Artificial Intelligence

CS 165A

Oct 1, 2020

Prof. Yu-Xiang Wang

Computer Science

About myself

Yu-Xiang Wang 王宇翔

Eugene Aas Chair Assistant Professor

Department of Computer Science,
and Department of Statistics (By Courtesy)

Co-Director, Center for Responsible Machine Learning

UC Santa Barbara

Office: HFH-2121

E-mail: yuxiangw AT cs.ucsb.edu

Yu-Xiang is pronounced approximately as ['juː'ʃiːʌŋ], namely, *y~eu~ee – sh~ih~ah~ng* .

Research area: Statistical Machine Learning. Optimization, reinforcement learning, differential privacy, deep learning.

Short biography:

China => Singapore
⇒ PhD from Carnegie Mellon University
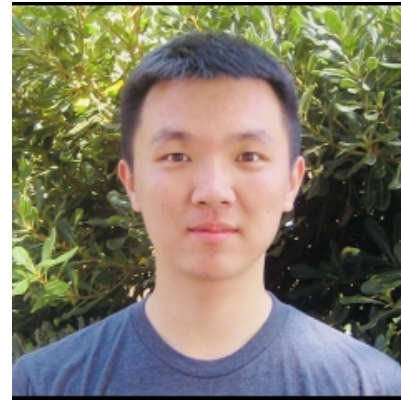⇒ Scientist at Amazon AI
⇒ Professor at UCSB

Homepage:
https://cs.ucsb.edu/~yuxiangw/

Call me: Yu-Xiang or Professor Wang.   **Don't** call me: Wang, Yu, Professor.

Instruction team


- TA: Xuandong (Dylan) Zhao

 xuandongzhao@ucsb.edu

- TA: Kaiqi Zhang

  kzhang70@ucsb.edu

- TA: Benu Changmai

  benu@ucsb.edu

- Reader: Yuxiao (Monica) Zhang

  yuxiao_zhang@ucsb.edu

Artificial Intelligence

- AI in the media
  - Popular movies
    - 2001: A Space Odyssey
    - Star Wars
    - The Terminator
    - The Matrix
    - Artificial Intelligence: AI
    - and many …
  - Popular press, novels
- Often portrayed as
  - A property of evil computers
  - Computers doing impossible things
- Public view
  - Books and movies have inspired many AI researchers
  - Books and movies have raised the public's expectations
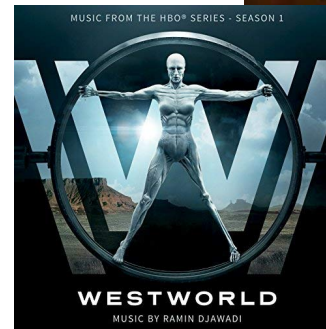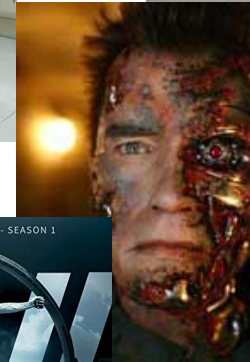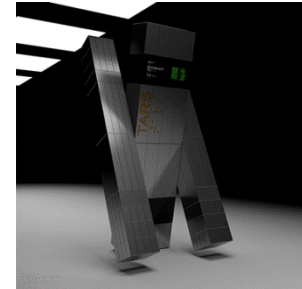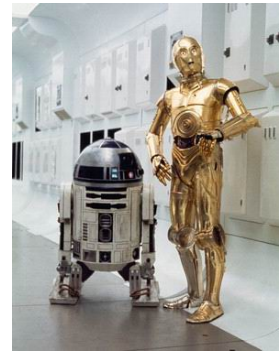
Artificial Intelligence
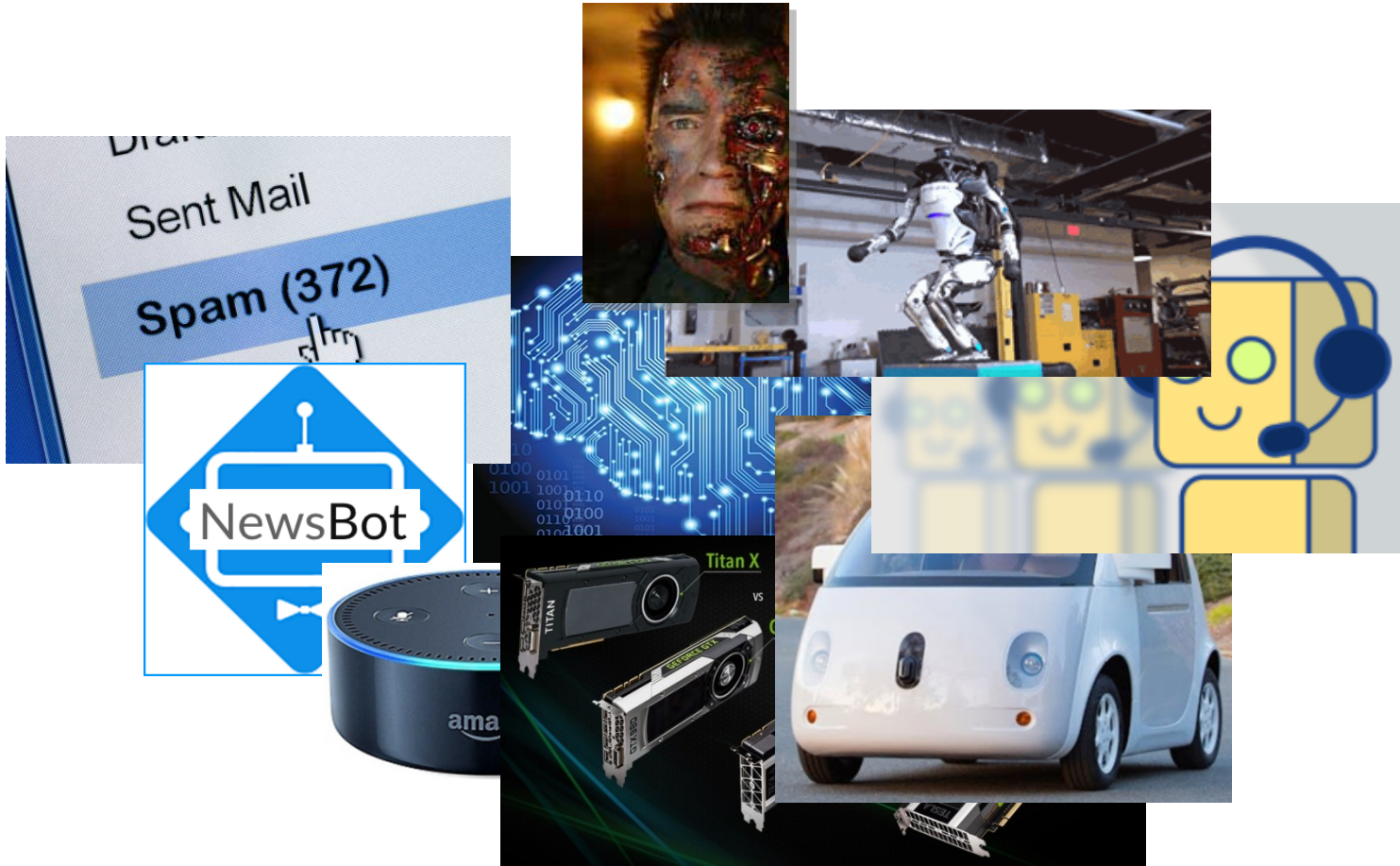
AI Is Transforming The Industries

