
Artificial Intelligence

Chris Callison-Burch

CIS 421/521

Fall 2018

<http://artificial-intelligence-class.org>



Welcome to CIS 521

- **Professor:**
 - Chris Callison-Burch, ccb@cis.upenn.edu
 - 3401 Walnut St., 4th floor B/C wing, room 463C
 - Key card access is required – make an appointment beforehand

- **TAs:**



Andrej Ilić



Deniz Beser



Ignacio Arranz



Jie Gao



Maria Kustikova



Simeng Sun



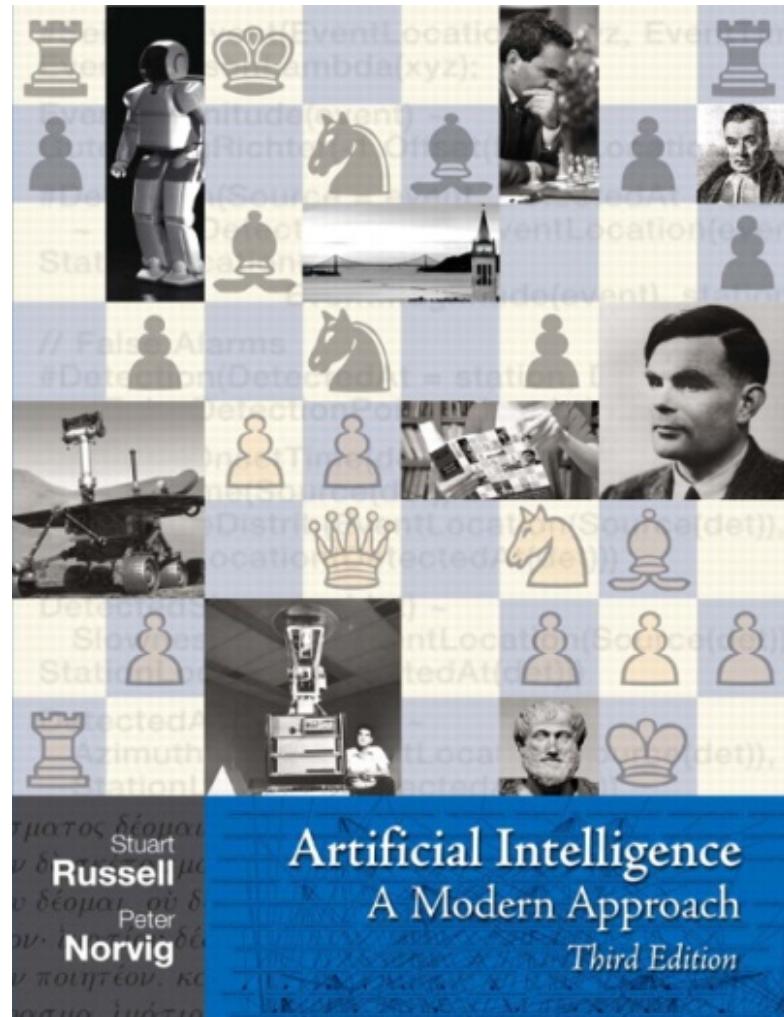
Welcome to CIS 521

- Course web page: <http://artificial-intelligence-class.org>
 - Lecture slides on web page
 - Homeworks on web page
- Discussion via Piazza (link on course home page)
- Lectures will be recorded using the Panopto system starting soon (if not now)
- Prerequisites:
 - Good knowledge of programming, data structures
 - Enough experience to *master* Python after two class lectures.
 - Introductory probability and statistics, and linear algebra will be *very useful*



Course Textbook

S. Russell and P. Norvig
Artificial Intelligence: A Modern Approach
Prentice Hall, 2009,
Third Edition (U.S.)



Grading & Homework

Grading:

- **55% Homeworks**
- **15% Midterm 1**
- **15% Midterm 2**
- **15% Midterm 3**

Homework:

- **Homework will be due at 11:59 on specified dates (usually Tuesdays).**
- **You can use up to 5 free late days.** Homeworks can be submitted at most two days late. If you are out of late days, then you will not be able to submit your homework. One “day” is defined as anytime between 1 second and 24 hours after the homework deadline.
- **Your lowest homework grade will be dropped.**



Academic Integrity

- *You can discuss homework problems with others (you will need to explicitly list who you discussed problems with on each homework submission), but*

ALL CODE MUST BE YOUR OWN INDEPENDENT WORK

- *Academic dishonesty as defined in Penn's Code of Academic Integrity will not be tolerated. In case of any question, cases will be handed to the Office of Student Conduct.*
- *The software we use to spot plagiarism is very good – it sees through*
 - *Variable changes*
 - *Spacing changes*
 - *Scrambling of independent code segments*



Waitlist

- There are 250+ students on the waitlist
- Currently, permits have been granted only to students who need this class to graduate (Robo master's students, and undergrads in the Computer and Cognitive Science in their final year).
- We are above the max enrollment for this course is 100 students
- Unlikely that more positions will be available.



On to the Real Stuff:

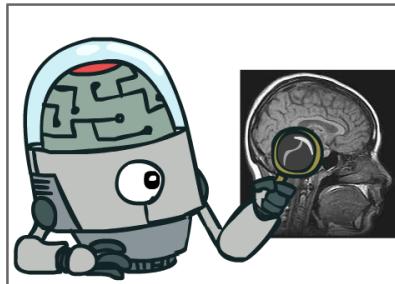
AI IN SCIENCE FICTION AND PHILOSOPHY



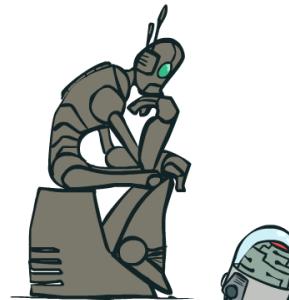
What is AI?

The science of making machines that:

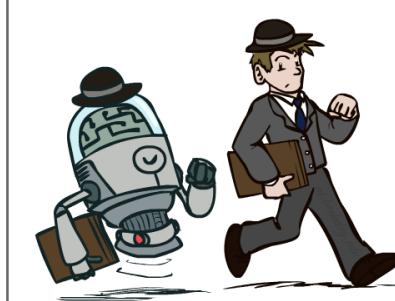
Think like people



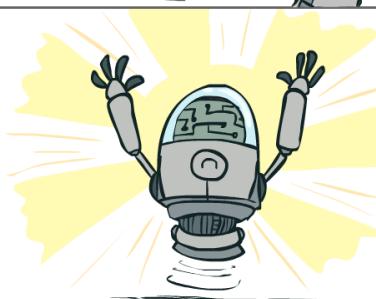
Think rationally



Act like people



Act rationally



AI in SciFi - Acting Human



STAR TREK
THE NEXT GENERATION

AI in SciFi - Acting Human



AI in SciFi - Acting Human



AI in SciFi - Acting Human



René Descartes (1596-1650)



cogito ergo sum

“I think, therefore I am.”

Principle of *dualism* – that the mind or thinking self is essentially incorporeal or spiritual – that the mind exists separately from the body: "if a foot or arm or any other part of the body is cut off, nothing has thereby taken away from the mind."



René Descartes (1596-1650)



Rene Descartes wondered if he could know for sure that others who looked and behaved like him weren't in fact automata.

Bodies of people and animals are nothing more than complex machines - the bones, muscles and organs could be replaced with cogs, pistons and cams.



