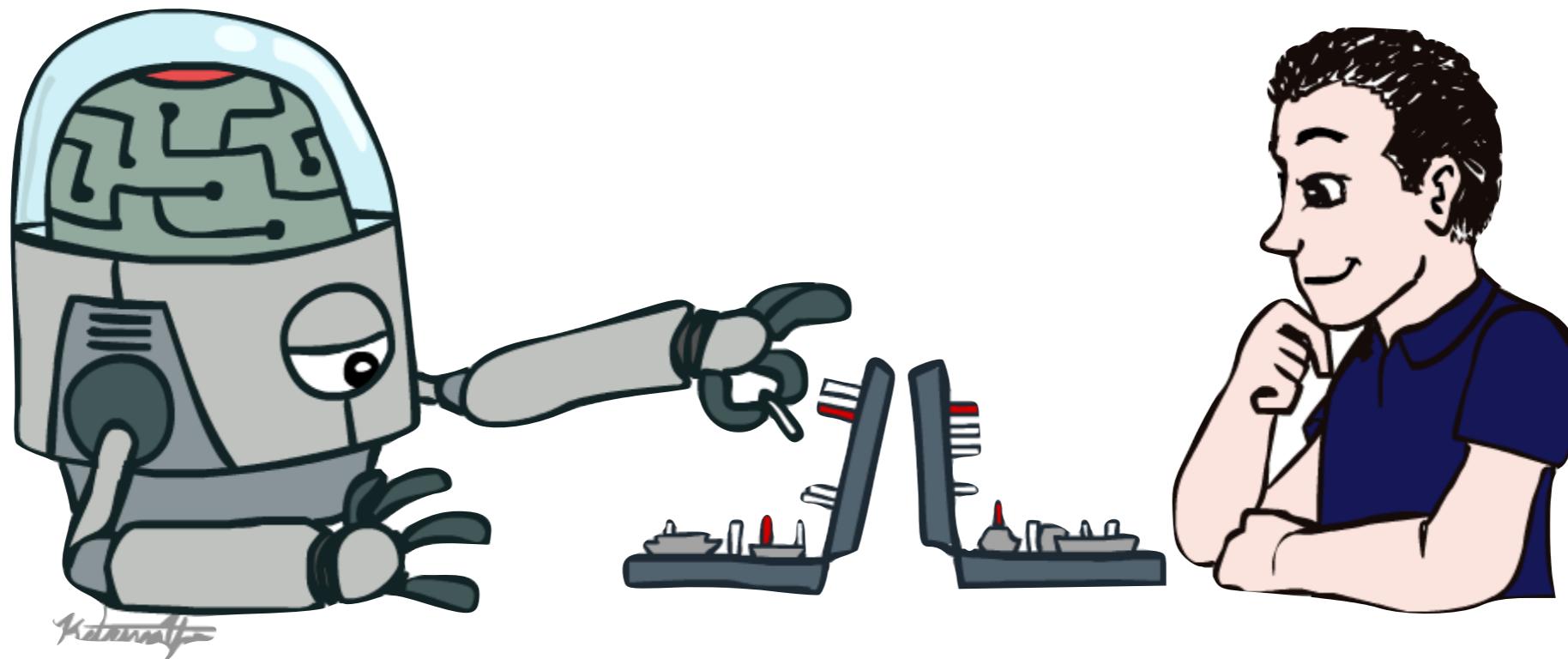


# Ve492: Artificial Intelligence

## Introduction



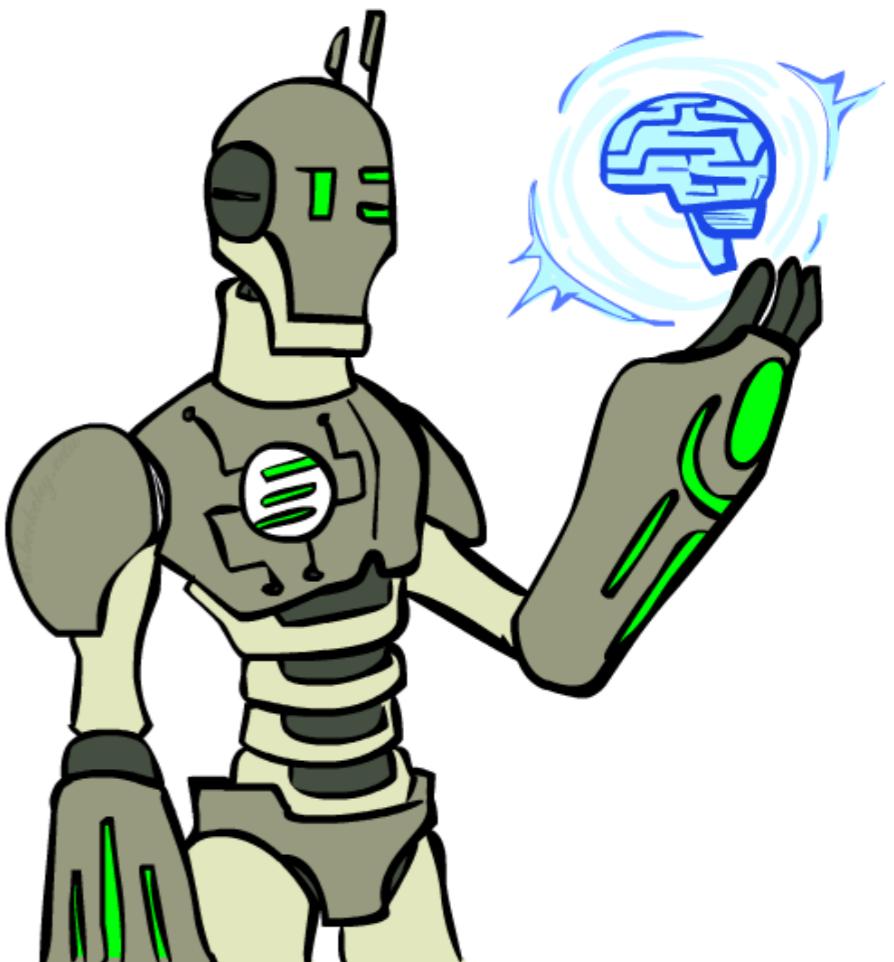
Paul Weng

UM-SJTU Joint Institute

Slides adapted from <http://ai.berkeley.edu>

# Today

- ❖ Course organization
- ❖ What is artificial intelligence?
- ❖ Past: How did the ideas in AI come about?
- ❖ Present: What is the state of the art?
- ❖ Future: Will robots take over the world?
- ❖ What is this course about?



# Course Staff

Instructor



Paul WENG

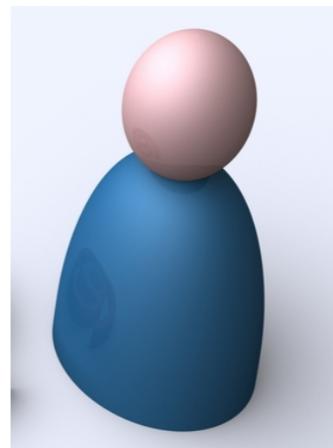
TAs



Jian YUAN



Shunyi ZHU



Third TA

---

# Course Organization

---

- ❖ **Classes**
- ❖ **Recitations:**
  - ❖ Not weekly, notably before exams
- ❖ **Office hours:**
  - ❖ To be determined the first week

---

# What is Flipped Classroom?

---

- ❖ **Poll:** Have you heard about flipped classroom before?
- ❖ **Idea:**
  - ❖ Outside class
    - ❖ Students read / watch teaching material
  - ❖ In class
    - ❖ Discussions about difficult concepts
    - ❖ Quizzes to check understanding
    - ❖ Exercises and problems solving
    - ❖ Demos

---

# Why Flipped Classroom?

---

- ❖ **Advantages:**
  - ❖ Students can study at their own pace
  - ❖ Content can be covered more in-depth (demo writing)
  - ❖ Class more interactive, more engaging, hopefully more fun?
  - ❖ Research shows that students learn / retain concepts better
- ❖ **Drawback:**
  - ❖ More efforts for both students and instructors
  - ❖ **Poll:** Do you want Ve280/Ve492 to be taught like this?

# Course Information

---

- ❖ **Prerequisites:**
  - ❖ Ve281 (Data Structures and Algorithms)
  - ❖ Recommended: Ve203 (Discrete Mathematics),  
Ve477 (Introduction to Algorithms)
  - ❖ Expect some math and a lot of programming
- ❖ **Work and Grading:**
  - ❖ In-class graded quizzes (15%)
  - ❖ ~10 homework assignments (20%)
    - ❖ Part 1: electronic, solve together, submit alone through online judge
    - ❖ Part 2: written, solve together, write up alone, electronic submission through Canvas
  - ❖ 5 graded programming projects: Python, groups of 1 or 2 (30%)
    - ❖ 5 late days for semester, maximum 2 per project
  - ❖ One midterm (15%), one final (20%)

---

# Honor Code

---

- ❖ Abide by the Honor Code
- ❖ Discussion (oral or via Piazza) is permitted and encouraged
- ❖ Never show your solutions to other students!
- ❖ After passing this course:
  - ❖ Please don't share your solutions
  - ❖ Please don't post the course materials online

# Online Tools

## ❖ Communication:

- ❖ Canvas
  - ❖ Announcements
  - ❖ Lecture slides
  - ❖ Project release
  - ❖ Grade release
- ❖ Piazza
  - ❖ Questions
  - ❖ Discussions