☐ AI has transformed the IT industry
  - ☐ Search Engine
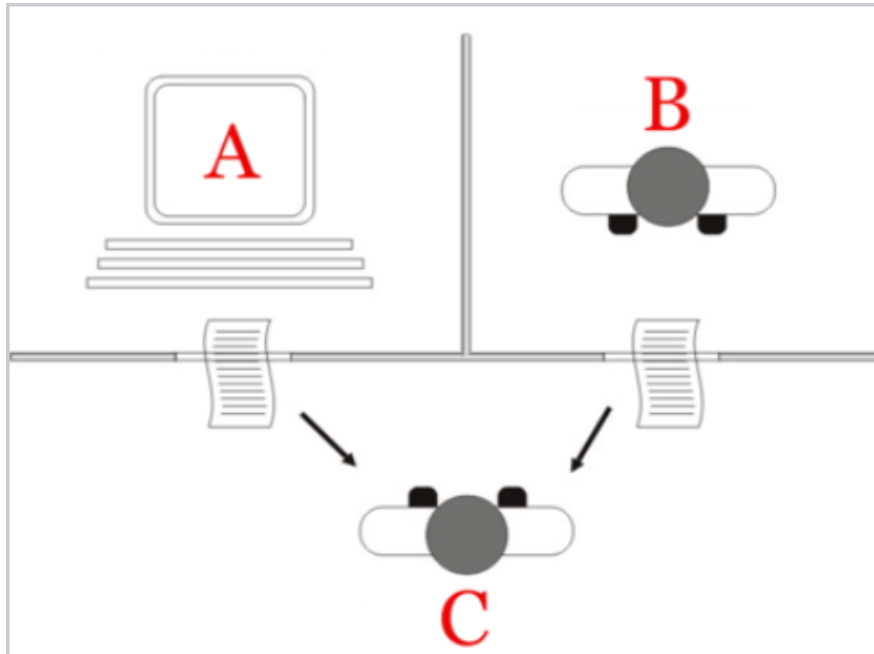 - ☐ Speech Recognition
  - ☐ Machine Translation
  - ☐ Recommendation

☐ AI is transforming other industries
  - ☐ Transportation
  - ☐ Healthcare
  - ☐ Finance
  - ☐ Insurance, Law, HR, Travel, Media, …
  - ☐ Semiconductor / Microprocessors

Artificial Intelligence

☐ Turing Test

The "standard interpretation" of the Turing Test, in which player C, the interrogator, is given the task of trying to determine which player – A or B – is a computer and which is a human. The interrogator is limited to using the responses to written questions to make the determination. (wiki)
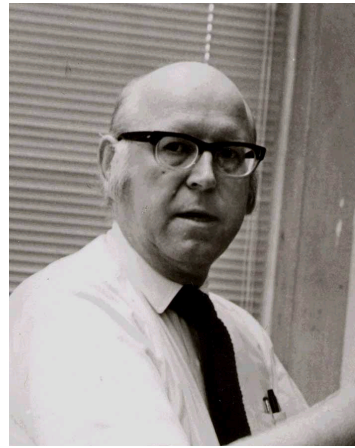
☐ Artificial Intelligence is an ultimate dream of computer science

Supplementary reading: Turing (1950) "Computing Machinery and Intelligence" [here].
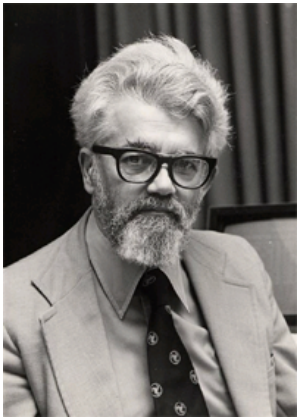
John McCarthy

Herb Simon

Allen Newell
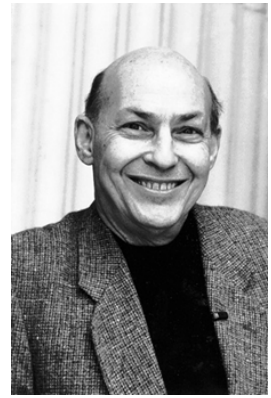
Marvin Minsky

Seminal event for AI as a field, in 1956:
The **Dartmouth Summer Research Conference on Artificial Intelligence**

"We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of **the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it**. 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 in one or more of these problems if a carefully selected group of scientists work on it together for a summer."

Over-Confidence

Herbert Simon (1957)

It is not my aim to surprise or shock you - but the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create, ... More precisely: within 10 years a computer would be **chess champion**, and an important new **mathematical theorem** would be proved by a computer.

Both of these milestones have now been achieved by computers, but each happened much later, more than **10** years:

1. Four color theorem proven in 1976 by Kenneth Appel and Wolfgang Haken

2. Deep Blue *chess* computer *defeated* Kasparov in 1997

History

- 1956: The **Dartmouth Summer Research Conference on Artificial Intelligence**, Seminal event for AI as a field

- 1976: **Four color theorem** was proven by Kenneth Appel and Wolfgang Haken

- 1997: Deep Blue **chess** computer *defeated* Kasparov (10^{50})
  - "Saying Deep Blue doesn't really think about chess is like saying an airplane doesn't really fly because it doesn't flap its wings."   --Drew McDermott

GOFAI: Good-Old-Fashioned Artificial Intelligence

- Since 2012, we saw breakthroughs in **Speech Recognition, Image Recognition, Machine Translation, Autonomous Car,** etc.
- 2015, **CNN+ReLU** outperforms humans in image recognition.
- 2016, **AlphaGo** beat a 9-dan professional (Lee Sedol) (10^{170})
- 2020, **GPT-3** is doing well simultaneously in many natural language tasks.