René Descartes (1596-1650)

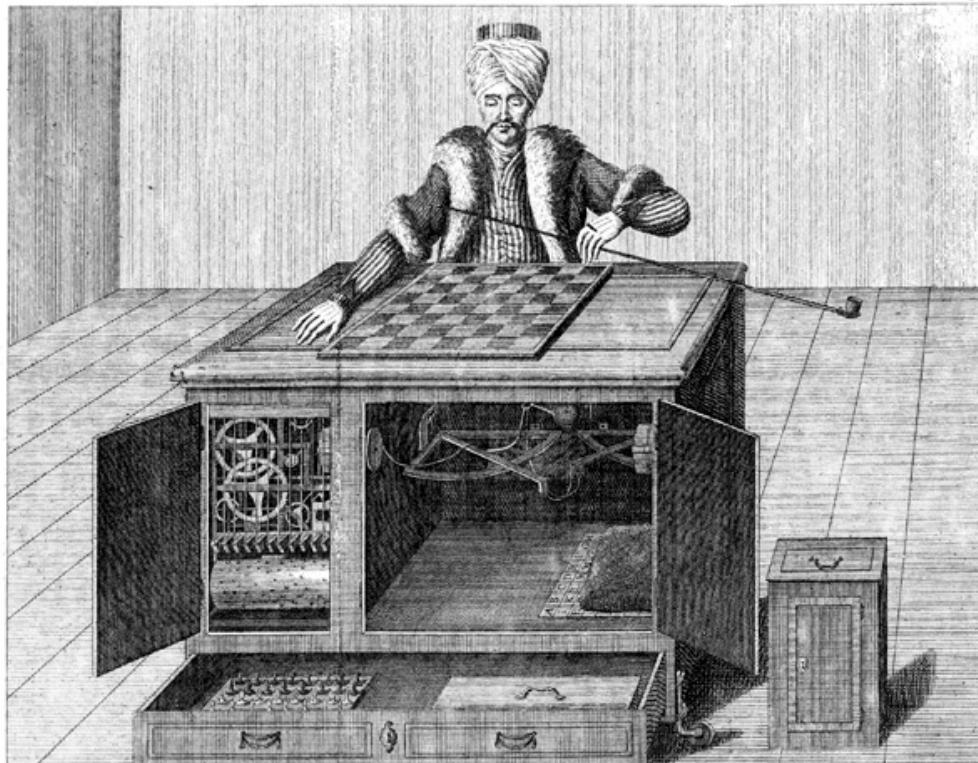


Rene Descartes wondered if he could know for sure that others who looked and behaved like him weren't in fact automata.

Bodies of people and animals are nothing more than complex machines - the bones, muscles and organs could be replaced with cogs, pistons and cams.



17th and 18th century automotons

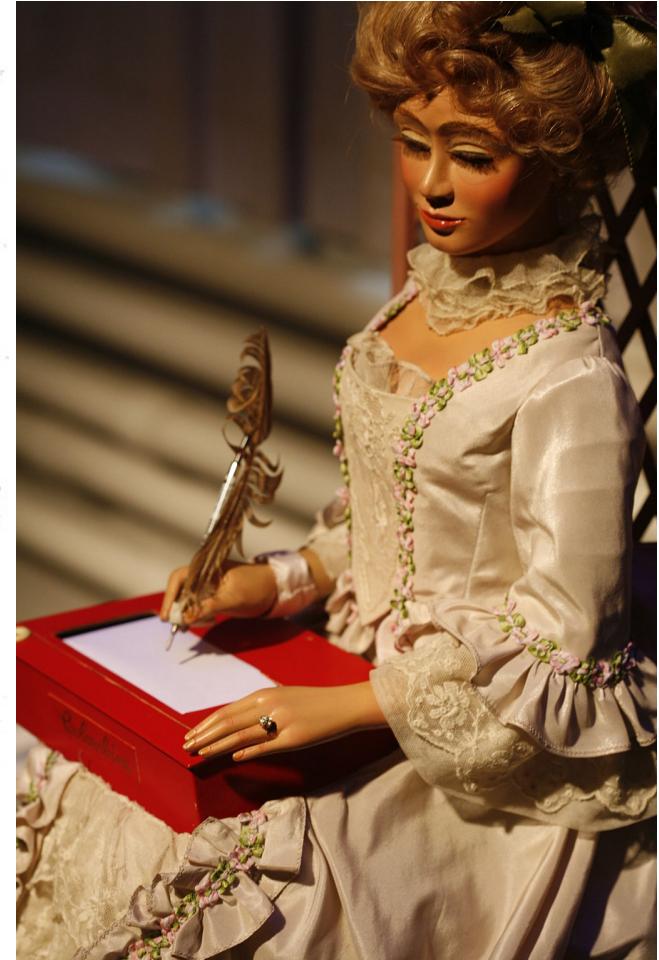


W. de Kempelen del.

Ch. a Meichel excud. Basilea.

P.G. Peutz sc:

Der Schachspielder wie er vor dem Spiel gezeigt wird von einem Le. Joucure d'echecs, tel qu'on le montre avant le jeu, par devant.



René Descartes (1596-1650)



“if there were machines bearing the image of our bodies, and capable of imitating our actions. For example, if touched in a particular place it may demand what we wish to say to it; if in another it may cry out that it is hurt. However there would be two tests to know that they were not really men.”... They could never use properly use language.



