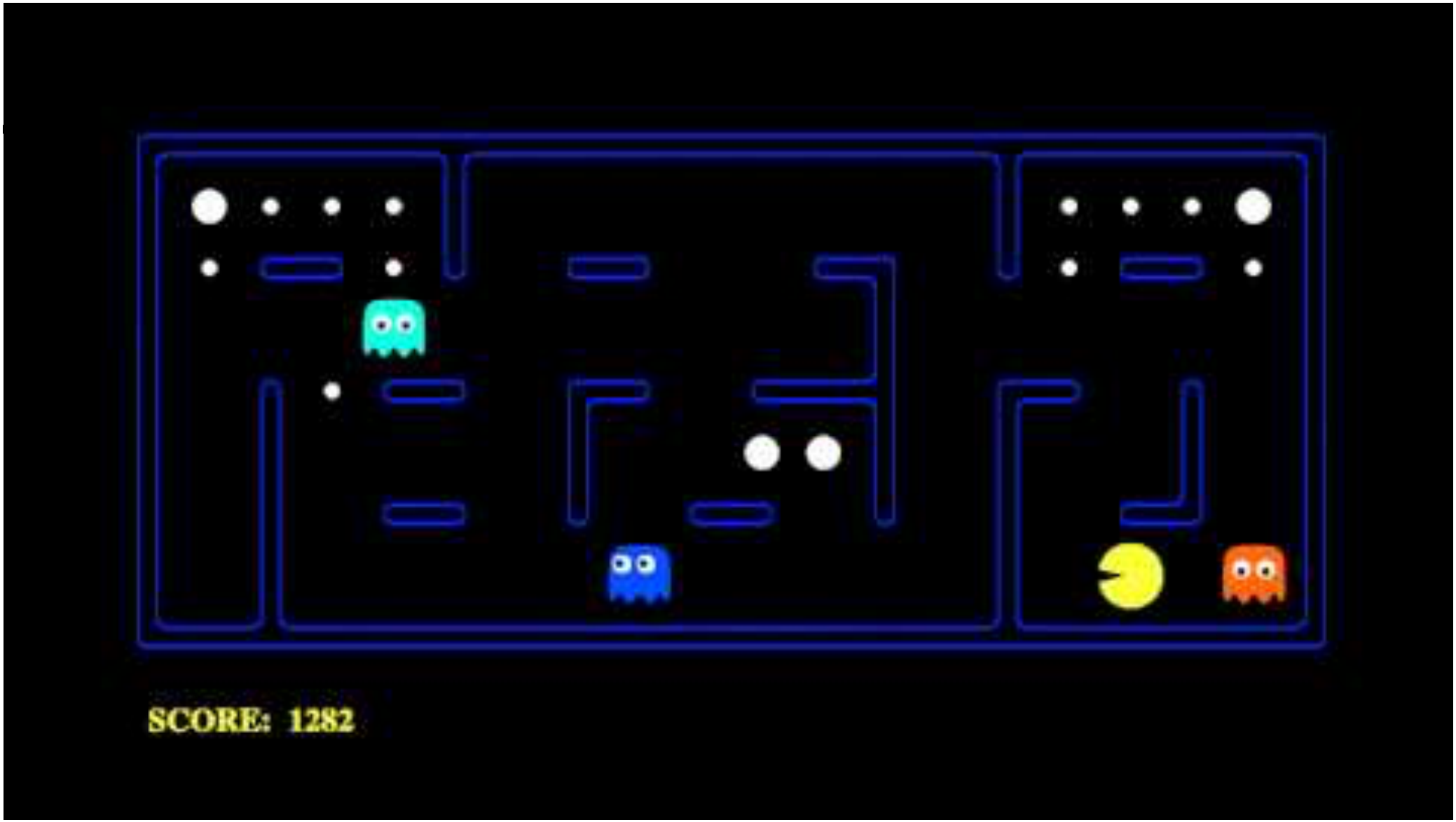
- Planning agents:
  - Ask “what if”
  - Decisions based on (hypothesized) consequences of actions
  - Must have a model of how the world evolves in response to actions
  - Must formulate a goal (test)
  - Consider how the world WOULD BE
- Optimal vs. complete planning
- Planning vs. replanning