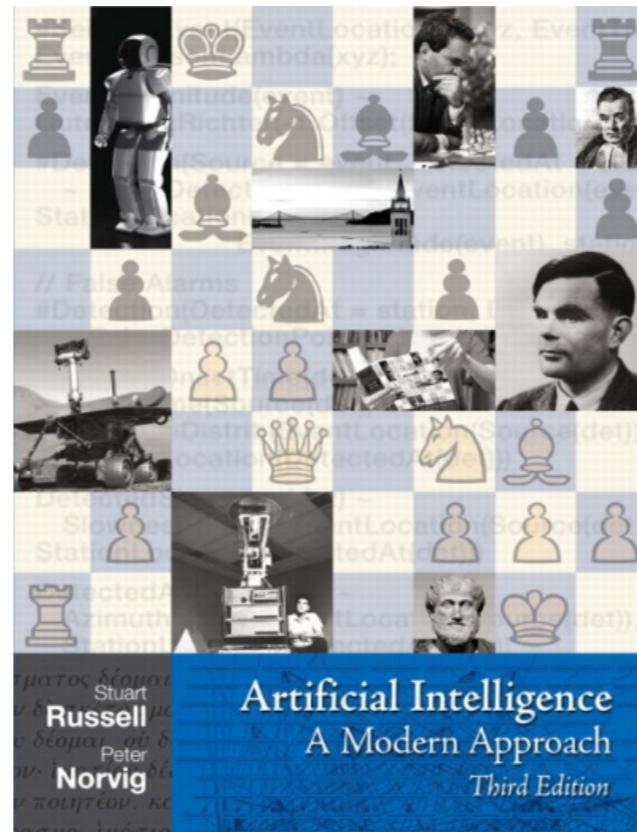
## ❖ Graded Assignments:

- ❖ Online judge
  - ❖ Part 1 of homework
  - ❖ Projects
- ❖ Canvas
  - ❖ Graded in-class quizzes
  - ❖ Part 2 of homework

The screenshot illustrates the integration of Canvas and Piazza. The Canvas sidebar is visible on the left, and a large red arrow points from its bottom towards the top of the Piazza interface. The Piazza interface shows a course navigation bar for 'VE492 > VE492, Intro to Artificial Intelligence'. The main area displays a feed of posts under 'Pinned' and 'Today', with a 'Welcome to Piazza!' message at the bottom. To the right of the Piazza feed, there is a 'Student Enrollment' section with an input field containing 'john@email.com, smith@email.com' and a 'Enroll Students' button. Below this, there is a section titled 'Allow TAs/other instructors' with 'yes' and 'no' buttons.

# Textbook

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



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

---

# Important Announcements

---

- ❖ This week:
  - ❖ Register for the class on Piazza --- our main resource for discussion and communication
  - ❖ P0: Python tutorial is out (due on May 18 at 11:59pm)
  - ❖ Poll to choose dates of Recitations and Office Hours
- ❖ Next week:
  - ❖ Office Hours start next week.

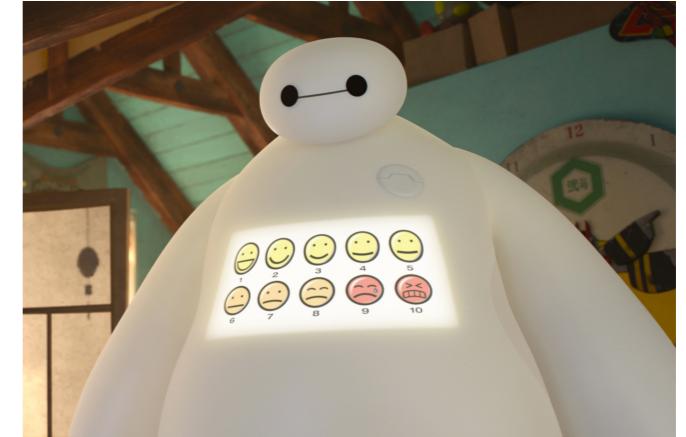
---

# Advice

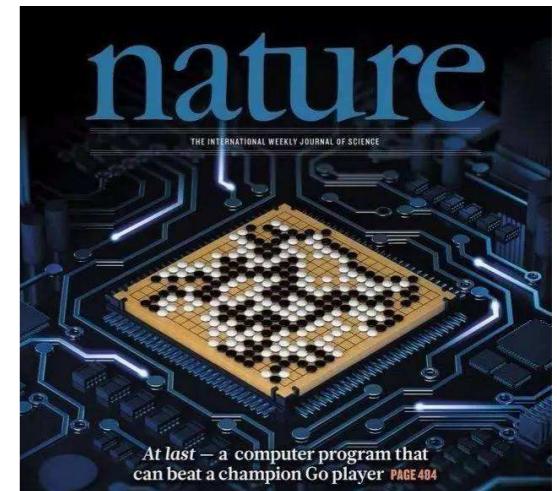
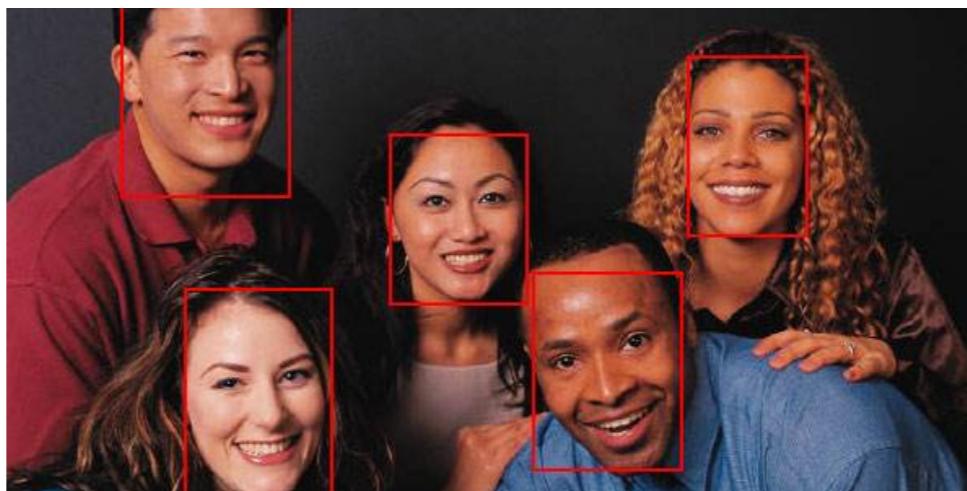
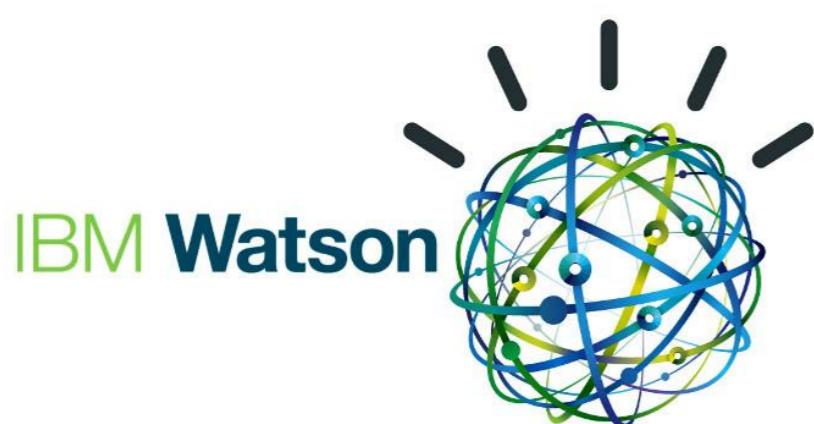
---

- ❖ Take writing notes
- ❖ Review the material regularly
- ❖ Follow the pace of the course
- ❖ Organize your time
- ❖ Don't hesitate to ask for help: Piazza, OH, email, schedule meeting
- ❖ Don't hesitate to provide us some feedback

# AI in Movies



# AI in the News



# AI in the News

Economics

## China's Got a Huge Artificial Intelligence Plan

Bloomberg News

July 2

Technology Intelligence

**AI is the biggest risk we face as a civilisation, Elon Musk says**



Musk was speaking to US governors CREDIT: AP

HOME > FINANCE > FINANCE TOPICS > DAVOS

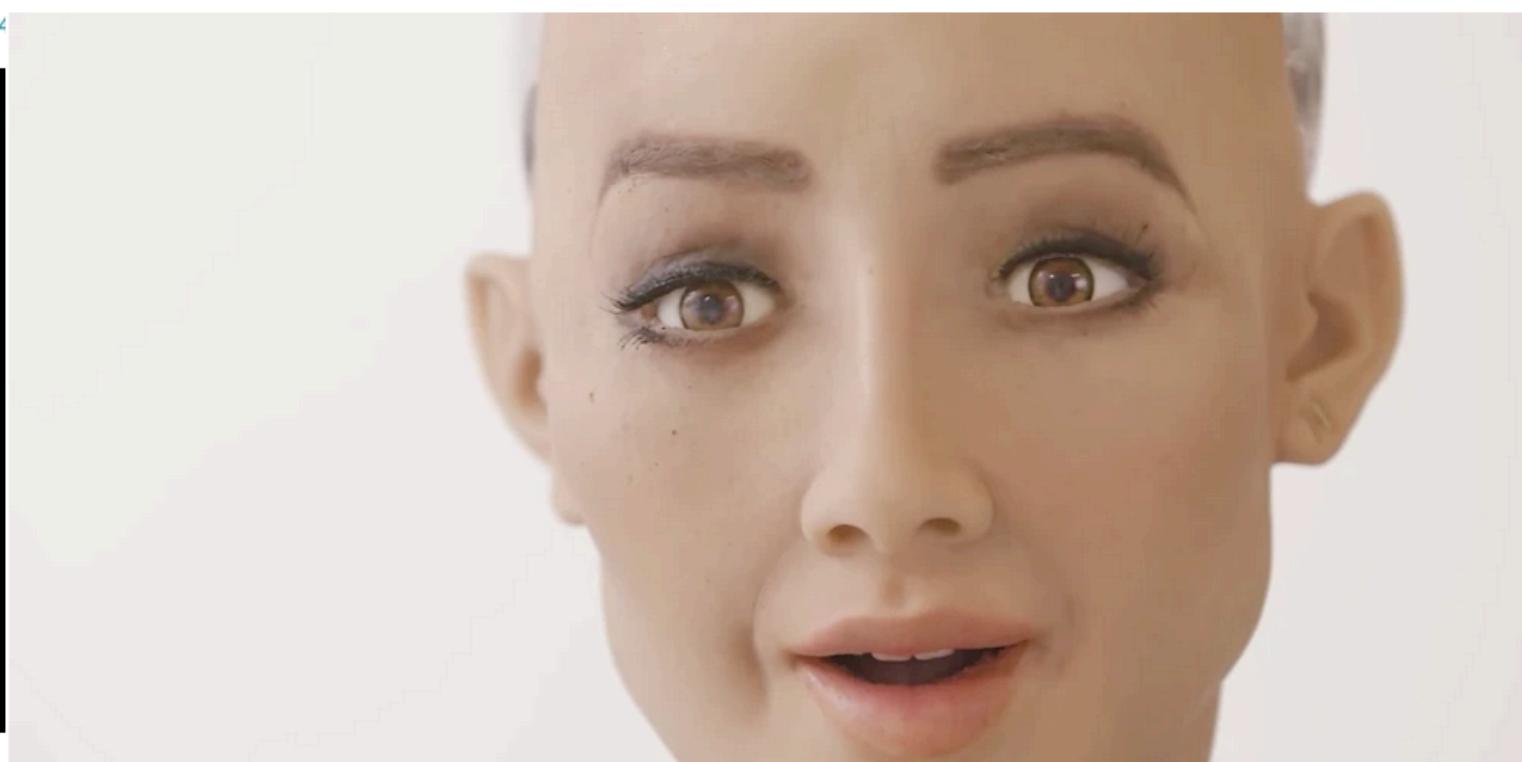
'Sociopathic' robots could overrun the human race within a generation