Blade Runner - Voight-Kampff Test



<https://www.youtube.com/watch?v=Umc9ezAyJv0>



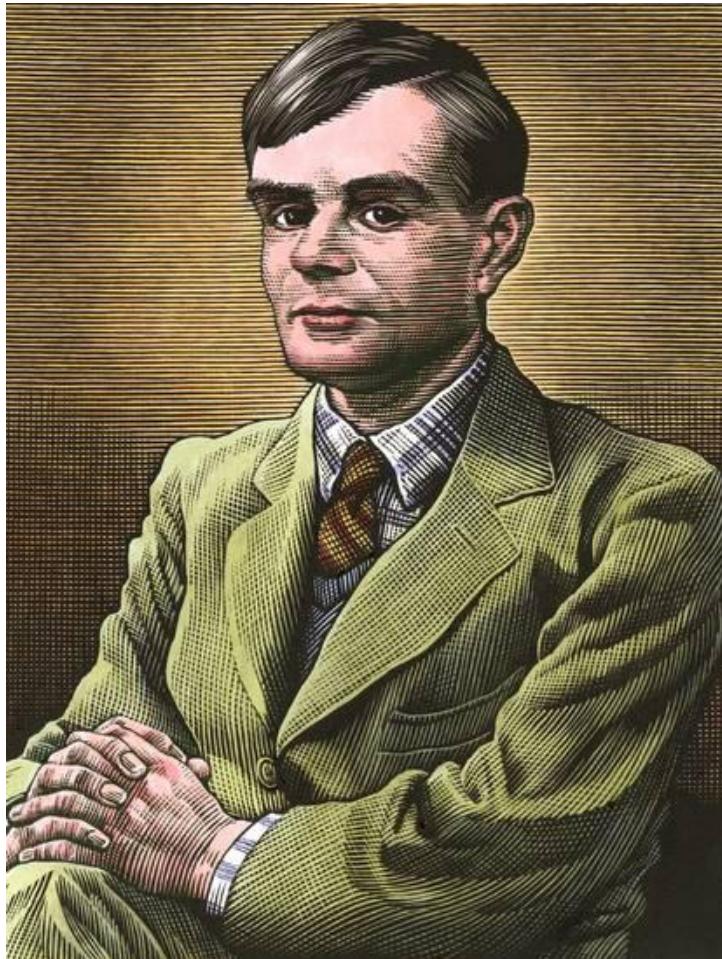
Descartes



Deckard



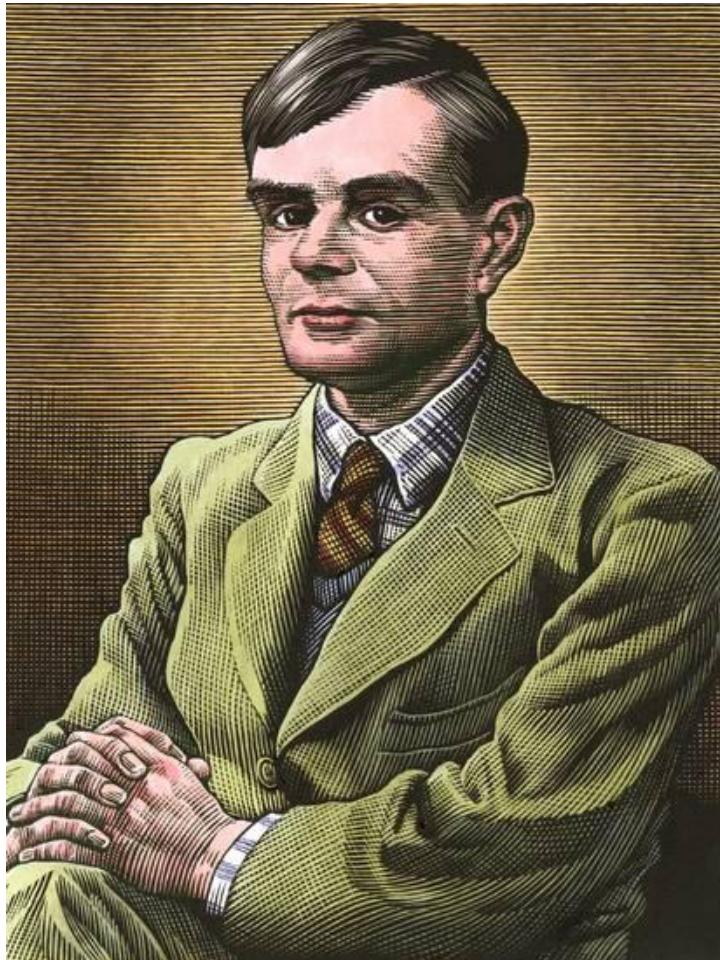
Alan Turing (1912-1954)



"I propose to consider the question, 'Can machines think?' Because "thinking" is difficult to define, Turing chooses to "replace the question by another, which is closely related to it and is expressed in relatively unambiguous words." Turing's new question is: "Are there imaginable digital computers which would do well in the *imitation game*?"



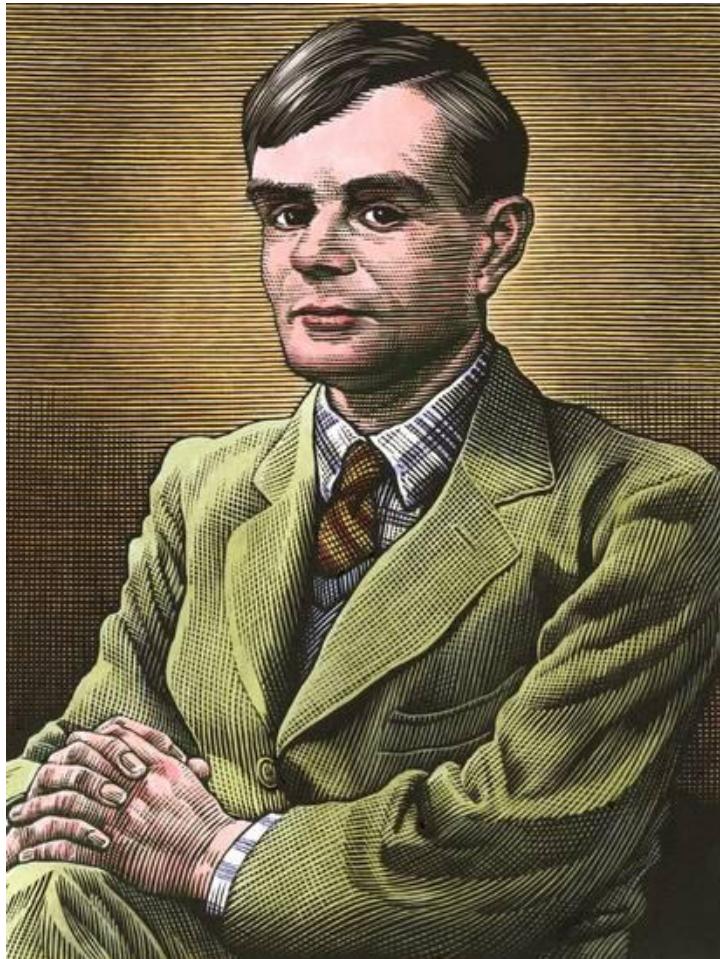
Alan Turing (1912-1954)



The “imitation game” is a party game, where a man and a woman go into separate rooms and guests try to tell them apart by writing a series of questions and reading the typewritten answers sent back. In this game both the man and the woman aim to convince the guests that they are the other.



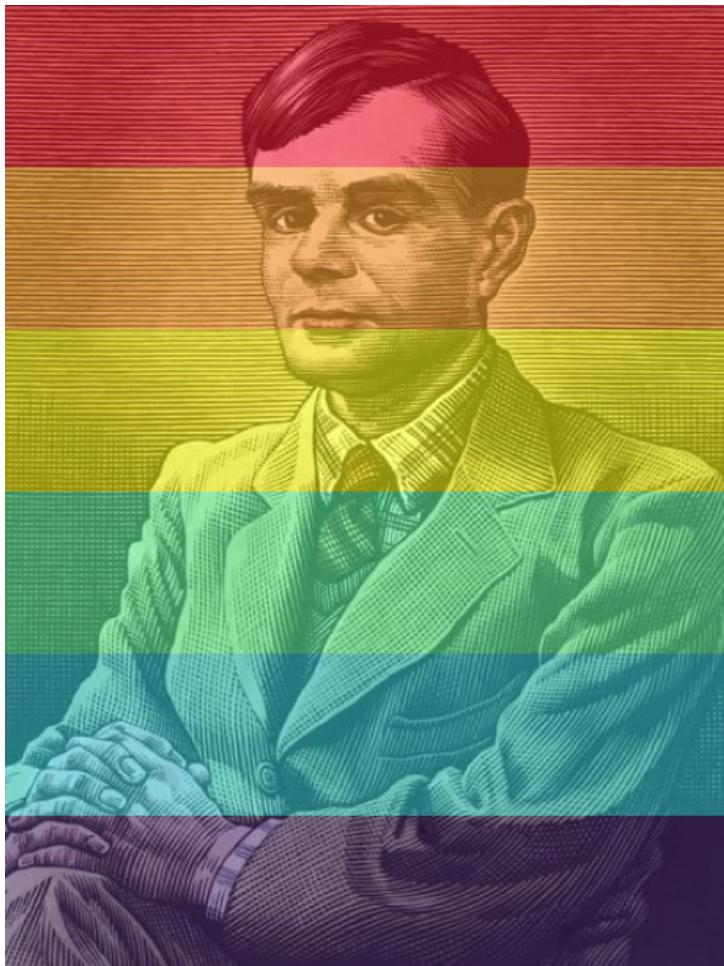
Alan Turing (1912-1954)



A human evaluator would judge text-based conversations between a human and a machine designed to generate human-like responses. If the evaluator cannot reliably tell the machine from the human, the machine is said to have passed the test. The test results do not depend on the ability to give correct answers to questions, only how closely one's answers resemble those a human would give.