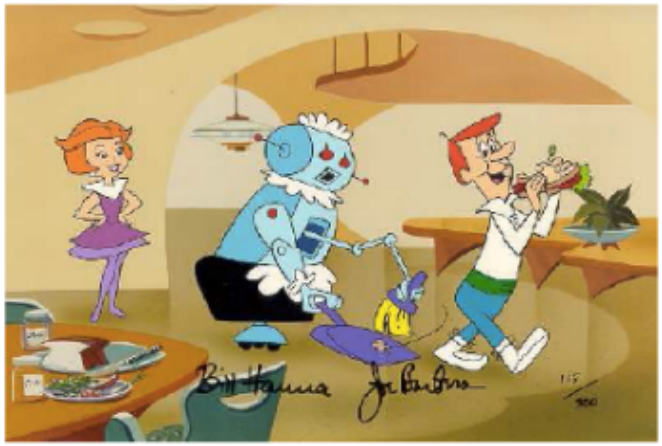
Video of Demo Re-planning

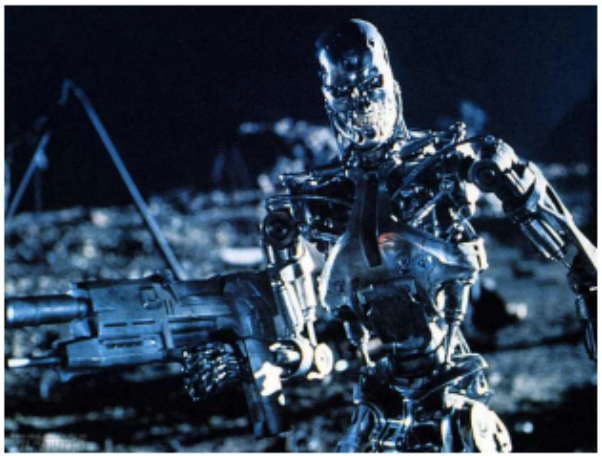
Video of Demo Mastermind


SCORE: 1282

Ethics and Implications

- Robust, fully autonomous agents in the real world
- What happens when we achieve this goal?