Computers should be trained to serve humans to reduce their threat to the human race, says a leading expert on artificial intelligence

**Saudi Arabia bestows citizenship on a robot named Sophia**

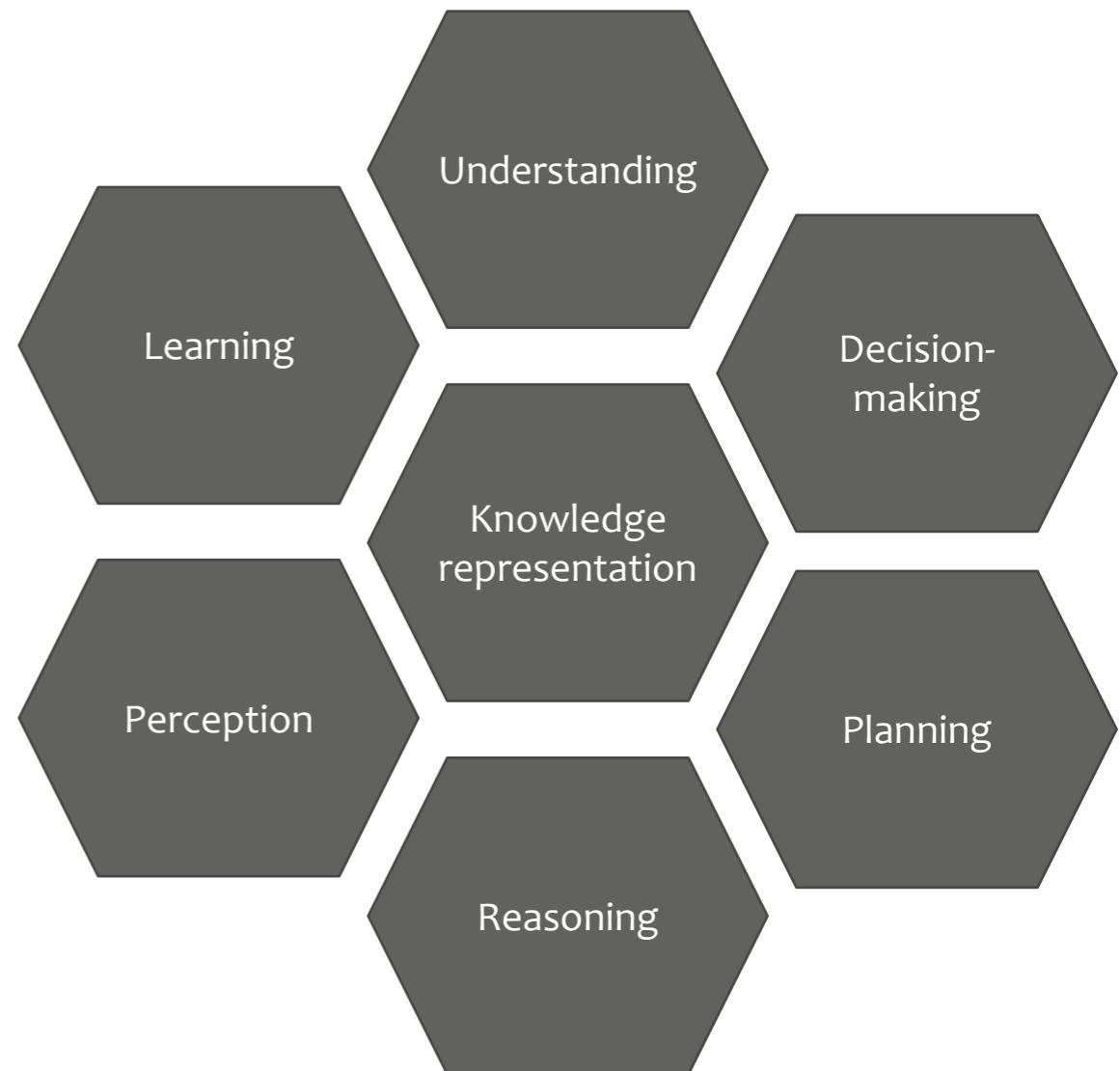
Taylor Hatmaker @tayhatmaker 7:59 am CST • October 27, 2017

Comment

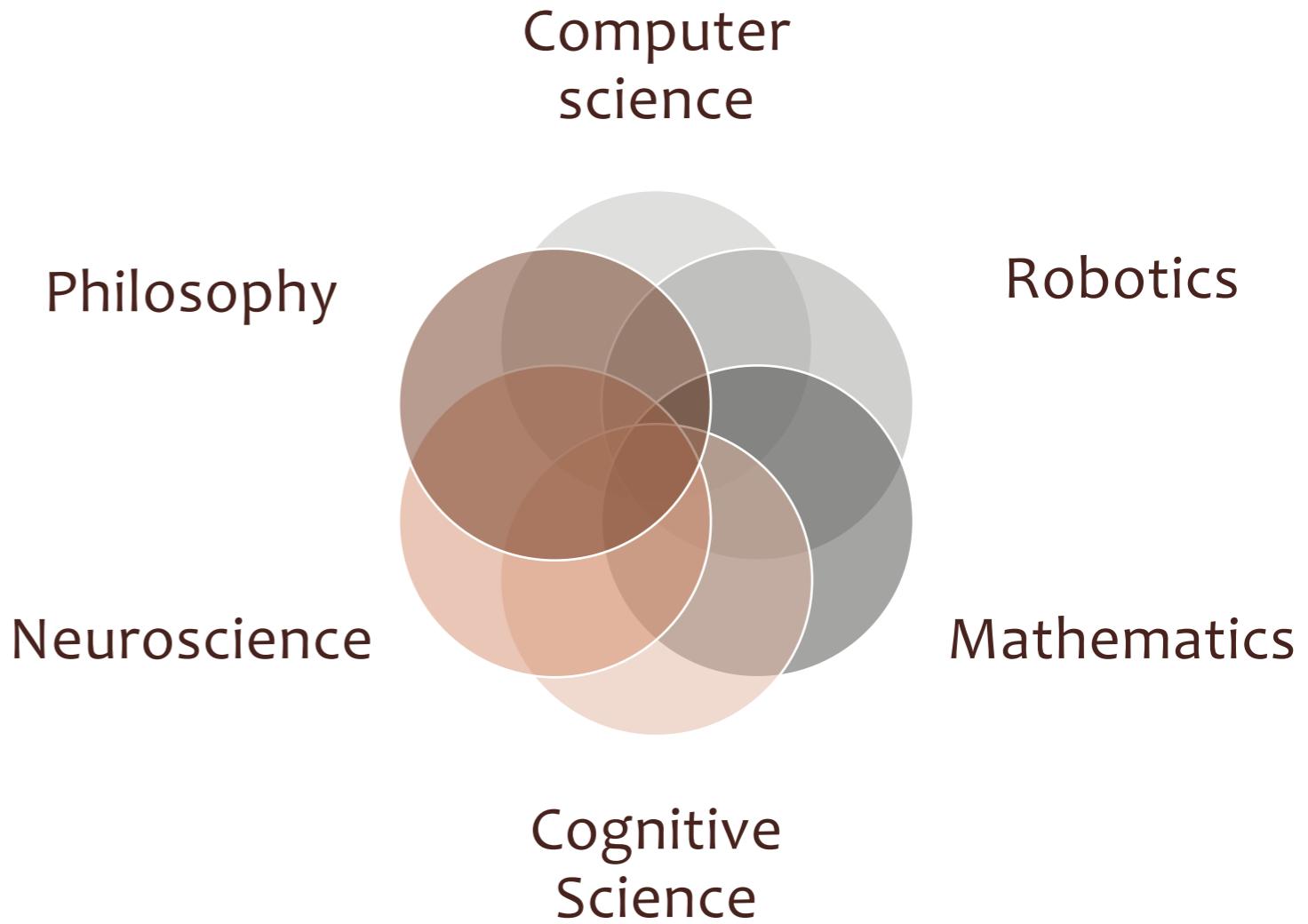
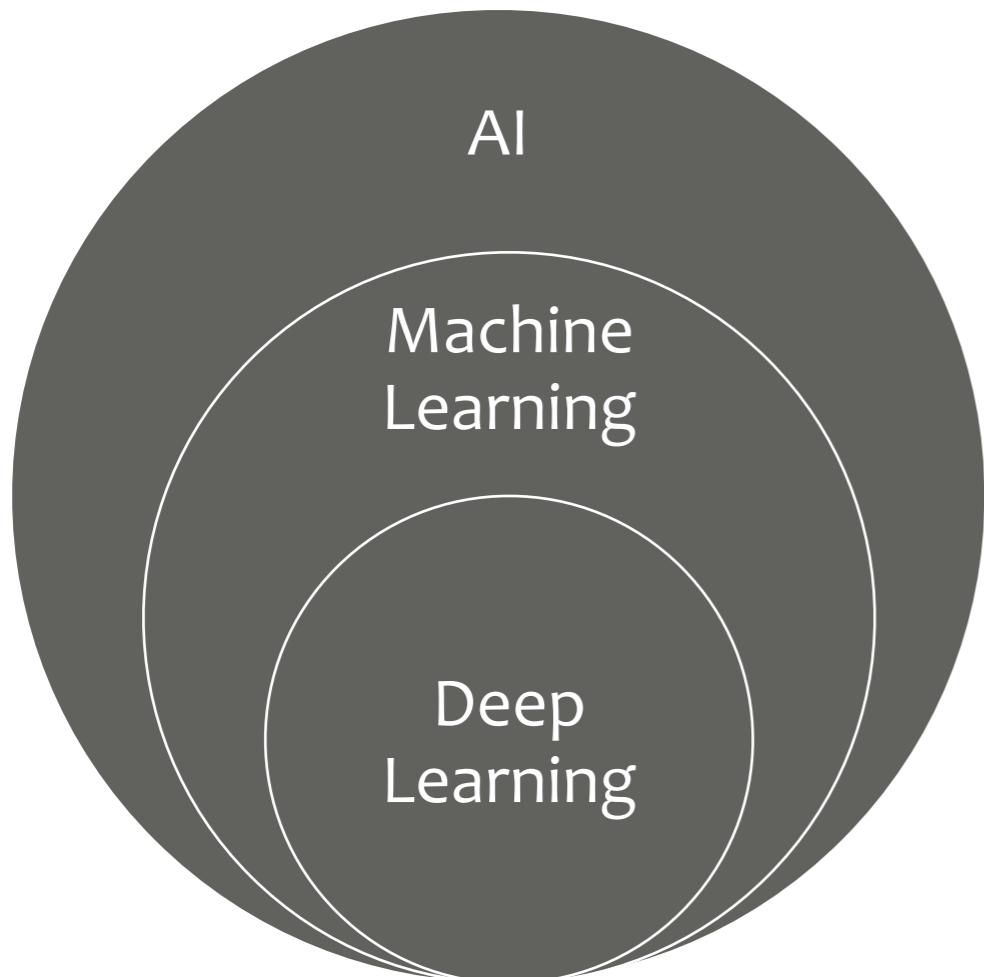