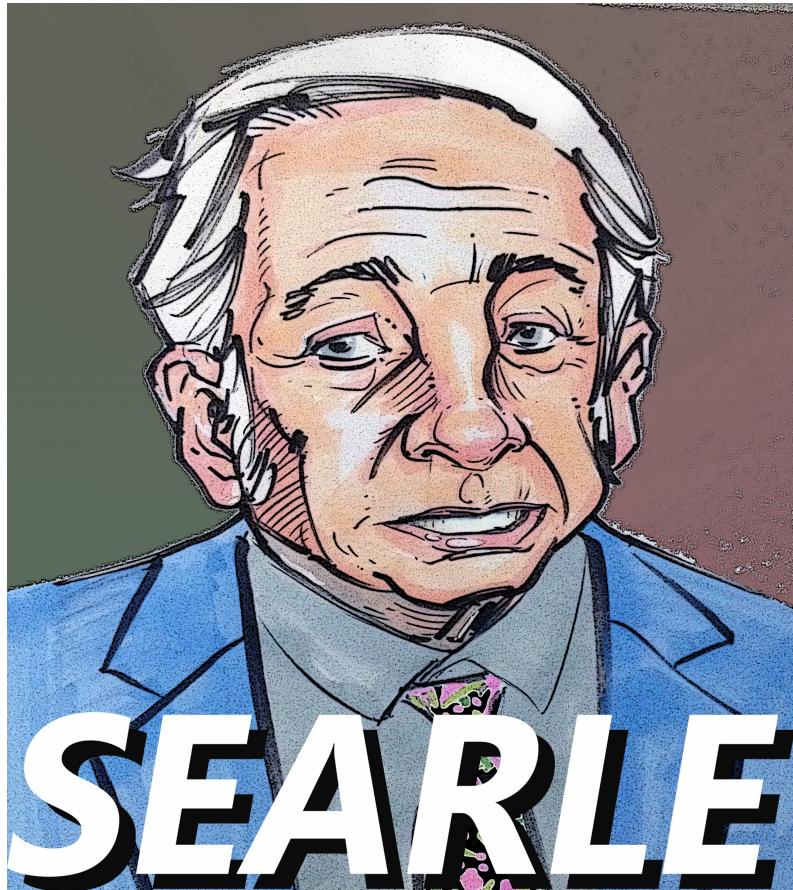
Alan Turing (1912-1954)



Turing was prosecuted in 1952 for being homosexual. He received chemical castration as an alternative to prison. Turing killed himself in 1954. It wasn't until the 2000s that Britain finally realized its gross injustice and its complicity in killing someone who should have been treated as a national hero.



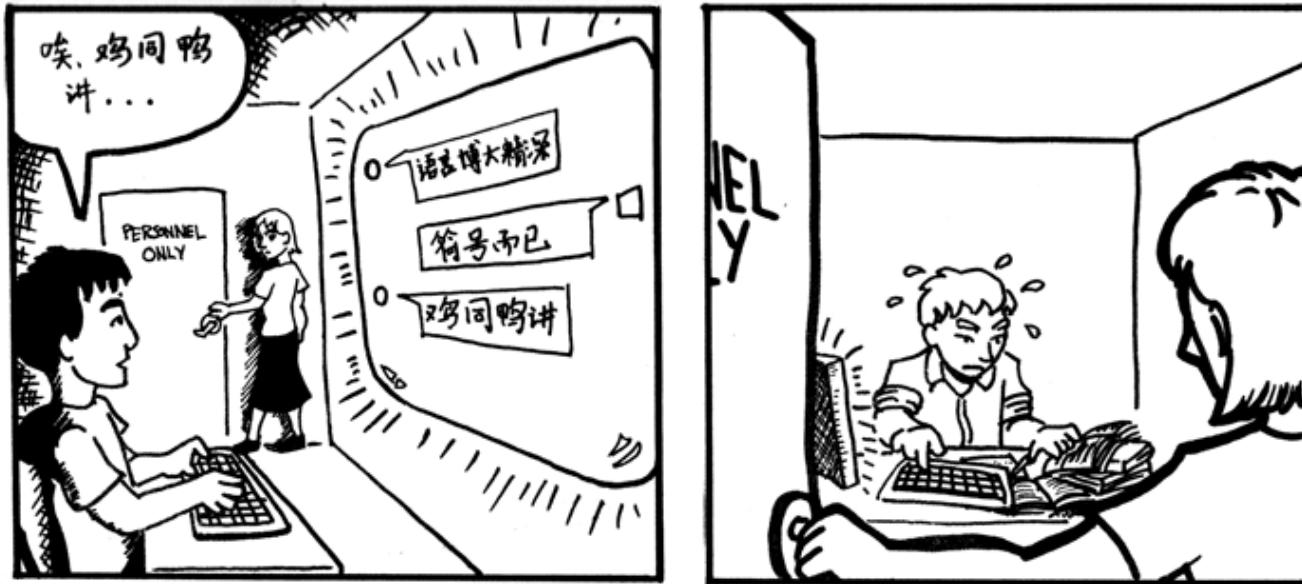
John Searle (1932-) The Chinese Room



Is the Turing Test a good test of whether machines possess intelligence? No, because the appearance of being a good conversational participant is achievable through simple symbol manipulation. Searle contends that the application of rules to input symbols is not true intelligence