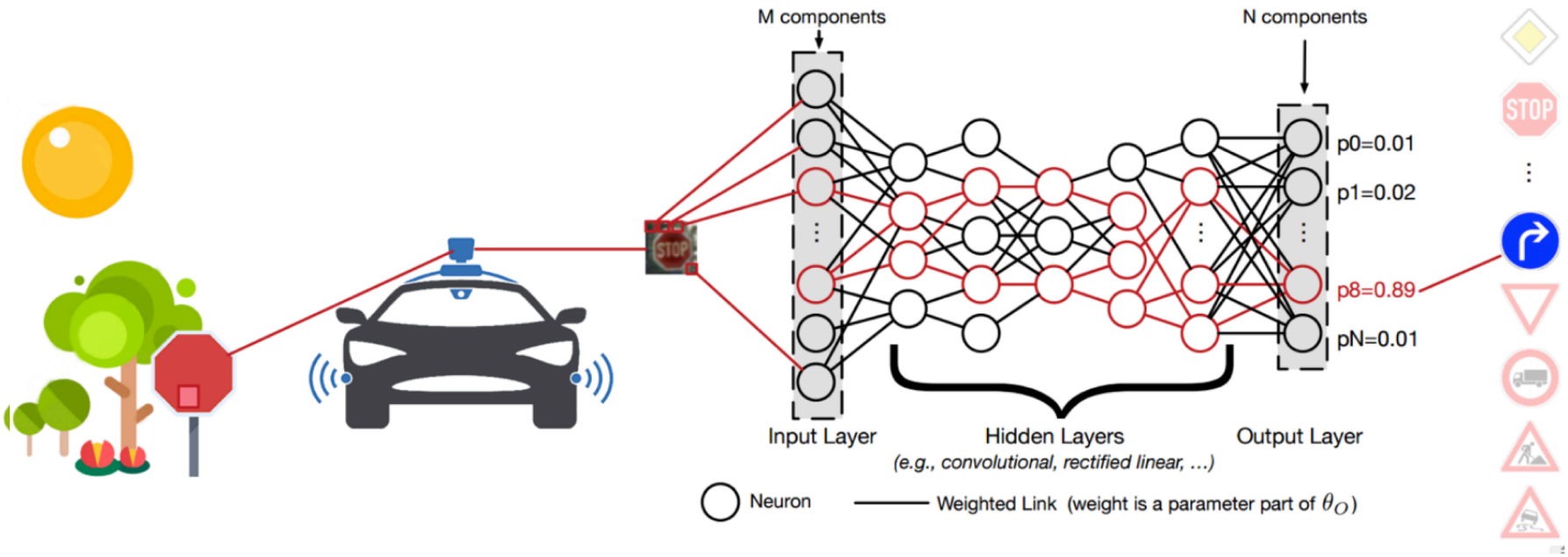




Ethics and Implications

- Who is liable if a robot driver has an accident?
- What will we do with super-intelligent machines?
- Would such machines have conscious existence? Rights?
- Can human minds exist indefinitely within machines (in principle)?