# AI as an Ability

- ❖ AI=cognitive capabilities displayed by a machine
- ❖ AI=machine that possesses such capabilities



# AI as a Research Area



---

# AI as a Field

---

The science of making machines that:

# AI as Computational Rationality

---

- ❖ Humans are intelligent to the extent that our actions can be expected to achieve our objectives
- ❖ Machines are intelligent to the extent that their actions can be expected to achieve their objectives
  - ❖ Control theory: minimize cost function
  - ❖ Economics maximize expected utility
  - ❖ Operations research: minimize sum of costs
  - ❖ Statistics: minimize loss function
  - ❖ AI: all the above, plus logically-defined goals
- ❖ AI agent ≈ computational rational agent

---

# Rational Decisions

---

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

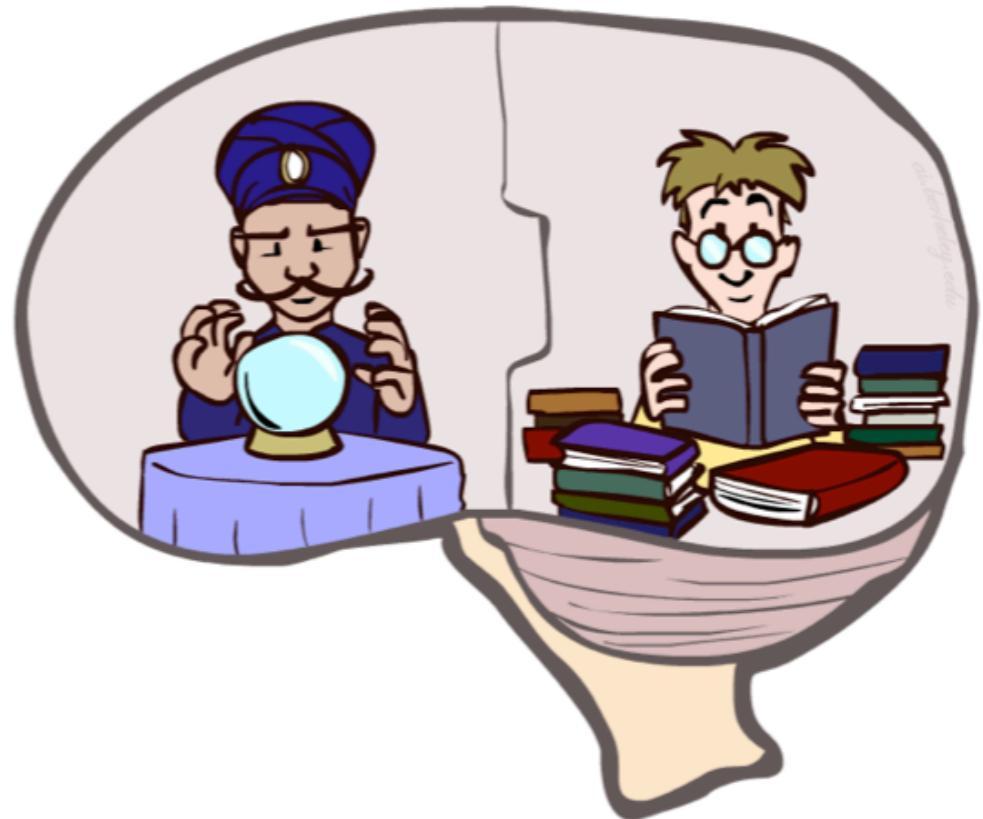
- Rational: maximally achieving pre-defined goals
- Rationality only concerns what decisions are made  
(not the thought process behind them)
- Goals are expressed in terms of the **utility** of outcomes
- Being rational means **maximizing your expected utility**

A better title for this course would be:

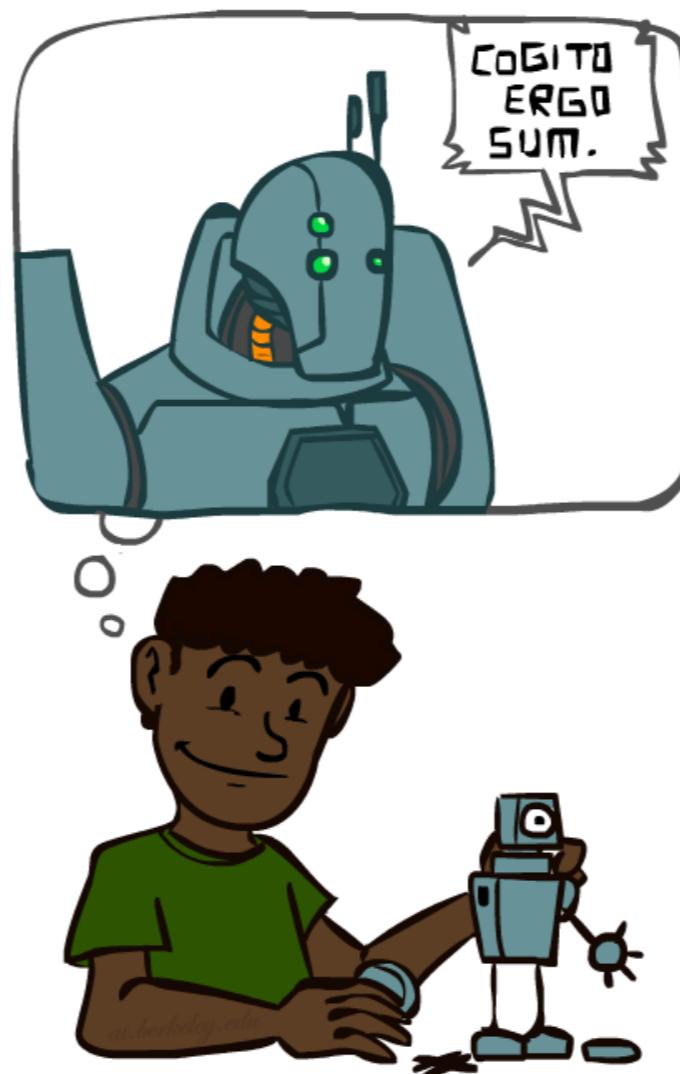
**Computational Rationality**

# What About the Brain?

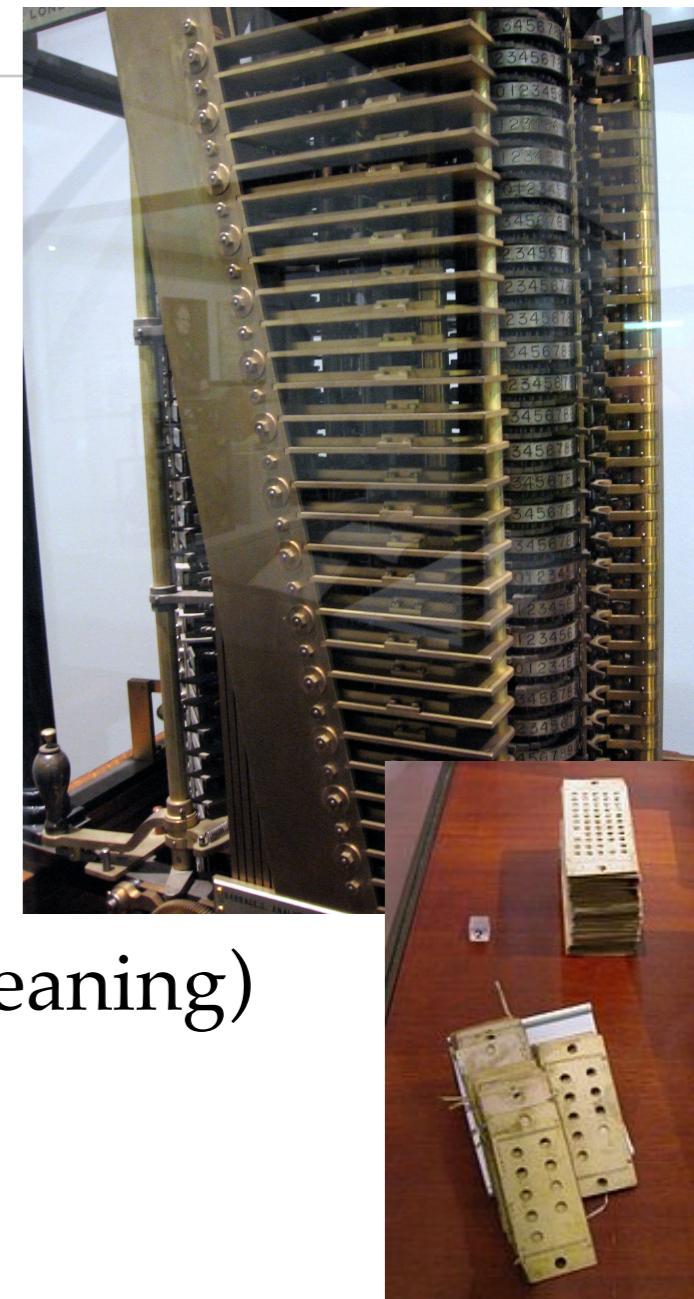
- ❖ Brains (human minds) are very good at making rational decisions, but not perfect
- ❖ Brains aren't as modular as software, so hard to reverse engineer!
- ❖ “Brains are to intelligence as wings are to flight”
- ❖ Lessons learned from the brain: memory and simulation are key to decision making



# A (Short) History of AI



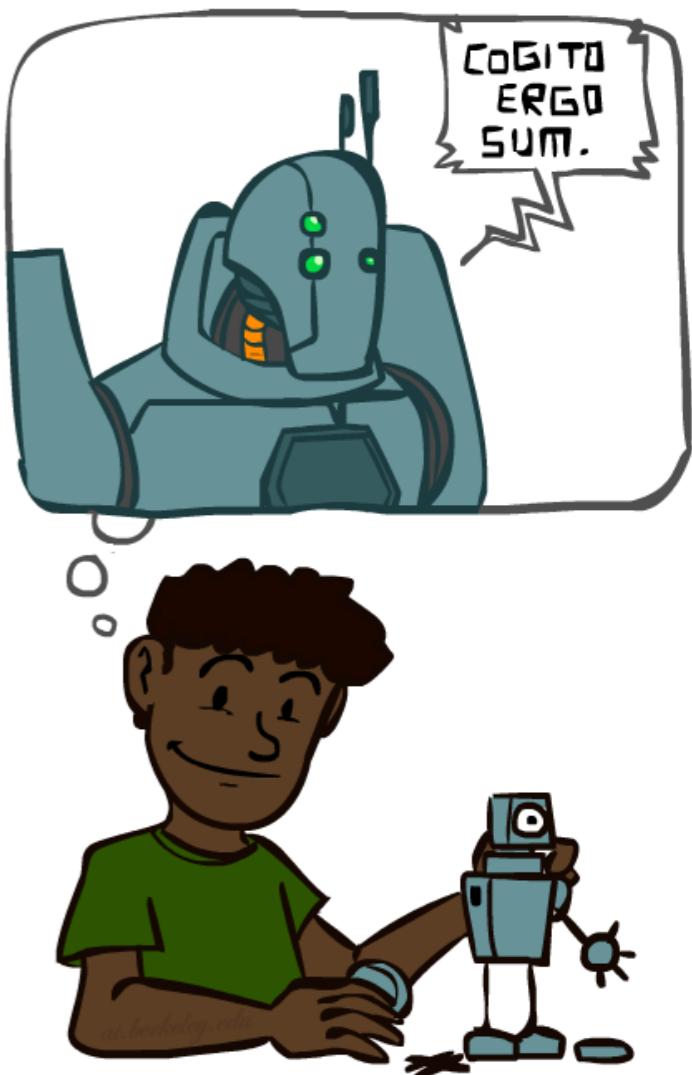
# A (Short) Prehistory of AI



- ❖ **Prehistory**
  - ❖ Philosophy from Aristotle onwards
  - ❖ Mathematics (logic, probability, optimization)
  - ❖ Neuroscience (neurons, adaptation)
  - ❖ Economics (rationality, game theory)
  - ❖ Control theory (feedback)
  - ❖ Psychology (learning, cognitive models)
  - ❖ Linguistics (grammars, formal representation of meaning)
- ❖ **Ancestor of computers (1834-1906)**
  - ❖ Babbage design for universal machine
  - ❖ Lovelace: “a thinking machine” for “all subjects in the universe”

# A (Short) History of AI

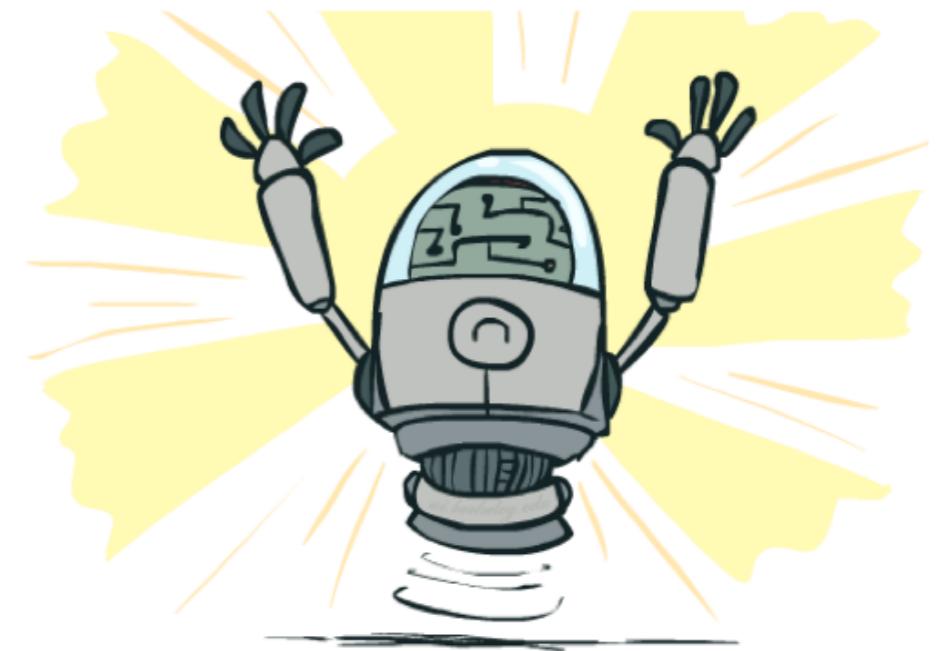
- ❖ 1940-1950: Early days
  - ❖ 1943: McCulloch & Pitts: Boolean circuit model of brain
  - ❖ 1950: Turing's "Computing Machinery and Intelligence"
- ❖ 1950-70: Excitement: Look, Ma, no hands!
  - ❖ 1950s: Early AI programs: chess, checkers programs, theorem proving
  - ❖ 1956: Dartmouth meeting: "Artificial 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
  - ❖ 1988—93: Expert systems industry busts: "AI Winter"
- ❖ 1990-2012: Statistical approaches + subfield expertise
  - ❖ Resurgence of probability, focus on uncertainty
  - ❖ General increase in technical depth
  - ❖ Agents and learning systems... "AI Spring"?
- ❖ 2012-Today: Excitement: Look, Ma, no hands!
  - ❖ Big data, big compute, neural networks
  - ❖ Some re-unification of sub-fields
  - ❖ AI used in many industries



# What Can AI Do?

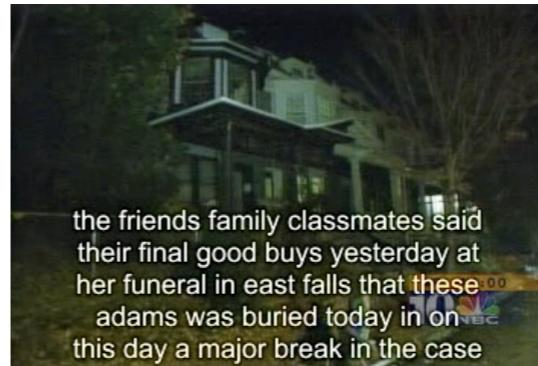
Which of the following can be done at present?

- ✓ Play a decent game of table tennis?
- ✓ Play a decent game of Jeopardy?
- ✓ Drive safely along a curving mountain road?
- ✗ Drive safely along Dongchuan road?
- ✓ Buy a week's worth of groceries on the web?
- ✗ Buy a week's worth of groceries at Auchan?
- ✗ 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?
- ✗ Compose original music?



# 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



## Trump confirme l'imposition dimanche de nouveaux tarifs douaniers sur des biens chinois

Les marchandises visées touchent un très large éventail de produits notamment les secteurs alimentaire, textile et ameublement.

Publié le 31 août 2019 à 00h34 - Mis à jour le 31 août 2019 à 12h28



## Trump confirms Sunday imposition of new tariffs on Chinese goods

The goods concerned cover a very wide range of products, particularly the food, textile and furniture sectors.