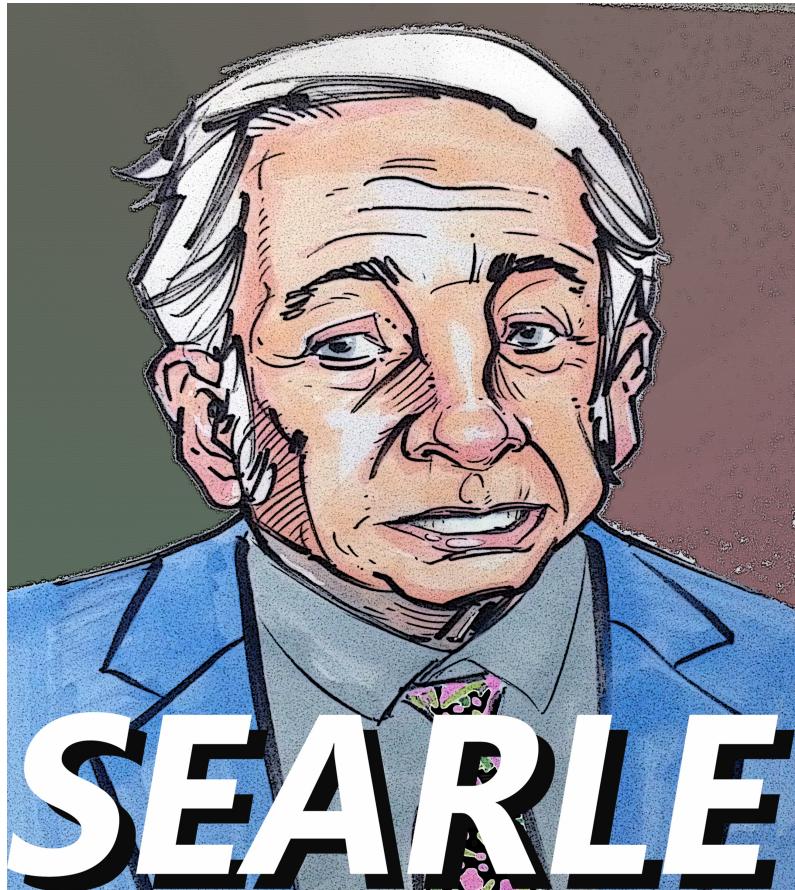
John Searle (1932-) The Chinese Room



© Daniel Dennett
and Neil Cohn



John Searle (1932-) The Chinese Room

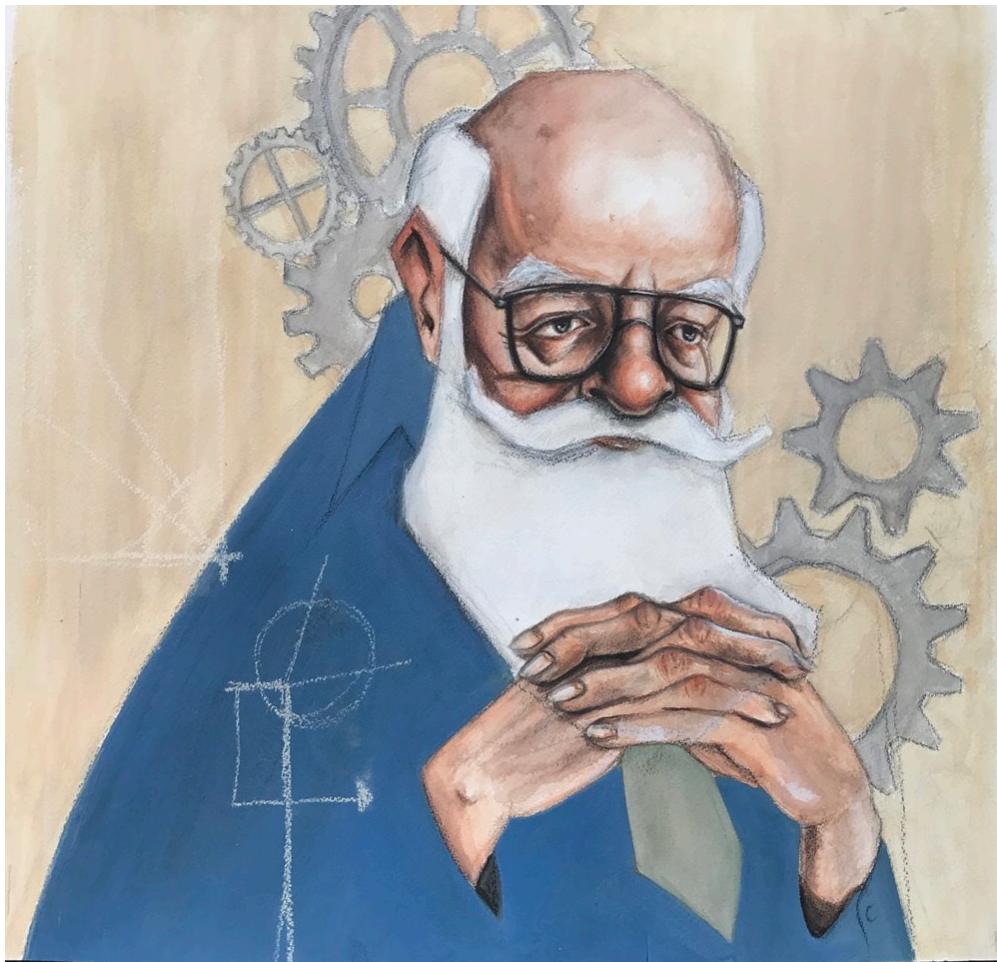


Searle contrasts strong AI with weak AI. In strong AI, the computer really *is* a mind in the sense that it can be literally said to *understand* and have other cognitive states. In *weak AI*, computers just *simulate* thought, their seeming understanding isn't real understanding.

He argues that (biological) *brains cause minds*.



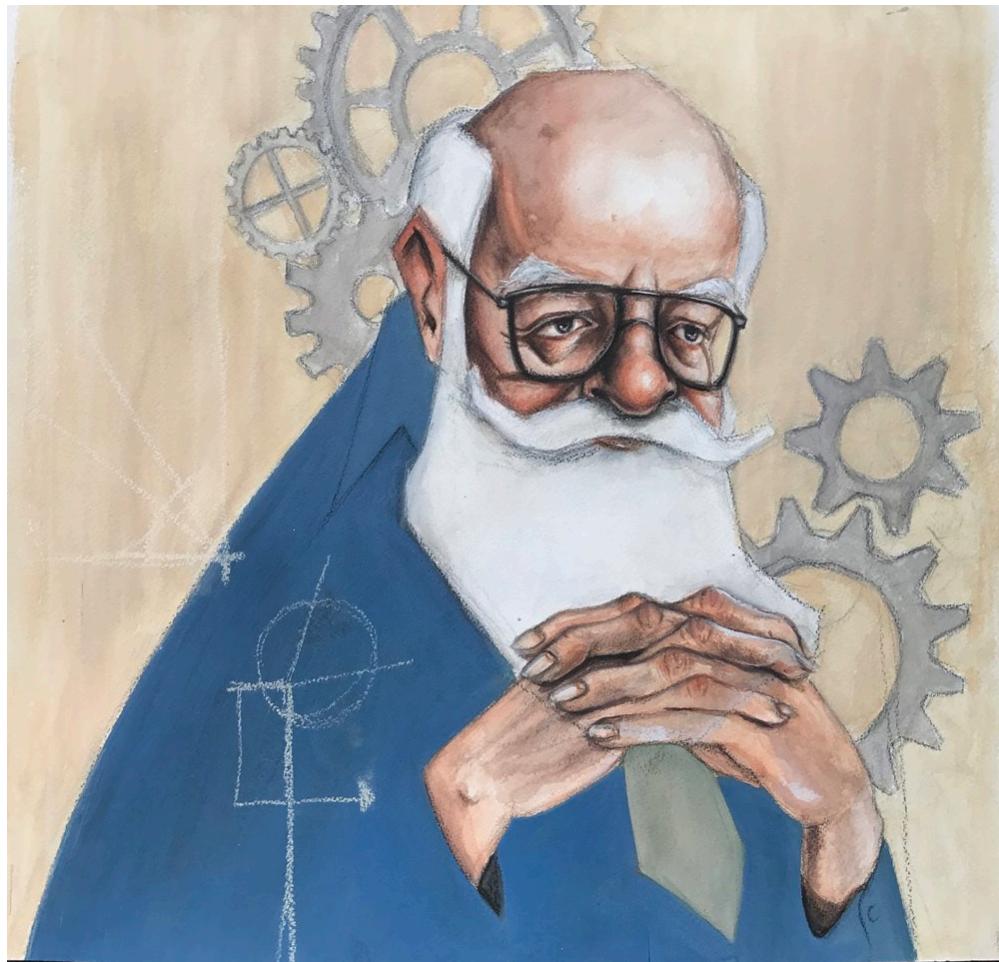
Daniel Dennett (1942-) Brain in a Vat



Daniel Dennett wrote a short story called “Where Am I?” where he describes being recruited by the Pentagon to have his brain removed from his body, and connected via radio links attaching his severed nerves. Body is sent on a secret mission to diffuse bomb without radiation harming brain.



Daniel Dennett (1942-) Brain in a Vat



His brain in a vat is processing symbols.

What makes this different than a computer processing symbols?