Importance of defending against adversarial attacks

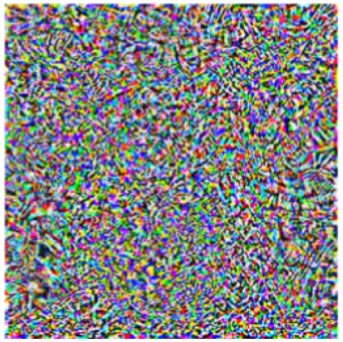


Adversarial examples

pig

x

+ 0.005 ×



η

=

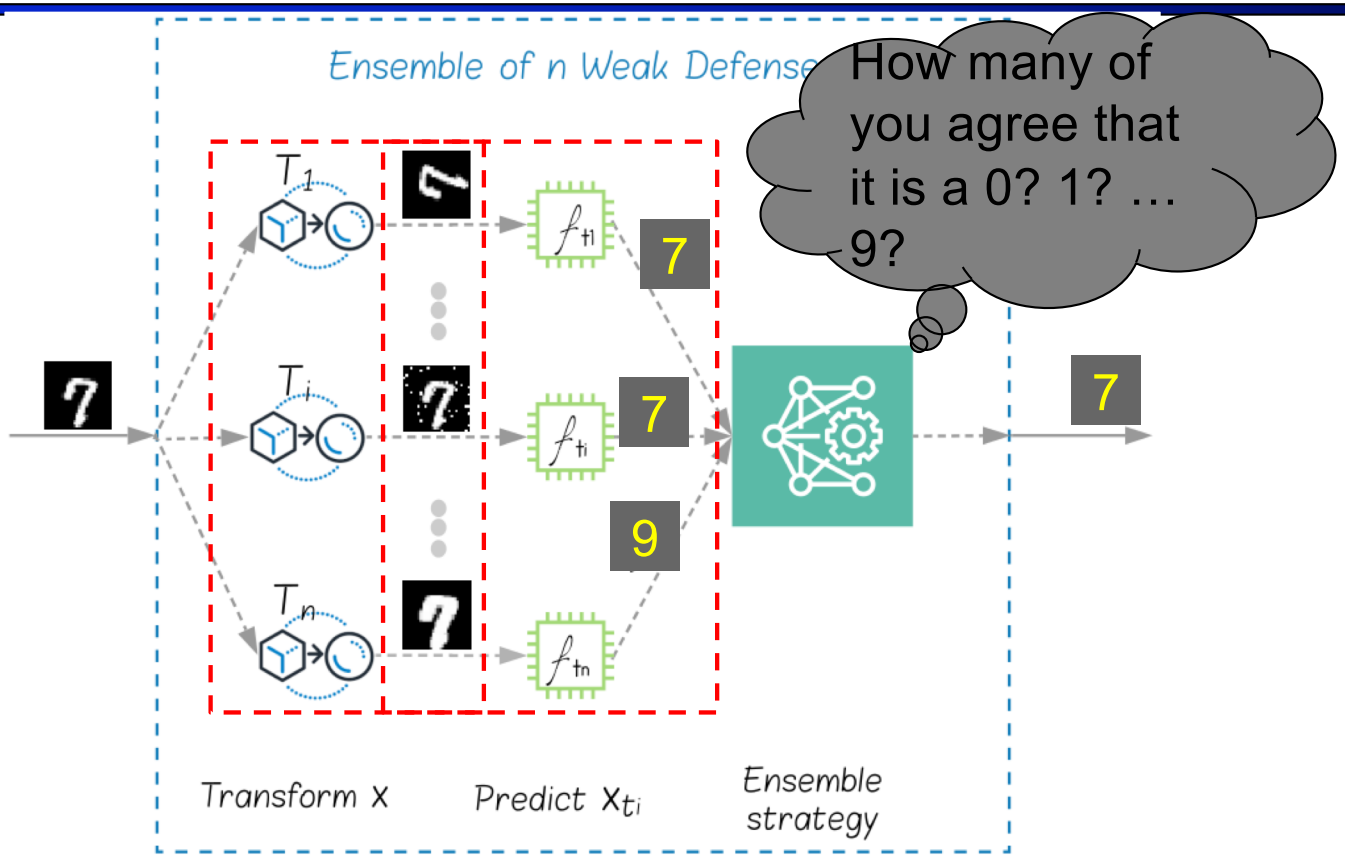
airliner

x’

FGSM

$\varepsilon \cdot \text{sign}(\nabla_x J(x, y))$

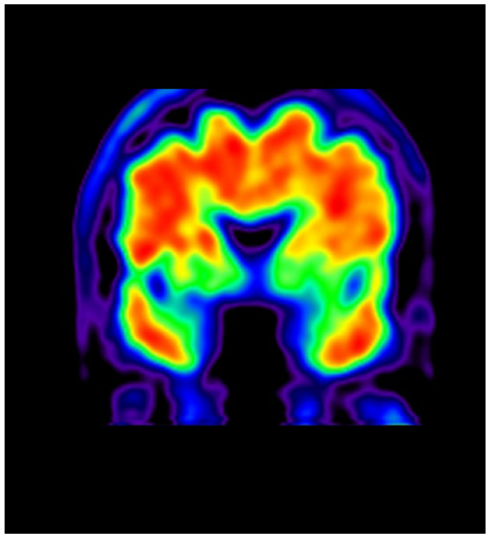
Approach: Ensemble of Many Weak Defenses

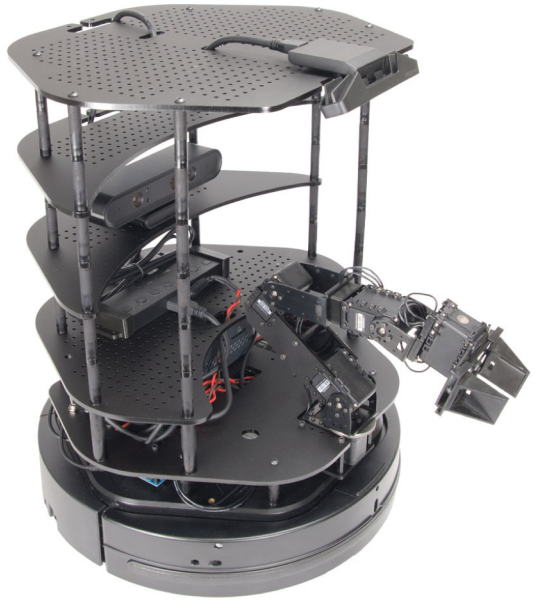


- 72 transformations
- 5 ensembles

Paper and Code

- **Paper**: https://arxiv.org/pdf/2001.00308.pdf
- **Code**: https://github.com/softsys4ai/athena





Announcements

- Project 0: Python Tutorial
  - I encourage team of 2 for doing the projects
  - I encourage pair programming
  - DO NOT SEPARATE THE TASKS BETWEEN EACH OTHER!
  - Due Monday at 11:59pm (pulse check to see you are in + get to know submission system)
- Homework 0: Math self-diagnostic
  - Important to check your preparedness for second half
  - You can either type in Latex or write by hand and upload the scanned PDF
- Project 1: Search
  - Will go out next week
  - Longer than most, and best way to test your programming preparedness
- Pinned posts on piazza


PAIR
PROGRAMMING