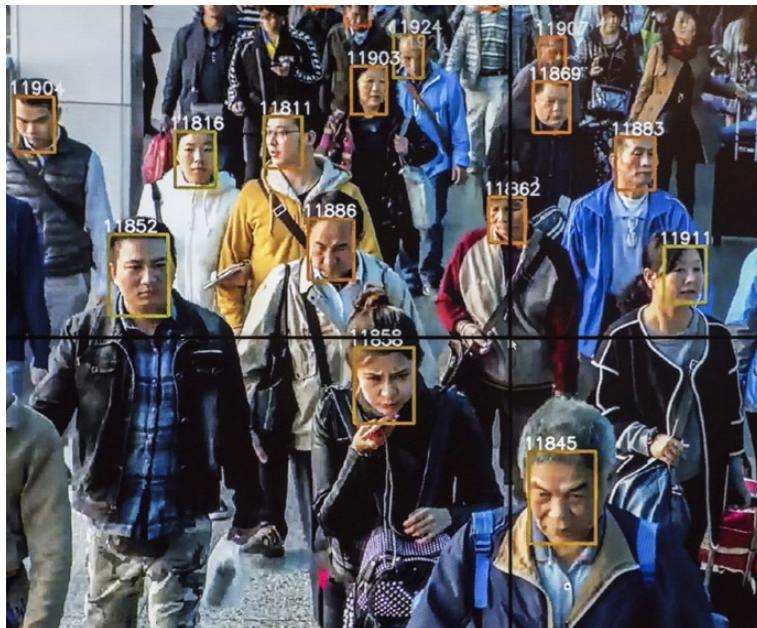
Posted on August 31, 2019 at 00:34 - Updated on August 31, 2019 at 12:28 pm



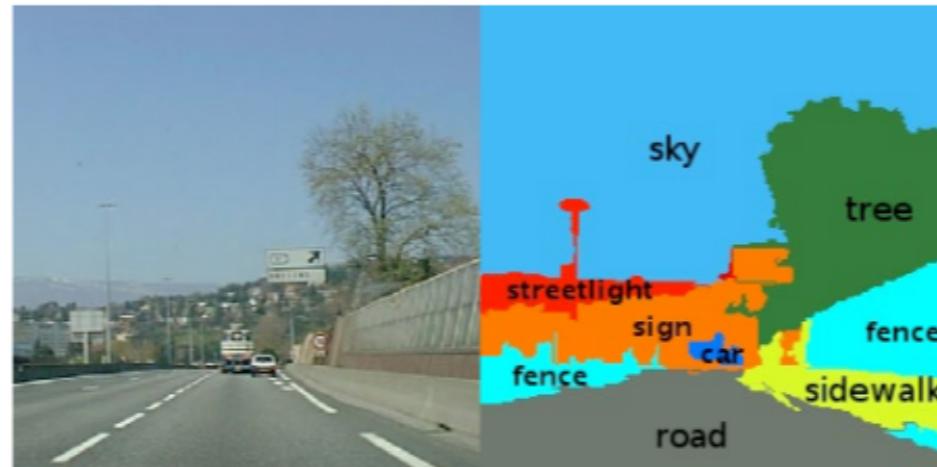
- ❖ Web search
- ❖ Text classification, spam filtering, etc...

# Vision (Perception)

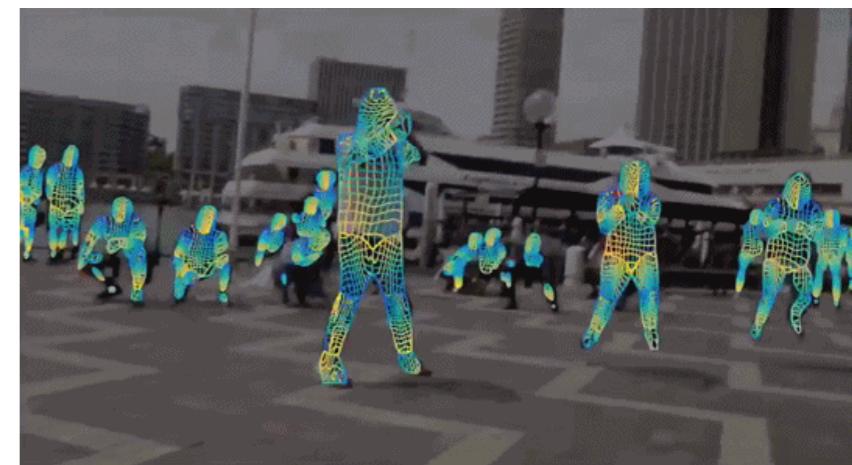
## Object and face recognition



Scene segmentation

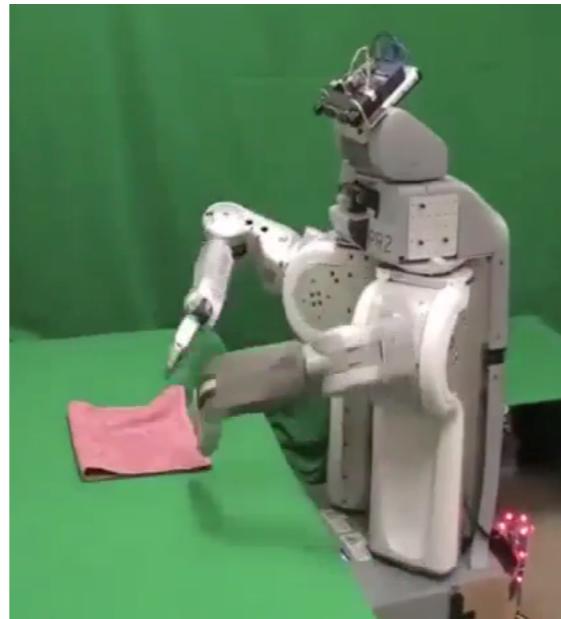


Pose estimation

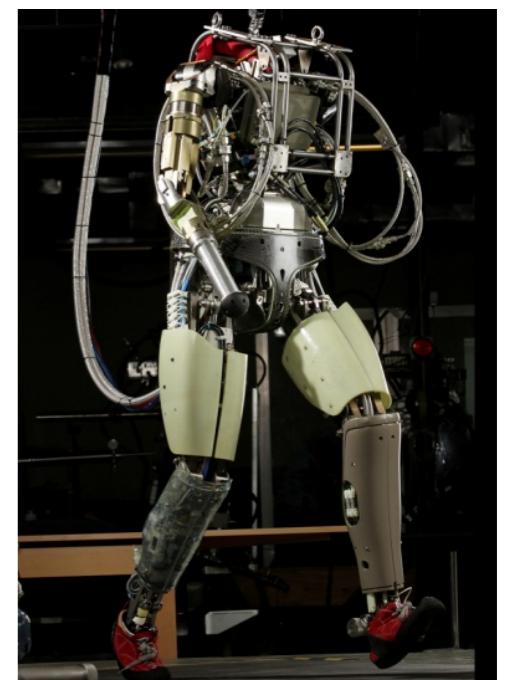


# Robotics

- ❖ **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

# Logic

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

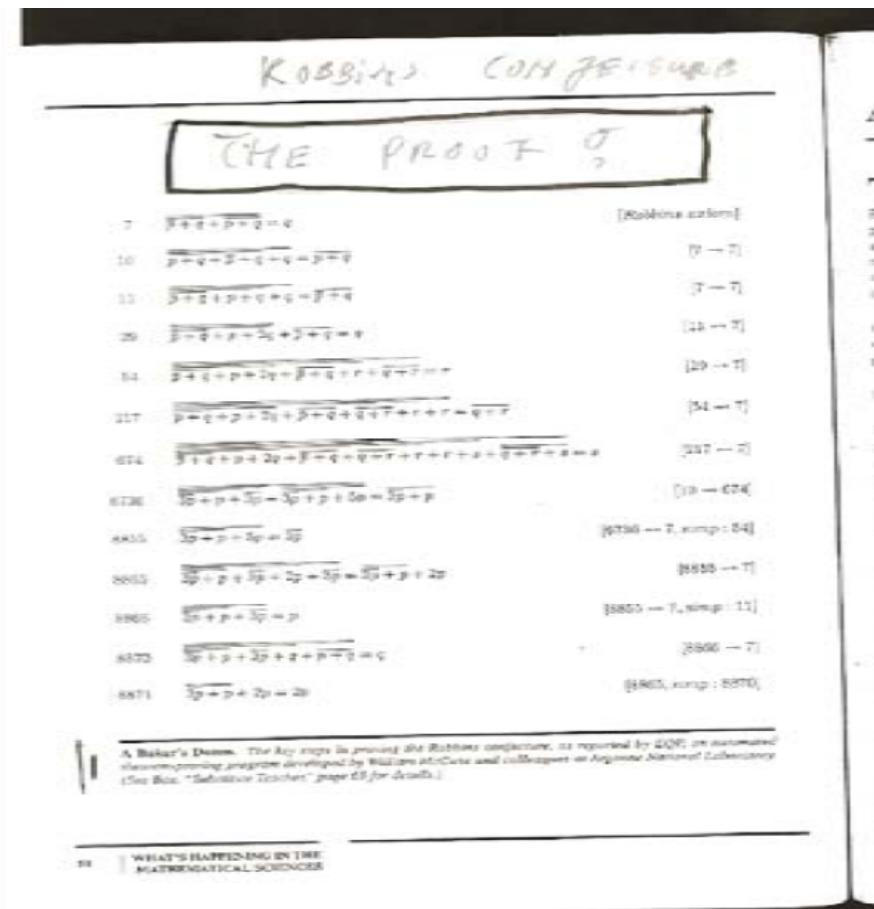


Image from Bart Selman

# 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

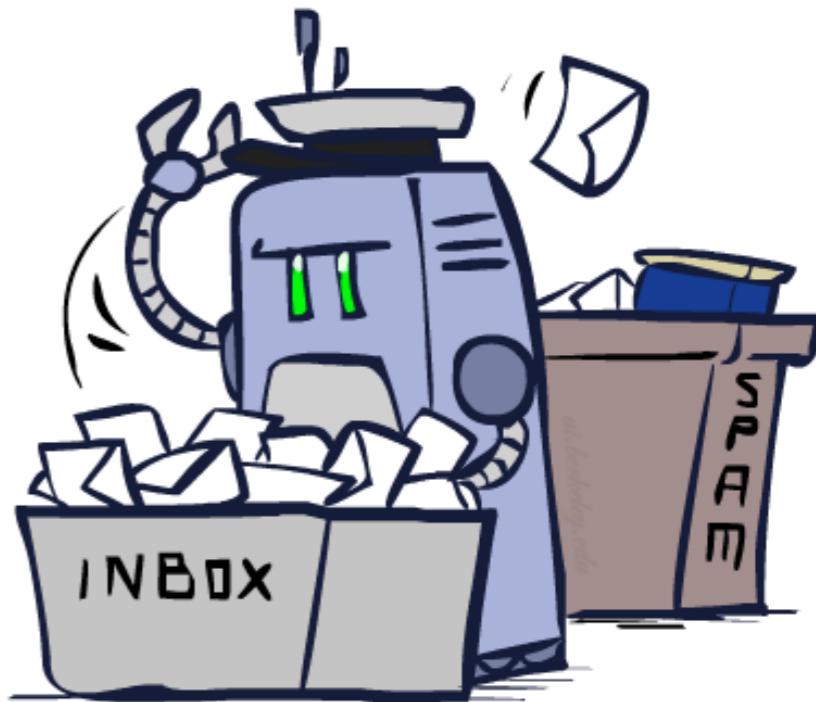
“Deep Blue hasn't proven anything.”
- ❖ 2017: Alpha Go Beats Ke Jie