Brain in a vat

FILOSOFIX
Philosophy animates

<https://www.youtube.com/watch?v=zO0sSJB1TrI>



Strong AI versus Weak AI

- *Actually thinking* versus merely *simulating thinking*
- Are you a brain in a vat?
- Would an AI program be equivalent?
- We will leave this to the philosophers and instead focus on practical AI. Programs that work.

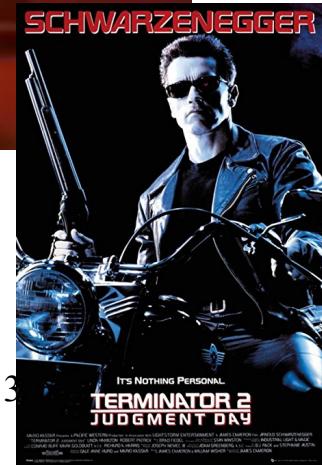
- (But I'll give you extra credit if you invent a sentient AI).



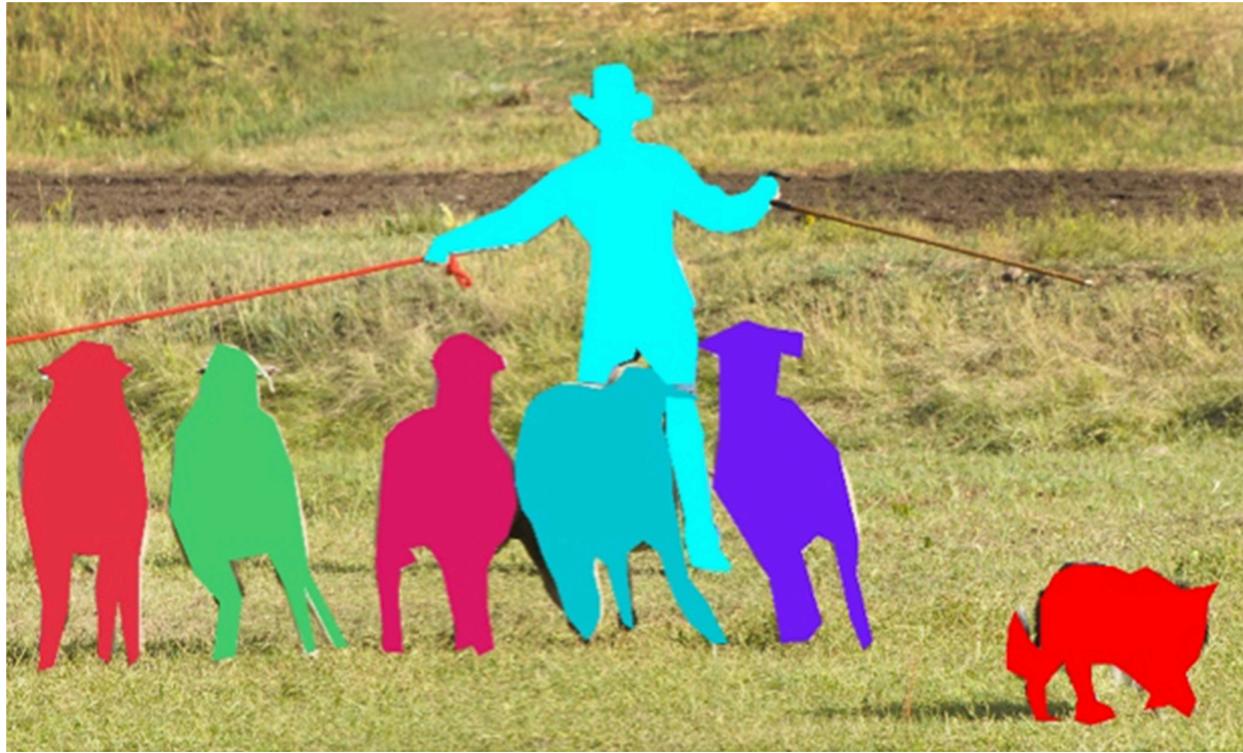
Science Fiction and Science Fact



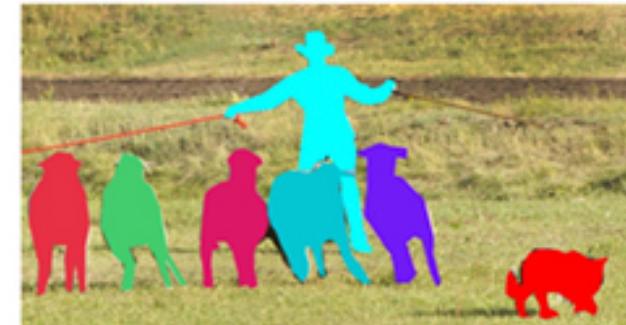
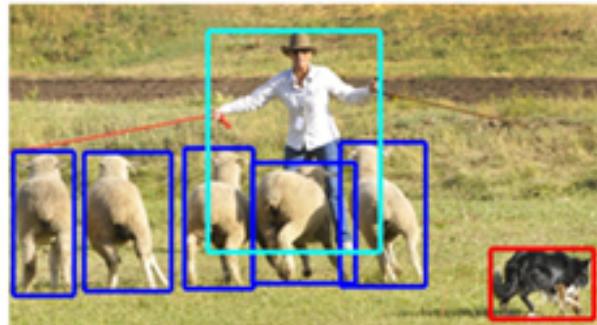
<https://www.youtube.com/watch?v=JknRYYrV5Vk>



Science Fiction and Science Fact



(a) classification



(c) segmentation

Science Fiction and Science Fact

Medium



Kenny Yin [Follow](#)

Developer. Cinematographer. Vocalist.

Jun 19, 2016 · 23 min read

Apple's updated Photos app recognizes thousands of objects, scenes and facial expressions



Science Fiction and Science Fact

Armchair, Armchairs, Art, Artistic Creation, Artistic Creations, Artistries, Artistry, Arts, Artwork, Artworks, Arthropods, Arthropod, Arthropods, Artichoke, Artichokes, Arugula, Arugulas, Garden Rocket, Garden Rockets, Rocket Salad, Rocket Salads, Roquette, Roquettes, Asparagus, Asparaguses, ATM, ATMs, Automated Teller, Automated Teller Machine, Automated Teller Machines, Automated Tellers, Automatic Teller, Automatic Teller Machine, Automatic Teller Machines, Automatic Tellers, Cash Dispenser, Cash Dispensers, Cash Machine, Cash Machines, ATV, ATVs, All Terrain Vehicle, All Terrain Vehicles, Auditorium, Amphitheater, Amphitheaters, Auditoriums, Concert Hall, Concert Halls, Music Hall, Music Halls, Aurora, Auroras, Australian Shepherd, Australian Shepherds, Automobile, Auto, Automobiles, Autos, Car, Cars, Motorcar, Motorcars, Avocado, Aguacate, Aguacates, Alligator Pear, Alligator Pears, Avocados, Axe, Ax, Axes, Baby, Babe, Babes, Babes, Infants, Infants