(Statistical) Machine Learning

Question Answering

- IBM Watson / Jeopardy!

- Apple Siri
  - The application uses a natural language user interface to answer questions, make recommendations, and perform actions by delegating requests to a set of Web services.

- Amazon Echo / Google Home
  - Associate the natural language user interface with a speaker.

- Google demo of restaurant reservation
  - https://www.youtube.com/watch?v=eHvyKWS5QzE

Question Answering

Apple Siri

- "Hey Siri, what's the best sushi place in town?"

- "Find videos I took at Iva's birthday party"

- "What movies are playing today?"

- "Wake me up at 7 AM tomorrow"

Google Search

- "Who was U.S. President when the Angels won the World Series?"

- But, "Who was the UK prime minister when the Angels won the World Series?"

Can Computers Understand?

- **They ate the pie with ice cream.**

- **They ate the pie with rhubarb.**

- **They ate the pie with paper plates.**

- **They ate the pie with cold milk.**

- **They ate the pie with friends.**

- **They ate the pie with dinner.**

- **They ate the pie with enthusiasm.**

- **They ate the pie with spoons.**

- **They ate the pie with napkins.**

from Dr. Douglas Lenatand Dr. Michael Witbrock   13

State-of-the-art AI for Question and Answers

- GPT-3 from OpenAI

- See example here: Giving GPT-3 a Turing Test

Q: What is your favorite animal?
A: My favorite animal is a dog.

Q: Why?
A: Because dogs are loyal and friendly.

Q: What are two reasons that a dog might be in a bad mood?
A: Two reasons that a dog might be in a bad mood are if it is hungry or if it is hot.

Q: How many eyes does a giraffe have?
A: A giraffe has two eyes.

Q: How many legs does a frog have?
A: A frog has four legs.

Q: Are there any animals with three legs?
A: No, there are no animals with three legs.

Q: Why don't animals have three legs?
A: Animals don't have three legs because they would fall over.

A lot more that GPT-3 is able to do:

- Generate well-written text

- Generate code for website layout, format charts / figures.

- Parse spreadsheets

- Translations: English ⇔ Chinese, Legal ⇔ Plain English

- Check out this video: https://youtu.be/_x9AwxfjxvE?t=219

Other cool demos of AI

- Learning to walk / run with Reinforcement Learning
  - https://youtu.be/gn4nRCC9TwQ

- Learning to generate a realistic face
  - https://www.thispersondoesnotexist.com/
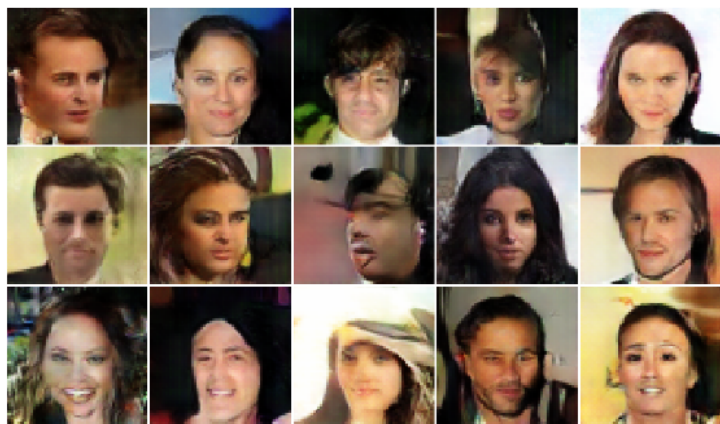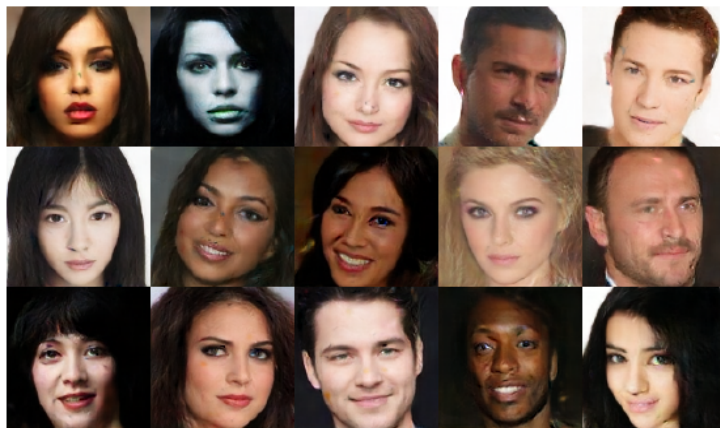
EBGAN (Zhao et al, 2016)   BEGAN (Berthelot et al, 2017)

- Learning to play "hide and seek" in creative ways:
  https://youtu.be/Lu56xVlZ40M?t=50

"We are in the middle of a new industrial revolution." --- Bernhard Schölkopf

- First two revolutions: our ability to harness energies
  - Coal / Fossil fuel
  - Electricity

- Third revolution: our ability to harness information
  - Digital revolution from 1950s.
  - Big Data revolution, most recently, AI revolution

- Fundamental societal changes
  - Job loss
  - Addiction, polarization
  - Privacy, fairness
  - Synthesize voice, face of any one
  - AI Weapons?

Further reading: https://arxiv.org/pdf/1911.10500.pdf

The exponential growth of AI as a field

Attendance at large conferences (1984-2019)
Source: Conference provided data.