“It plays in a third, almost alien, way.”
- ❖ 2018: Fine Art Beats Ke Jie with two-stone handicap



# 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!



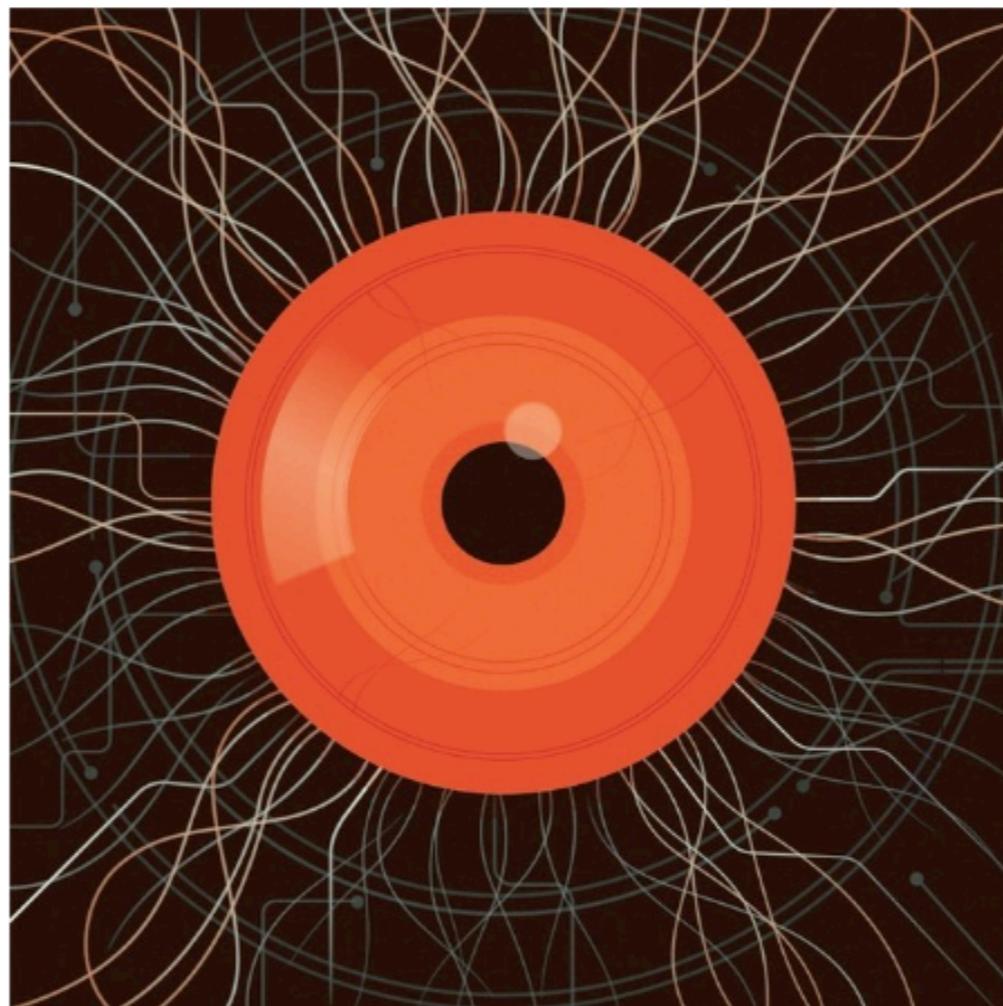
# Future

## HOW FRIGHTENED SHOULD WE BE OF A.I.?

*Thinking about artificial intelligence can help clarify what makes us human—for better and for worse.*



By **Tad Friend** May 7, 2018



*An A.I. system may need to take charge in order to achieve the goals we gave it.*

Illustration by Harry Campbell

# What Could Go Wrong?

AI risks replicating tech's ethnic minority bias across business

Diverse workforce essential to combat new danger of 'bias in, bias

BRIEFING • ARTIFICIAL INTELLIGENCE

## A.I. Expert Says Automation Could Replace Millions of Jobs in 15 Years

By Don Reisinger January 10, 2019

An artificial intelligence expert and venture capitalist predicts automation will cause major changes in the workforce.



Speaking to CBS News' Scott Pelley in an interview for *60 Minutes* on Sunday, Lee said that he believes 40% of the world's jobs will be replaced by automation over the next 15 years. He said that both blue collar and white collar professionals will be affected, but he believes those who drive for a living could be most at risk.

"Chauffeurs, truck drivers, anyone who does driving for a living—that's probably going to be disrupted more in the 15-25 year time frame," he said in the interview. "There are things that seem a little bit complex, chef, waiter, a lot of things will become automated."

**technology** ➤ **home entertainment > computers**

Aliya Ram MAY 31, 2018

"We used to talk about garbage in, garbage out," Microsoft CEO Satya Nadella said during a review into artificial intelligence commissioned by the company. "Now, when we talk about AI, we talk about bias in, bias out."

## Microsoft CEO: tech sector needs to pro

MICROSOFT boss Satya Nadella has issued an ominous warning about the future of AI, suggesting it could lead to a dystopian, 1984-esque future.

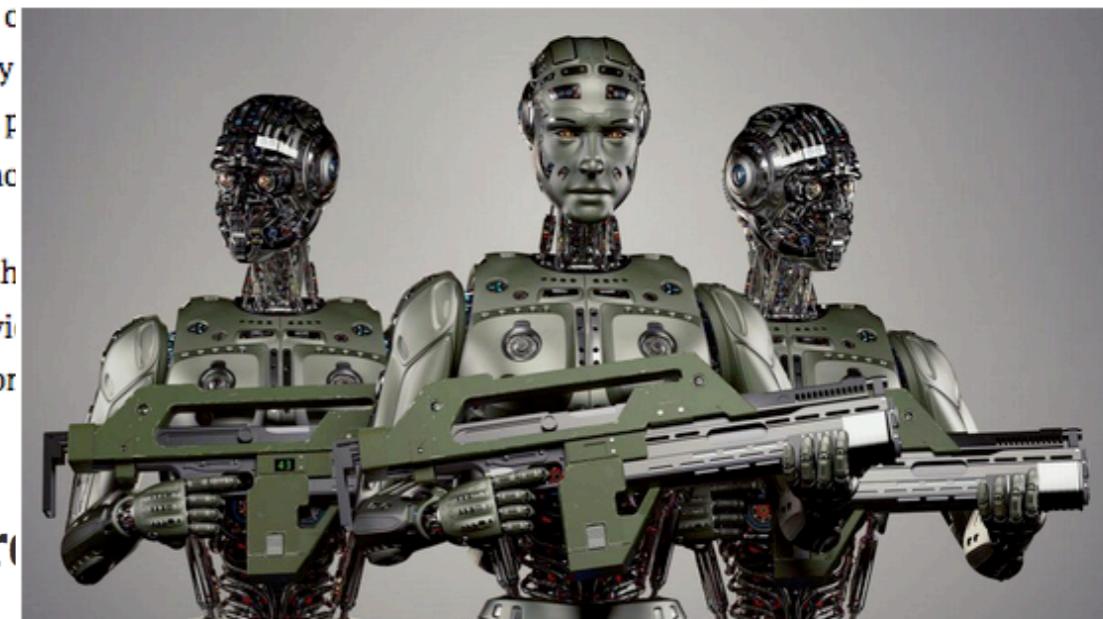
## The Weaponization Of Artificial Intelligence

Jayshree Pandya Contributor

COGNITIVE WORLD Contributor Group

AI & Big Data

Jayshree Pandya is Founder of Risk Group & Host of Risk Roundup.



The reality of the rise of autonomous weapons systems SHUTTERSTOCK ENHANCED BY COGWORLD

---

# Some Issues Raised by AI

---

- ❖ Ethical questions
  - ❖ If an AI system causes some harm, who's responsible?
- ❖ Societal impact
  - ❖ Job automation
  - ❖ Privacy
- ❖ Philosophical questions
  - ❖ What is being human?
- ❖ Existential consequences
  - ❖ Is AI a threat to humanity?