Science Fiction and Science Fact



<https://www.youtube.com/watch?v=s93KC4AGKnY>



Science Fiction and Science Fact



Science Fiction and Science Fact



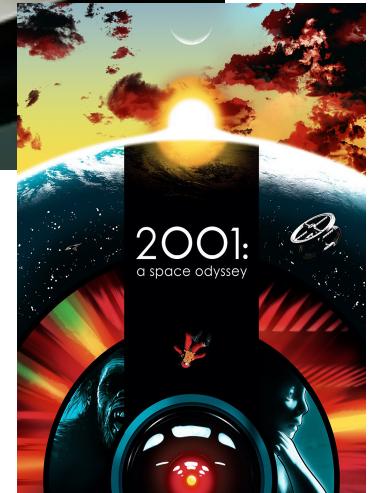
Science Fiction and Science Fact



Science Fiction and Science Fact



<https://www.youtube.com/watch?v=Lk8qgMFqaY8>



Science Fiction and Science Fact



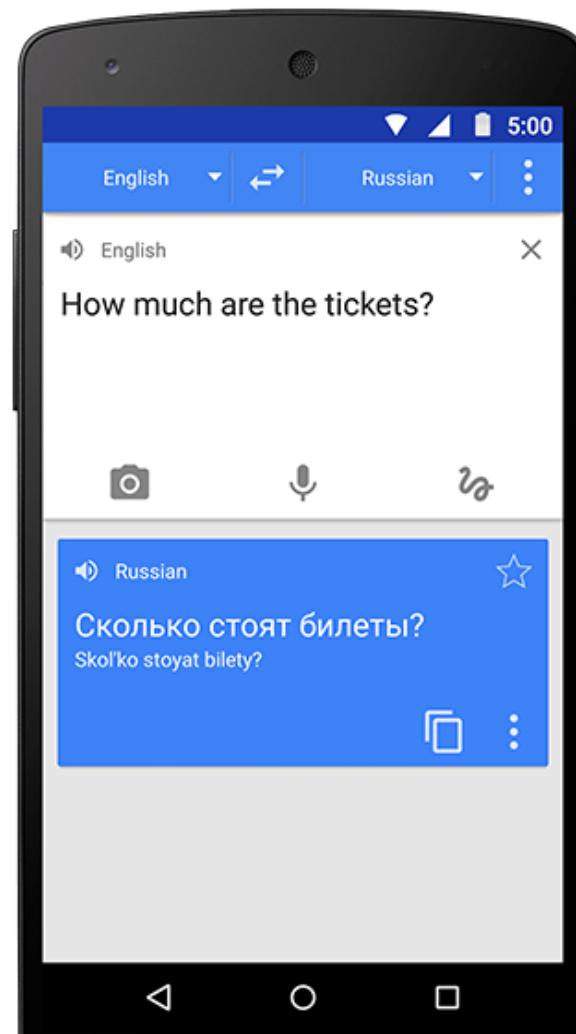
Science Fiction and Science Fact



Science Fiction and Science Fact



Science Fiction and Science Fact



Science Fiction and Science Fact



Science Fiction and Science Fact



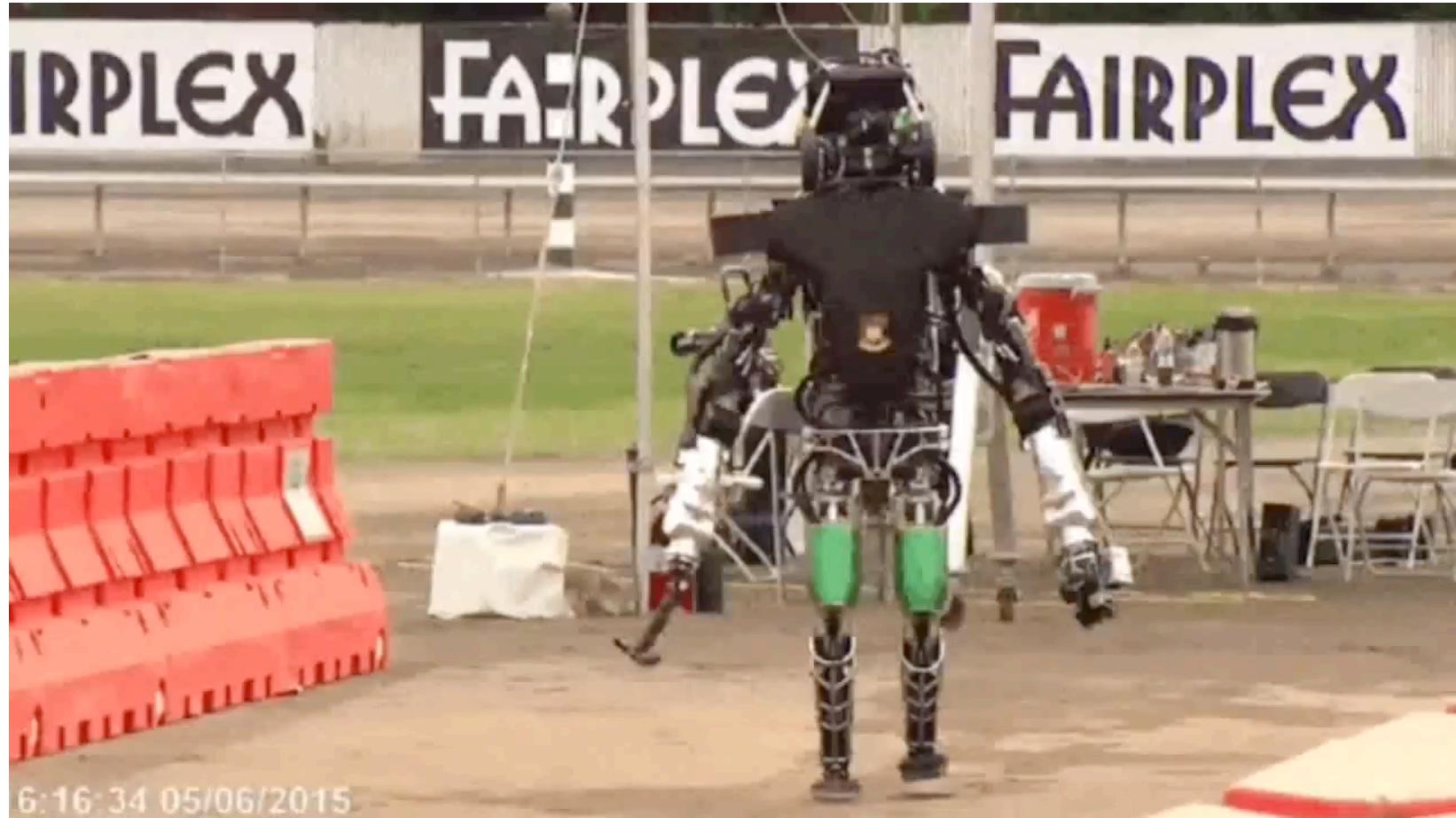
Science Fiction and Science Fact



Science Fiction and Science Fact



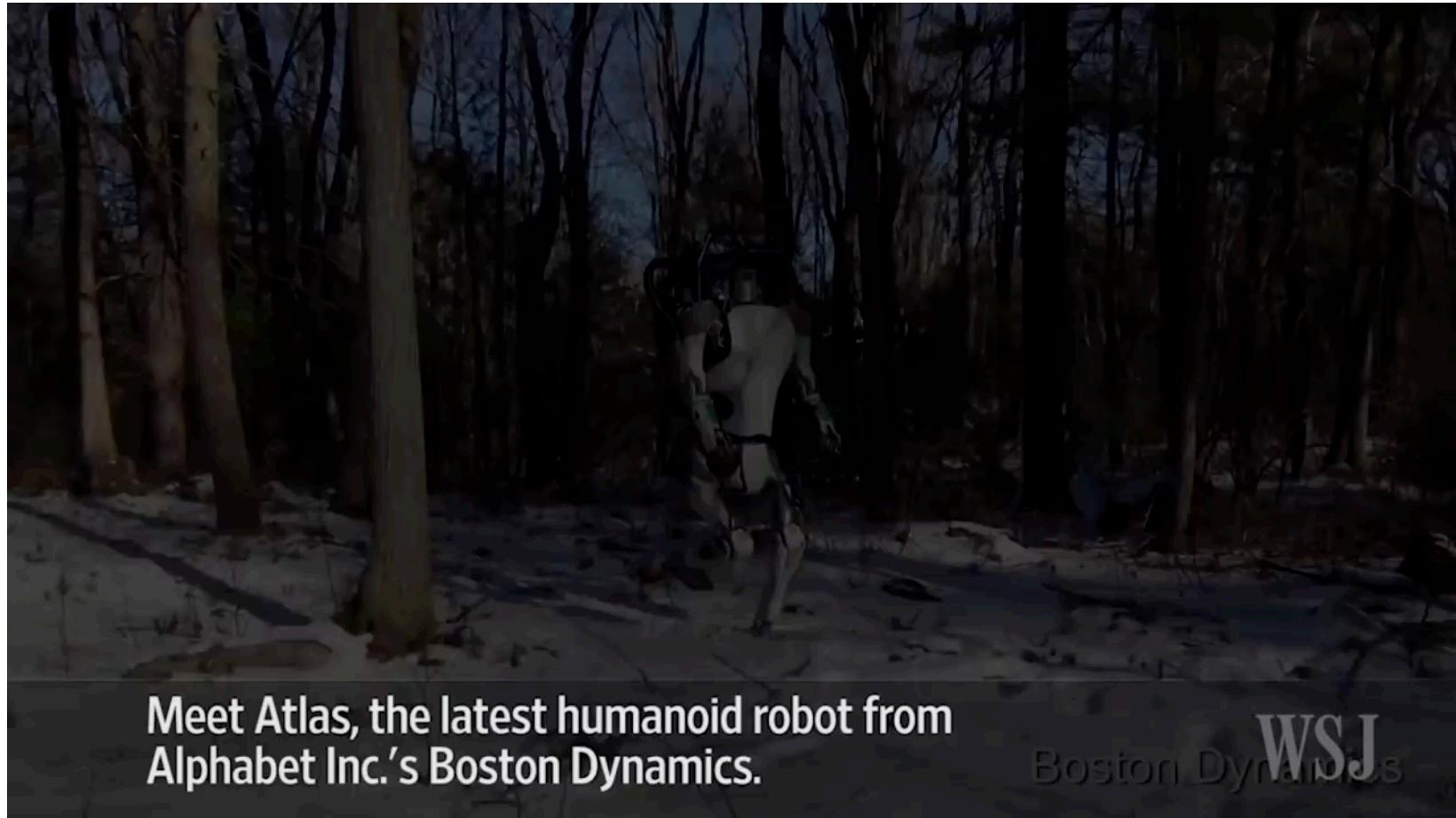
Science Fiction and Science Fact



<https://www.youtube.com/watch?v=g0TaYhjpOfo>



Science Fiction and Science Fact



Meet Atlas, the latest humanoid robot from
Alphabet Inc.'s Boston Dynamics.

Boston Dynamics **WSJ**

<https://www.youtube.com/watch?v=M91ISnATDQY>



Applied AI

- **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**



Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

What is AI?

Views of AI fall into four categories:

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

We will focus on "acting rationally"



Thinking humanly: Cognitive modeling

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

- 1960s "cognitive revolution": information-processing psychology, a.k.a. *Cognitive Psychology*
- Requires scientific theories of internal activities of the brain
- How to validate? Requires
 - 1) Predicting and testing behavior of human subjects
or
 - 2) Direct identification from neurological data (bottom-up) :