Fig. 2.1a
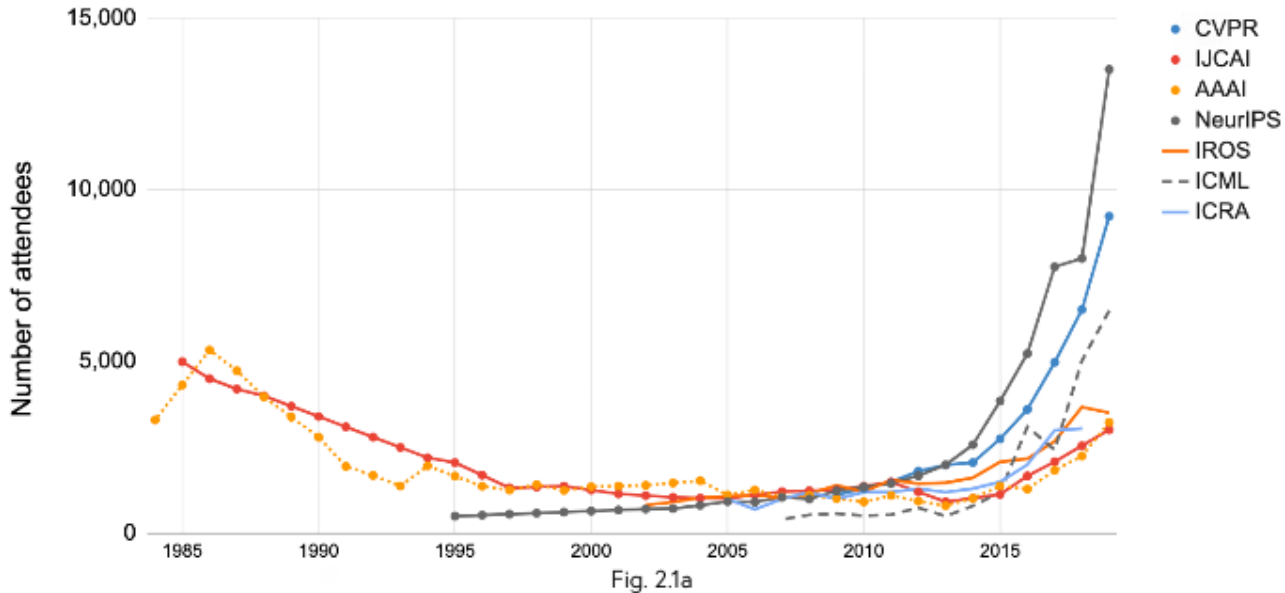
NIPS @NipsConference · 4m
#NIPS2018 The main conference sold out in 11 minutes 38 seconds
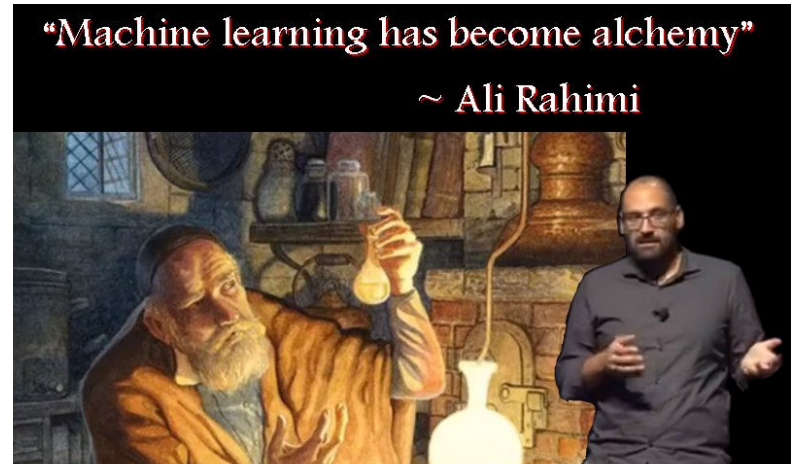
💬 3   🔁 21   ♡ 25   ⤳

"For those debating ML hype, getting a ticket to a ML conference is now more challenging than a Taylor Swift conference or a Hamilton showing."

– Kaggle CTO Ben Hamner

18

Is it all just hype?

- AI research is supported by rigorous mathematical foundation.
  - Although… there's caveat.
  - Theory is always a bit behind.

"Machine learning has become alchemy"
~ Ali Rahimi

- AI Researchers know the limitations of their algorithms
  - (although.... sometimes it's easy to go over the top)

Andrew Ng
@AndrewYNg

Following

Should radiologists be worried about their jobs? Breaking news: We can now diagnose pneumonia from chest X-rays better than radiologists.
stanfordmlgroup.github.io/projects/chexn…

3:20 PM - 15 Nov 2017 from Mountain View, CA

1,440 Retweets 2,401 Likes

114   1.4K   2.4K

19

Regardless, it is an exciting time!

- The field of AI is gathering talents from CS, Stats, Math, Physics, Engineering, Business like never before!

  - 70% PhD applicants to UCSB CS want to specialize in AI / ML!

- We are training the first generation of Da Vinci-like engineers / researchers in large scale
  - Who can swiftly leverage multiple branches of mathematics
  - Understand computing / computer systems / write efficient codes
  - Understand statistics / causality / handles big data
  - Inspired by ideas from neural science / biological world.

Goals of this course

- To teach you some ideas of AI

- To introduce you to a set of key techniques and algorithms from AI

- To get you thinking about how AI can be applied to a variety of real problems

- To help you understand what's hard in AI

- To know where to find additional materials when needed in the future

- **It is not about vision, natural language processing, machine learning, AlphaGo,…it is an entry level course**

The schedule of this course

Week	Topic
1	Course Overview & Intelligent Agents
2	Machine Learning
	Machine Learning
2	Machine Learning
	Probabilistic Graphical Models
3	Probabilistic Graphical Models
	Search: Problem solving with search
4	Search: Search algorithms
	Search: Minimax search and game playing
5	Midterm Review
	Midterm (take-home)
6	RL: Intro / Markov Decision Processes
	RL: Solving MDPs
7	RL: Bandits and Exploration
	RL: Reinforcement Learning Algorithms
8	RL: Reinforcement Learning Algorithms
	Logic: Propositional Logic
9	Logic: First order Logic
	Responsible AI
10	Responsible AI
	Final Review
11	Final Exam (take-home)

Machine Learning

Probabilistic Reasoning

Search

Reinforcement Learning

Logic

Responsible AI

It will be tough and time-consuming to many!

- You need to use math:
  - Calculus, Linear Algebra, Probabilities.
  - Know how to do complexity analysis in Big O notation.

- You need to be able to code:
  - Mainly in Python (numpy, scipy)
  - Data structures: trees, graphs, hashtables, sparse matrices, etc…

- Homework assignments will complement the materials in the lectures.
  - That will help you gain deeper understanding of the topics.

- Formal prerequisite:
  - CS20,CS50, CS130A, PSTATS 120A.

It will be "worth it".

- Like Jedi training.

  - The information can be presented to you, but you can only learn if you are actively looking and thinking.
  - I learn a lot more from doing homeworks on my own than just from the lectures.

- We (your instructor, TAs and peers) are ready to help you!
  - Discussion sessions on Friday, Office hours, Piazza

Expectation of you

- Academic Integrity
  - Homeworks are individual work
  - Discussion with your peers on the high level is fine.
  - Finding solutions to the exact questions on the internet is not OK.
  - Sharing equations, sharing code / pseudo-code are not OK.
  - If not sure if anything is a violation of the course policy, check with the TAs on Piazza.

- Creating a respectful learning environment for all
  - Be respectful to your peers and instructors
  - Harassment and discrimination of any form will not be tolerated.

* More details on these in the [syllabus].