---

# Value Alignment Problem

---

- ❖ AI may exploit any defects in how we specify its goal
  - ❖ Three rules of robotics (Asimov, 1942)
    - ❖ First Law: A robot may not injure a human being or, through inaction, allow a human being to come to harm.
    - ❖ Second Law: A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
    - ❖ Third Law: A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.
  - ❖ Rationality based on Expected Utility

---

# Instrumental Goals

---

- ❖ For any **primary** goal, the odds of success are improved by
  - ❖ Maintaining one's own existence
  - ❖ Acquiring more resources
- ❖ With value misalignment, these lead to obvious problems for humanity

# AI as Computational Rationality

---

- ❖ Humans are intelligent to the extent that our actions can be expected to achieve our objectives
- ❖ Machines are intelligent to the extent that their actions can be expected to achieve their objectives
  - ❖ Control theory: minimize cost function
  - ❖ Economics maximize expected utility
  - ❖ Operations research: minimize sum of costs
  - ❖ Statistics: minimize loss function
  - ❖ AI: all the above, plus logically-defined goals
- ❖ We don't want machines that are intelligent in this sense!
- ❖ Machines are beneficial to the extent that their actions can be expected to achieve our objectives
- ❖ We need machines to be provably beneficial

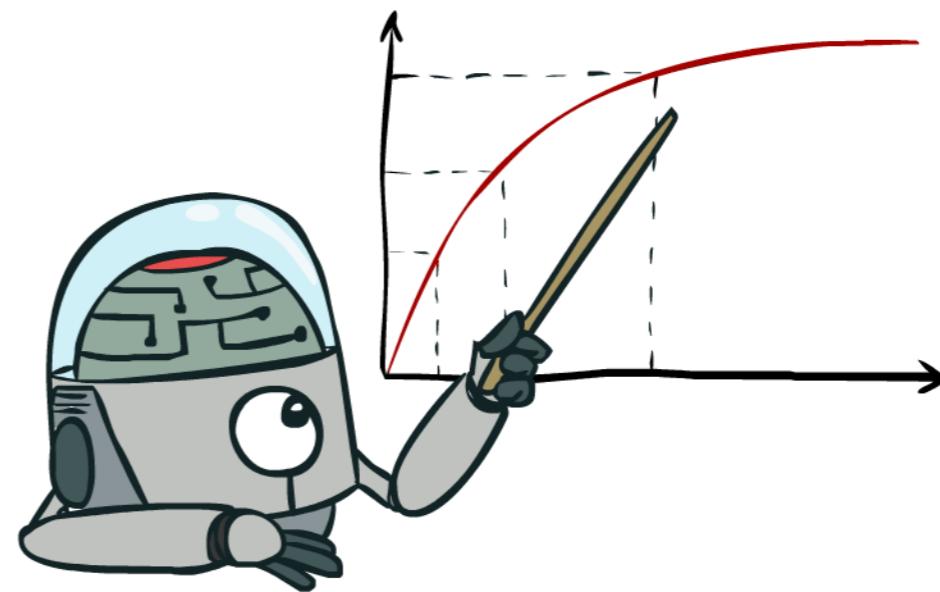
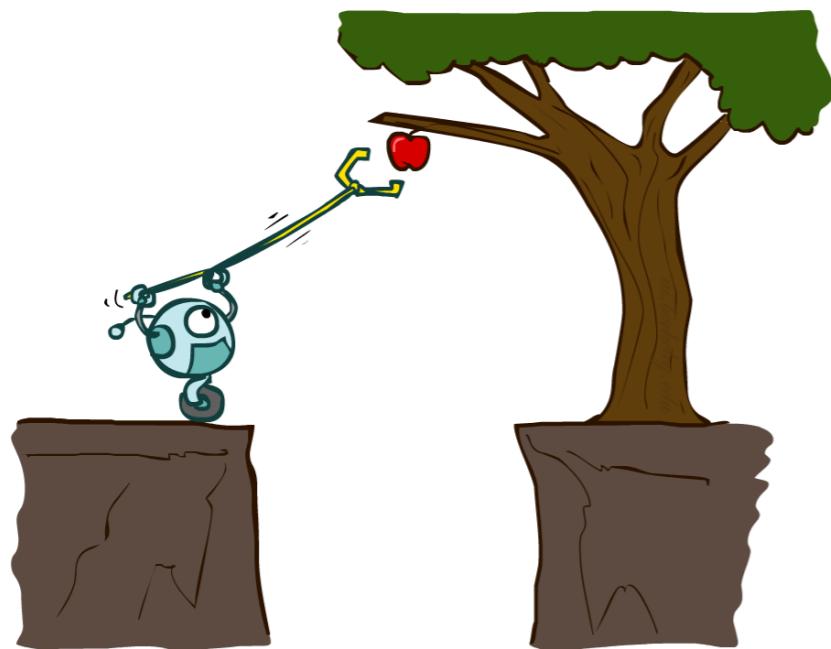
---

# Towards Human-Level AI?

---

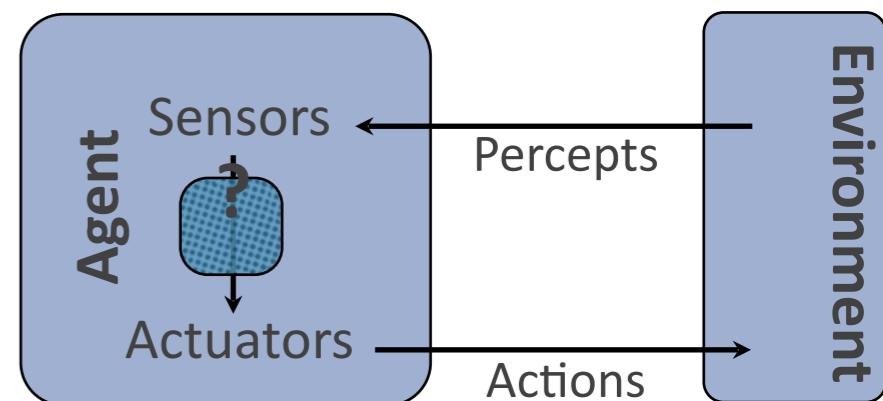
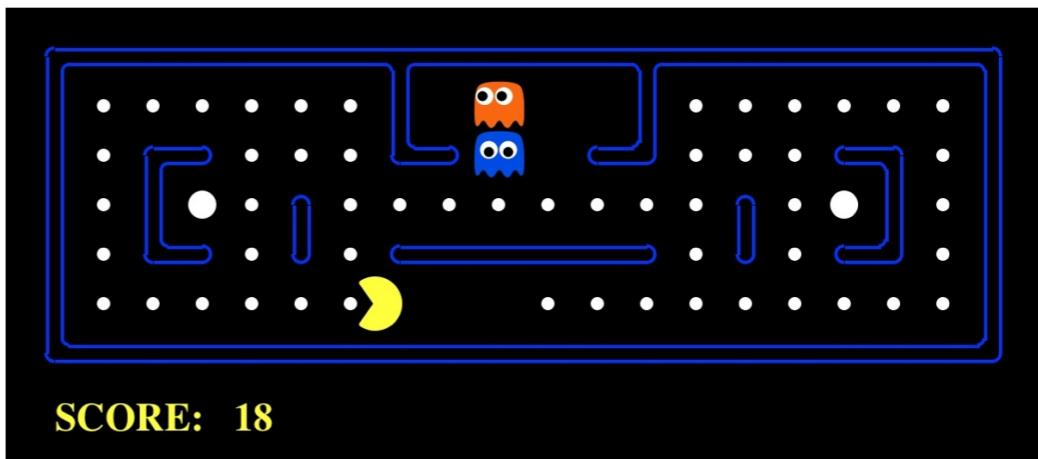
- ❖ Still missing:
  - ❖ Real understanding of language
  - ❖ Integration of learning with knowledge
  - ❖ Long-range thinking at multiple levels of abstraction
  - ❖ Cumulative discovery of concepts and theories
- ❖ Date: Unpredictable
  - ❖ Weak AI: system specialized in one task
  - ❖ Strong AI: general system

# Course Content



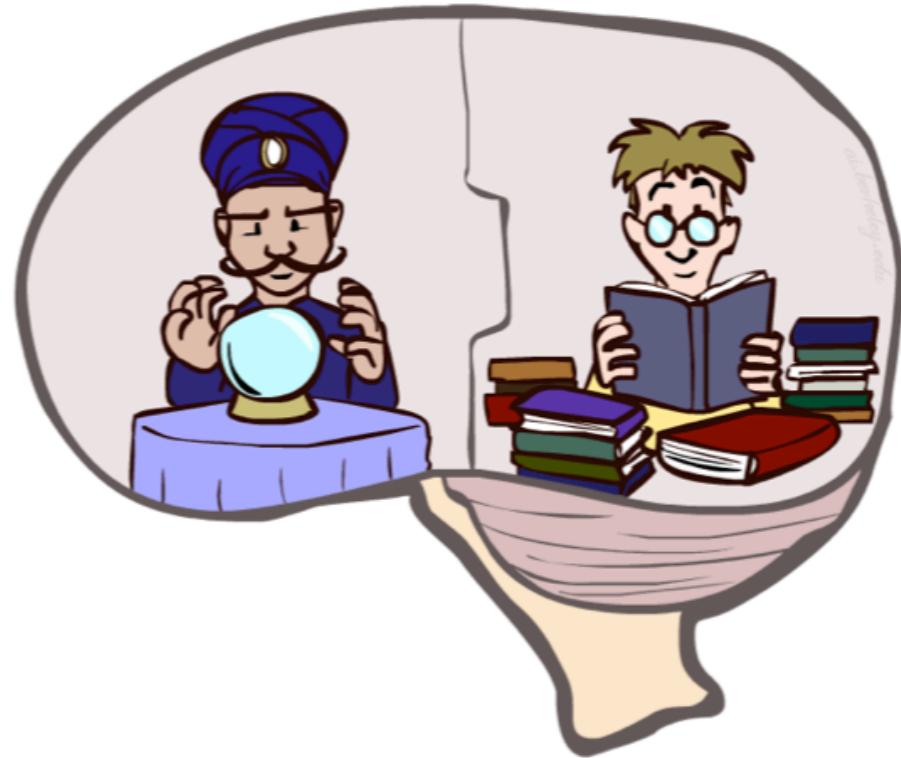
# Designing Rational Agents

- ❖ 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



# Course Topics

- ❖ Reasoning
  - ❖ Search; Constraint satisfaction
- ❖ Decision-making
  - ❖ Markov Decision Processes
- ❖ Knowledge representation
  - ❖ Logic
- ❖ Reasoning under uncertainty
  - ❖ Bayesian networks
- ❖ Learning
  - ❖ Reinforcement learning; neural networks



---

# Relation to Other Courses

---

- ❖ JI:
  - ❖ Ve485 Optimization for Machine Learning
  - ❖ Ve445 Intro to Machine Learning
  - ❖ Ve414 Bayesian Analysis
  - ❖ Ve572 Methods and Tools for Big Data
  
- ❖ Outside SJTU:
  - ❖ UM: EECS 492 Intro to AI
  - ❖ Berkeley: AIMA, CS188 Intro to AI
  - ❖ CMU: AI: Representation and Problem Solving