Cognitive Neuroscience



Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

Thinking rationally: “Laws of thought” aka Logic

- Aristotle: what are correct arguments/thought processes?
- Several Greek schools developed various forms of *logic*: *notation* and *rules of derivation* for thoughts; may or may not have proceeded to the idea of mechanization
- Direct line through mathematics and philosophy to modern AI
- Problems:
 1. Not all intelligent behavior is mediated by logical deliberation
 2. What is the purpose of thinking? What thoughts should I have?
 3. Ignores the hard problem of perception
 4. *All attempts to encode what we know in logic have failed*
 5. *Most logical inference is intractable*



Acting rationally: rational agents

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

- **Rational** behavior: doing the right thing
- The right thing: that which is *expected to maximize goal achievement, given the available information*
- Doesn't necessarily involve thinking – e.g., blinking reflex – but thinking should be in the service of rational action



Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

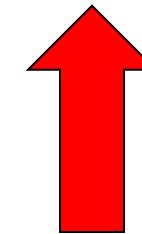
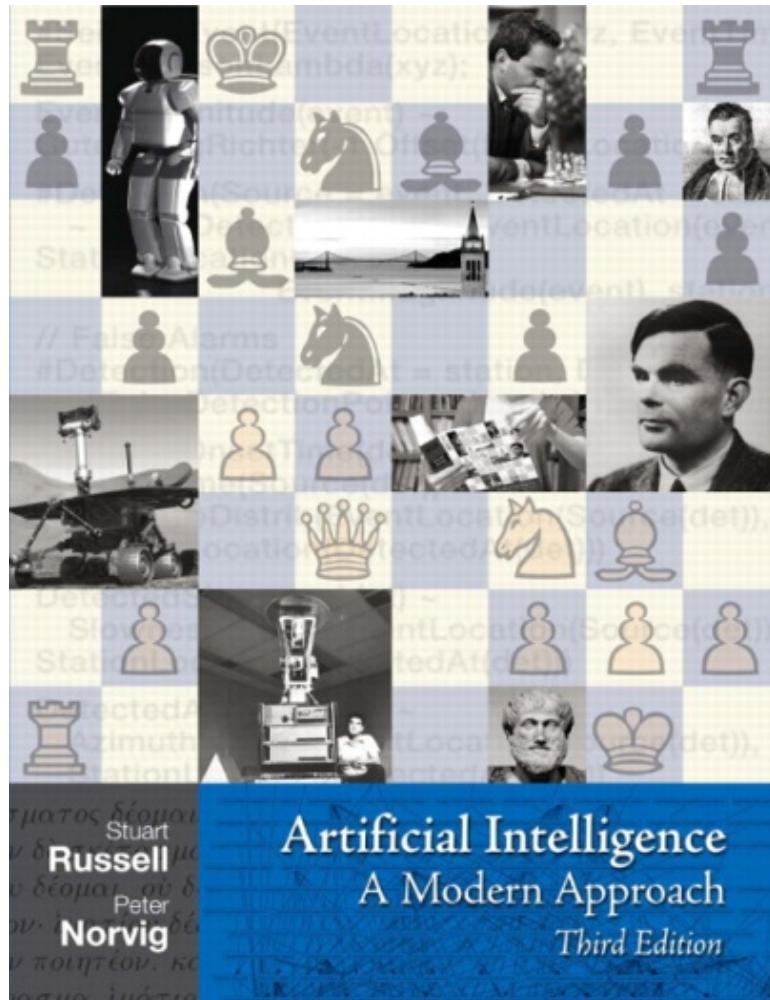
Rational agents

Rational agent: An *agent* is an entity that perceives and acts

This course is about *effective programming techniques* for designing *rational agents*



Homework: Read Chapters 1 and 26



1 and 26,
not 1 through 26



Next Two Lectures

A Brief Introduction to Python for those who know Java

The first Python
programming HW
has been released.

It is due Sept 4.

Find it on the
course web site:

<http://artificial-intelligence-class.org/>