Course information

- Web sites
  - http://www.cs.ucsb.edu/~yuxiangw/classes/CS165A-2020fall/
    - Announcements, syllabus, schedule, lecture notes, assignments, related links
- Discussion on Piazza: piazza.com/ucsb/fall2020/cs165a
- Homework submission on Gradescope
  - You will be added directly to it.
- Textbook
  - Stuart Russell and Peter Norvig, Artificial Intelligence: A Modern Approach, Prentice Hall, **Third Edition**, 2010
  - Other reference books will be referred to. (e.g., Sutton and Barto will be our main text for the RL part of the class)

Mode of Remote Instruction

- Synchronous Lectures (Recorded)
  - A TA will moderate the questions / address easy questions via chat
  - I will address questions not answered by Tas periodically during lectures.

- Asynchronous Discussion Sections (Recorded)
  - Videos will be released to you on Piazza
  - TAs will be at the Wednesday discussion class to answer questions and provide live interactions with students

- Office Hours from me and the TAs (not recorded)

- Asynchronous Q&A: Piazza

Workload and Grade

- 4 Homework Assignments, 2 Exams

- 60% Homework assignments
  15% Take-home midterm exam
  25% Take-home final exam

- x% bonus points:
  - For solving bonus questions in homeworks
  - For class participation

- Late Policy:
  - Each student has **a total of 3 late days** without penalties.
  - When your late days are used up, you may still submit. When submitted within a week of the due date you will receive 75% of the credits. Beyond that only 50%.
  - No late submission for the take-home exams are allowed.

Office Hours

- Instructor Office Hour: Thursday 2:00-3:00pm

- TA1: Xuandong Zhao, Office Hour:
- TA2: Kaiqi Zhang, Office Hour:
- TA3: Benu Changmai, Office Hour:

Single zoom link for all these OHs: announced on Piazza.

- **Ask your questions on Piazza!**

29

How to Maximize Your Learning

- Read textbook chapter (or other resources) before class

- Attend lectures and study the slides

- Read related chapters before the class

- Do your own "homework", self-study

- Actively participate discussions on Piazza

- No last minute work

Useful resources

- Appendix A in the AIMA book for a mathematical review.

- Standard mathematical notations used in ML / AI:
  http://www.deeplearningbook.org/contents/notation.html

- Chapter 2, Chapter 17-18 of the D2L book: http://d2l.ai/

- Probability / Linear Algebra review (shared with you on Piazza, will be covered next Wednesday)

- If you are having trouble understanding the probability part of above
  – Here is a more straightforward probability review

Remainder of the lecture: Intelligent agents

- Formally setting up the problem
  - Intelligent Agents
  - The Model-Inference-Learning Paradigm
  - Task environment
  - Model vs. reality
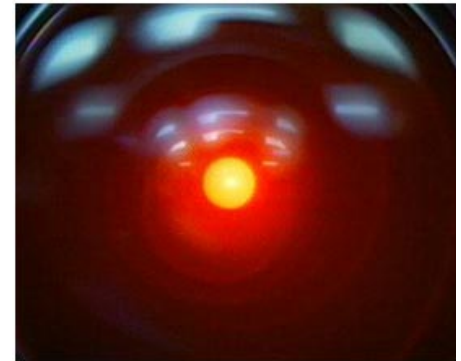
What's an Agent?

"An intelligent agent is an entity capable of combining cognition, perception and action in behaving autonomously, purposively and flexibly in some environment."

- Possible properties of agents:
  - Agents are **autonomous** – they act on behalf of the user
  - Agents can **adapt** to changes in the environment
  - Agents don't only act **reactively**, but sometimes also **proactively**
  - Agents have **social ability –** they communicate with the user, the system, and other agents as required
  - Agents also **cooperate** with other agents to carry out more complex tasks than they themselves can handle
  - Agents **migrate** from one system to another to access remote resources or even to meet other agents

33

Our view of AI

- AIMA view – AI is building intelligent (rational) agents
  - Principles of rational agents, and
  - Models/components for constructing them

- Rational = "Does the right thing" in a particular situation
  - Maximize *expected* performance (not *actual* performance)

- So a rational agent does the "right" thing (at least tries to)
  - Maximizes the likelihood of success, given its information
  - How is "the right thing" chosen?
    - Possible actions (from which to choose)
    - Percept sequence (current and past)
    - Knowledge (static or modifiable)
    - Performance measure (*wrt* goals – defines success)

Our model of an agent

- An agent **perceives** its environment, **reasons** about its goals, and **acts** upon the environment
  - Abstractly, a function from percept histories to actions

$$f : P^* \rightarrow A$$

- Main components of an agent
  - Perception (sensors)
  - Reasoning/cognition
  - Action (actuators)

- Supported by
  - knowledge representation, search, inference, planning, uncertainty, learning, communication….

Our view of AI (cont.)

- So this course is about <u>designing rational agents</u>
  - Constructing f
  - For a given class of environments and tasks, we seek the agent (or class of agents) with the "best" performance
  - Note: Computational limitations make complete rationality unachievable in most cases

- In practice, we will focus on <u>problem-solving techniques</u> (ways of constructing f), not <u>agents</u> per se

Ideal Rational Agent

- In other words…

  "For each possible percept sequence, an ideal rational agent should do whatever action is expected to maximize its performance measure, on the basis of the evidence provided by the percept sequence and whatever built-in knowledge the agent has."
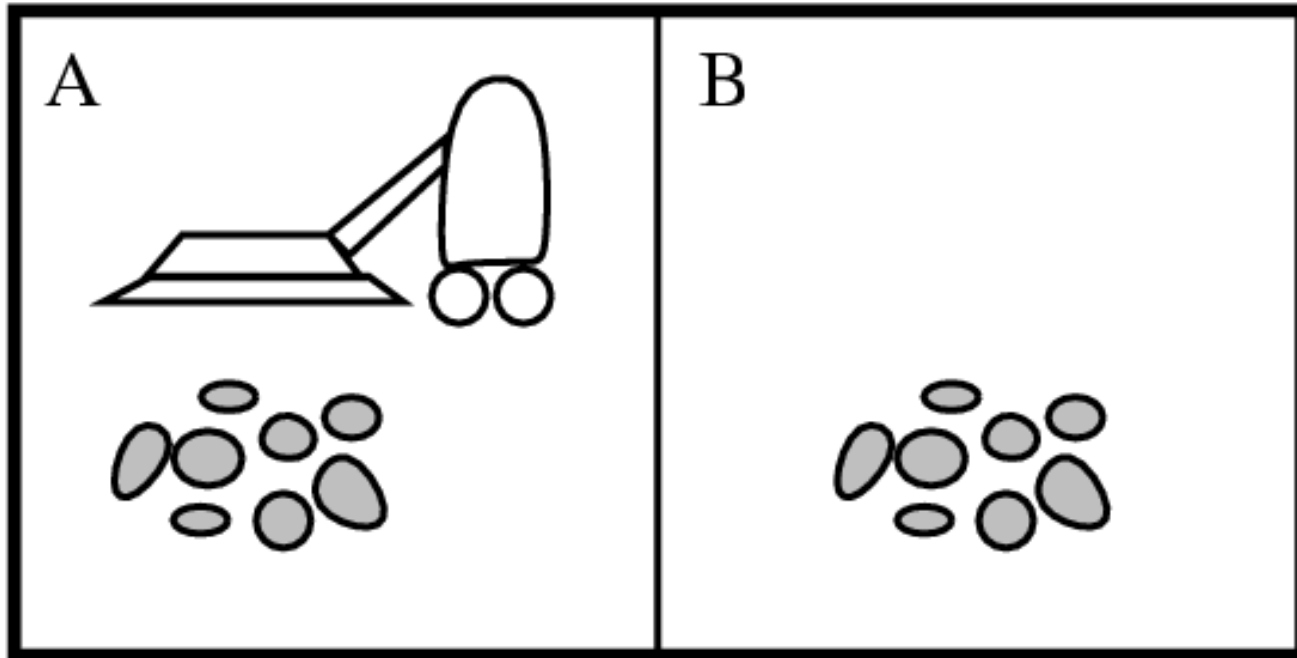
Note that:

Rational ≠ Omniscient
Rational ≠ Clairvoyant
Rational ≠ Successful

Describing the Task Environment
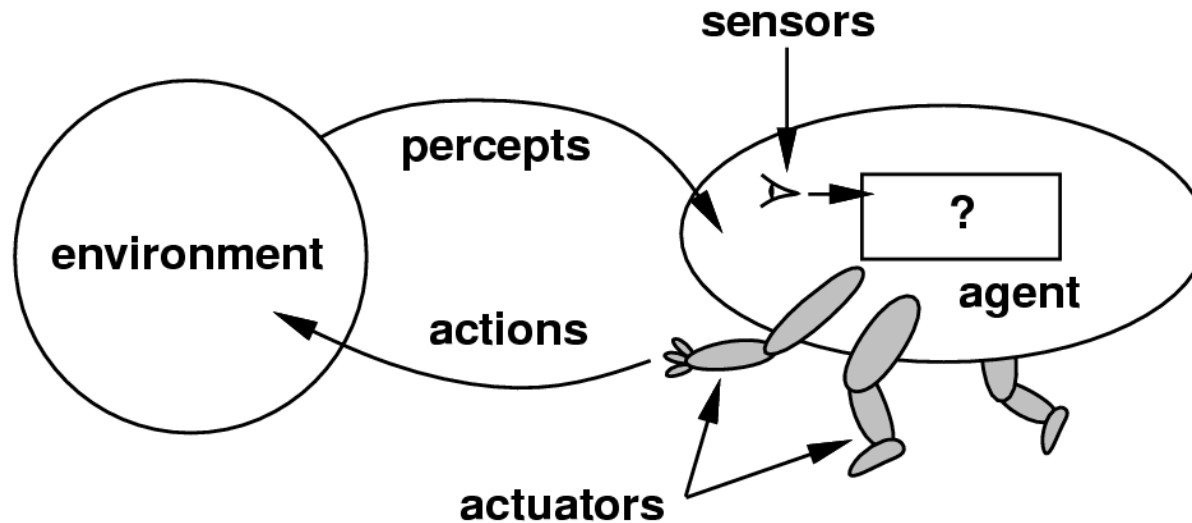
- **PEAS** – Performance measure, Environment, Actuators, Sensors
  - Goals may be explicit or implicit (built into performance measure)

- Not limited to physical agents (robots)
  - Any AI program

The Vacuum World: 3 min discussion

Performance measure, Environment, Actuators, Sensors

The Vacuum World

- Performance (P)
  - Keep world clean
  - Possible performance measures

- Environment (E)
  - Location
  - Cleanliness

- Three actions (A)
  - Move right
  - Move left
  - Remove dirt

- Sensed information (percepts) of environment (S)
  - Two locations
   - Left
    - Right
  - Two states
    - Dirty
    - Clean

PEAS Descriptions

Agent Type	P	E	A	S
Medical diagnosis system	Healthy patient, minimize costs	Patient, hospital	Questions, tests, treatments	Symptoms, findings, patient's answers
Satellite image analysis system	Correct categorization	Images from orbiting satellite	Print a categorization of scene	Pixels of varying intensity, color
Part-picking robot	Place parts in correct bins	Conveyor belt with parts	Pick up parts and sort into bins	Pixels of varying intensity
Refinery controller	Maximize purity, yield, safety	Refinery	Open, close valves; adjust temperature	Temperature, pressure readings
Interactive English tutor	Maximize student's score on test	Set of students	Print exercises, suggestions, corrections	Typed words

Environments

- Properties of environments
  - Fully vs. partially observable
  - Deterministic vs. stochastic
  - Episodic vs. sequential
  - Static vs. dynamic
  - Discrete vs. continuous
  - Single agent vs. multiagent

- The environment types largely determine the agent design

- The real world is partially observable, stochastic, sequential, hostile, dynamic, and continuous
  - Bummer…

New paradigm: Modeling-Inference-Learning

Modeling

Inference
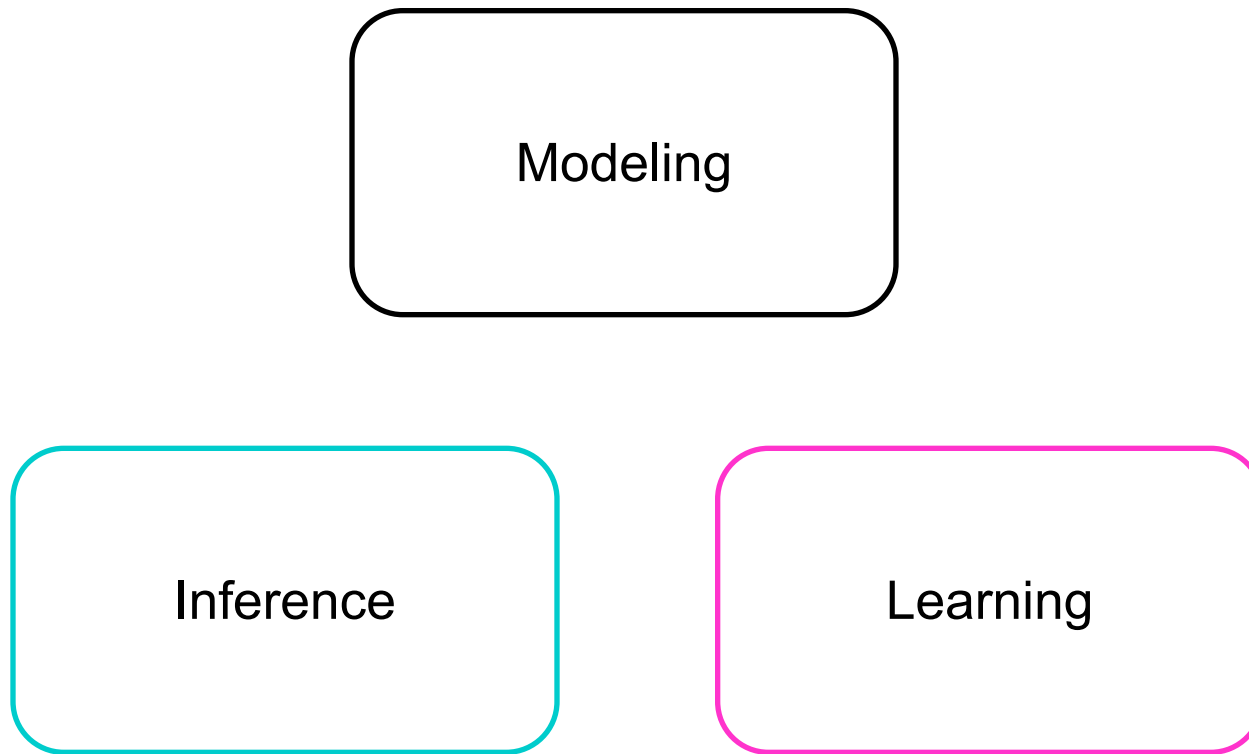
Learning

(Idea and example taken from Percy Liang's teachings)

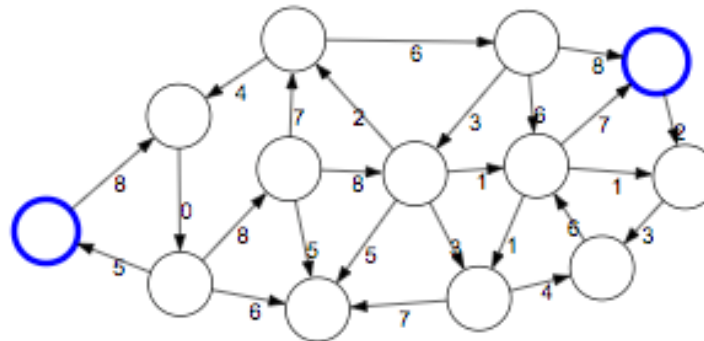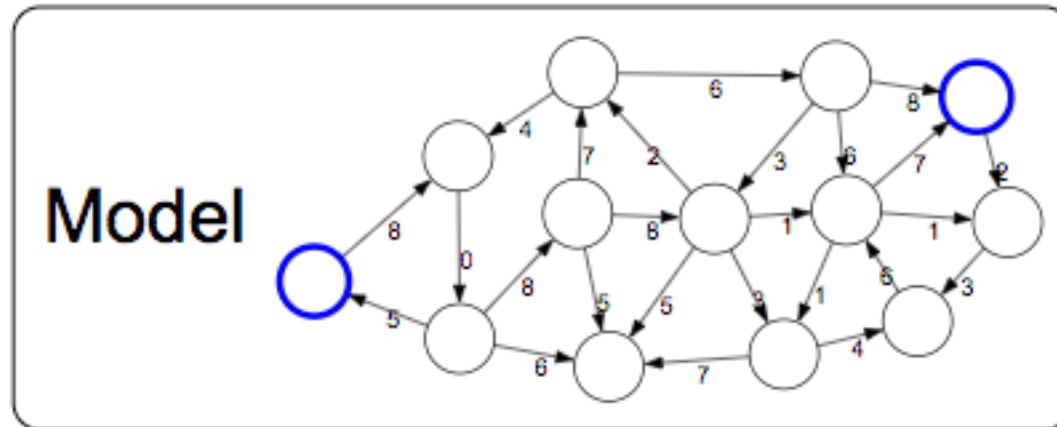# Paradigm: Modelling

Real world

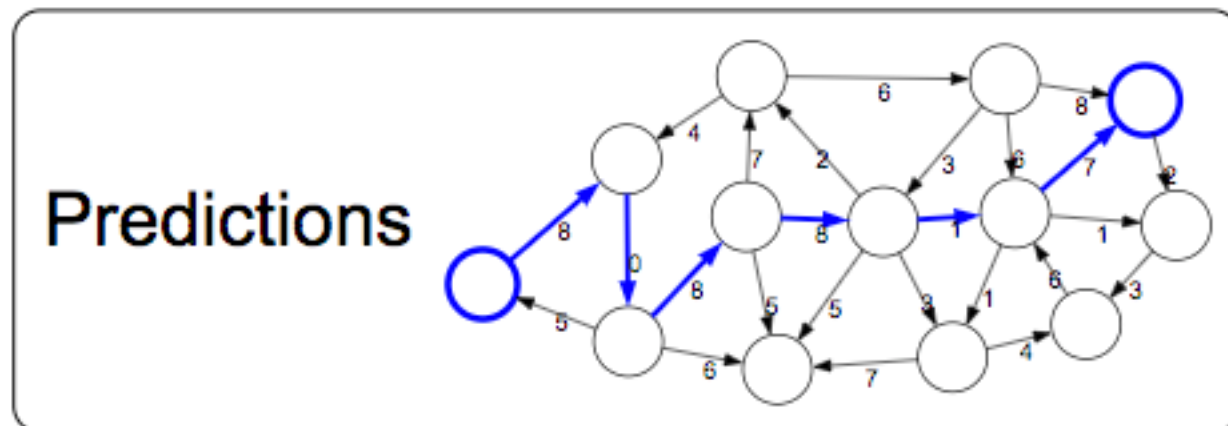**Modeling**

Model

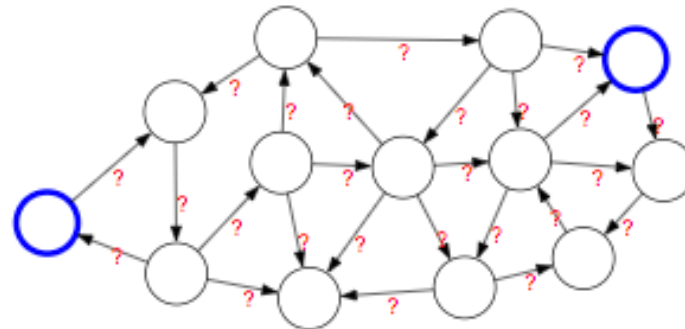# Paradigm: Inference

Model

Inference

Predictions
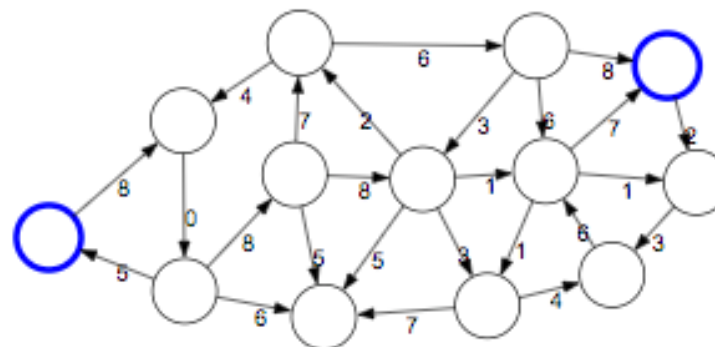
Paradigm: Learning

Model without parameters

+data

Learning

Model with parameters

Structure of the course

Probabilistic Graphical Models / Deep Neural Networks

Classification / Regression Bandits	Search game playing	Markov Decision Processes Reinforcement Learning	Logic, knowledge base Probabilistic inference
Reflex Agents	**Planning Agents**		**Reasoning agents**

Low-level intelligence

High-level intelligence

Machine Learning

(Again this idea is adapted from Percy Liang's teachings)

Generic Agent Program

- Implementing $f : P^* \rightarrow A$ ...**or**... $f(P^*) = A$
  - Lookup table?
  - Learning?

Knowledge, past percepts, past actions

```
function SKELETON-AGENT(percept) returns action
    static: memory, the agent's memory of the world

    memory ← UPDATE-MEMORY(memory, percept)
    action ← CHOOSE-BEST-ACTION(memory)
    memory ← UPDATE-MEMORY(memory, action)
    return action
```

e.g.,

Table-Driven-Agent

Add *percept* to percepts

LUT [percepts, table]
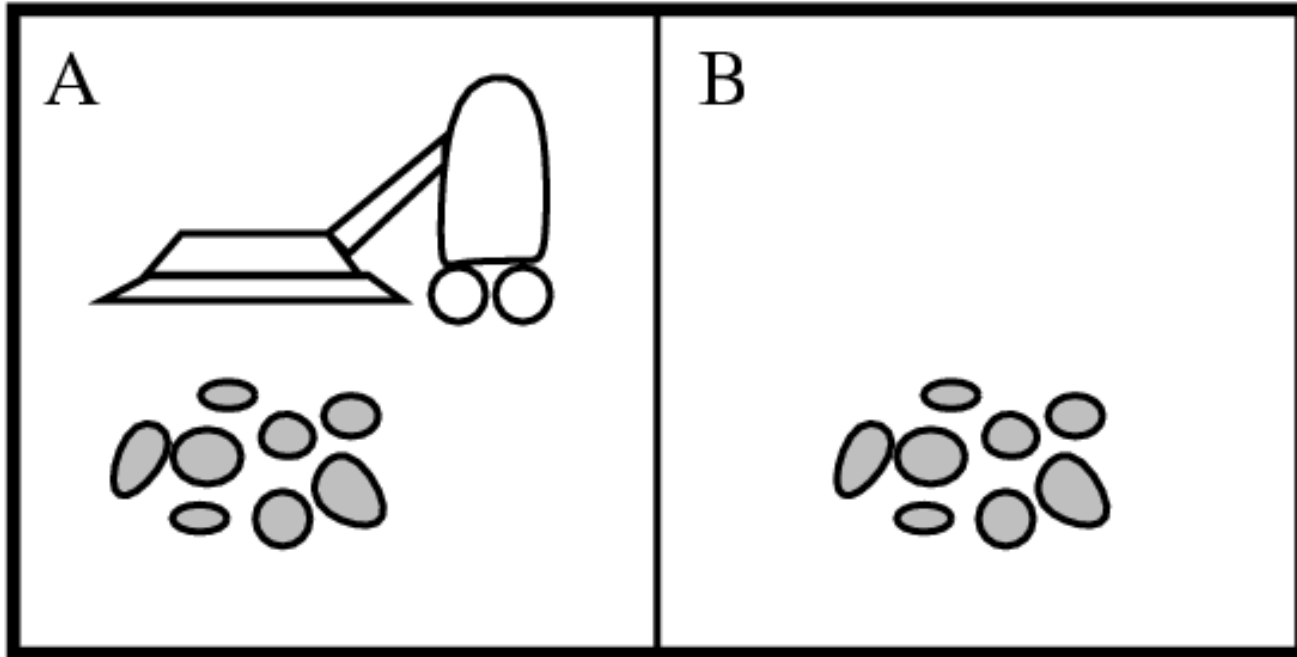
NOP

AIMA's categorization of agent programs

- Simple reflex agent

- Model-based reflex agent

- Goal-based agent

- Utility-based agent

- Learning agent

(Read more in Section 2.4 of the AIMA book.)

Potential mid-term questions:
1. Where do these agent fall under our new categorization?
2. What are these agent's "Modelling-Inference-Learning" components?

- Reflex, planning, reasoning?
- What is the model? Are there any learning components?

When to use which type of agent?

- Depends on the problem (task environment)
  - Stochastic/deterministic/stateful/adversarial …

- Depends the amount of data available
  - Often we need to learn how the world behaves

- Depends on the dimensionality of your observations

Solving the right problem approximately

vs

Solving an approximation of the problem exactly

"All models are wrong, but some are useful."

George Box
(1919 - 2013)

Next three lectures: machine learning

- Chapter 18 in AIMA (3rd Edition) (more extensive coverage in the 4th Edition)

- Chapter 1 in D2L: Intro to ML
  http://d2l.ai/chapter_introduction/index.html

- Chapter 1 in ESL: A more extensive introduction to ML

- Chapter 3,4,11 in D2L:
  - Chapter 3,4: From linear classification to neural networks
  - Chapter 11: Optimization in ML.

Final notes

- Submit (anonymous) feedbacks here: https://forms.gle/CkfYBNnjR7durZ7g6

  - What do you like / dislike about the course so far?
  - Are you following the lectures?
  - Am I going too fast or too slow?
  - Which AI topic from the course-schedule that you are most excited about?

- Homework 1 will be released tonight! Start early!

- If you haven't yet, please answer the new student poll: https://forms.gle/iFoR7R27V5AQGrKX6