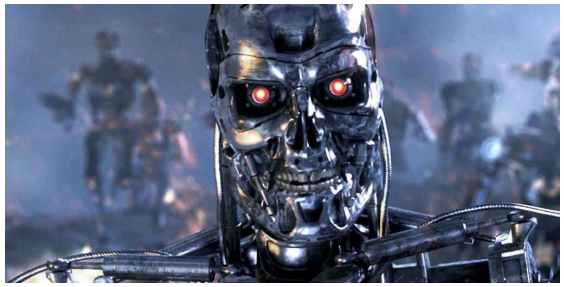
CS440/ECE448: Artificial Intelligence
Lecture 1: What is AI?

Slides by Svetlana Lazebnik, 9/2016
Modified by Mark Hasegawa-Johnson, 8/2017



CS440/ECE448: Artificial Intelligence

Website:
http://courses.engr.Illinois.edu/ece448/

What is AI?

- Candidate definitions from the textbook:

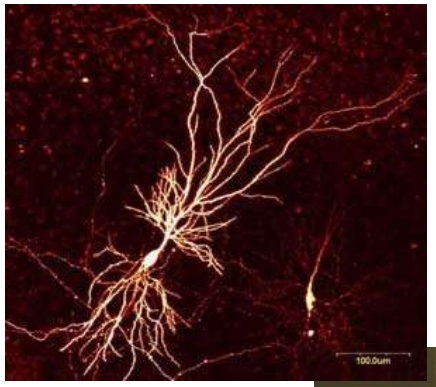
1. Thinking humanly	**2. Acting humanly**
3. Thinking rationally	**4. Acting rationally**

Defining AI: Narrow vs. Broad AI

- What's the advantage of defining AI in terms of how you act, rather than how you think?
  - Sufficient to pass Turing test
  - Lots easier to measure
  - Thought process might be alien; that's OK
  - Possible to use "intelligence" to describe behavior of materials etc.
- What's the advantage of defining AI in terms of how you think, rather than how you act?
  - React to new situations not previously planned/observed
  - Easier to improve it because you know rationale behind actions
  - Doesn't have to always respond to same situation in same way
  - Space of actions is too large; can only define AI in terms of strategy

AI definition 1: Thinking humanly

- Need to study the brain as an information processing machine: cognitive science and neuroscience



AI definition 1: Thinking humanly

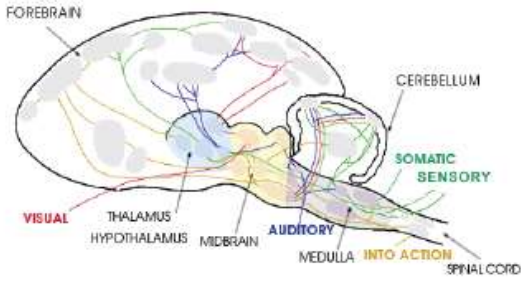
- Should we build a brain?

  The “embodied cognition” movement exemplified: https://www.youtube.com/watch?v=uoXAPSdXiZg

Source: L. Zettlemoyer

AI definition 1: Thinking humanly

- CAN we build a brain?



10^{11} neurons
10^{14} synapses
cycle time: 10^{-3} sec

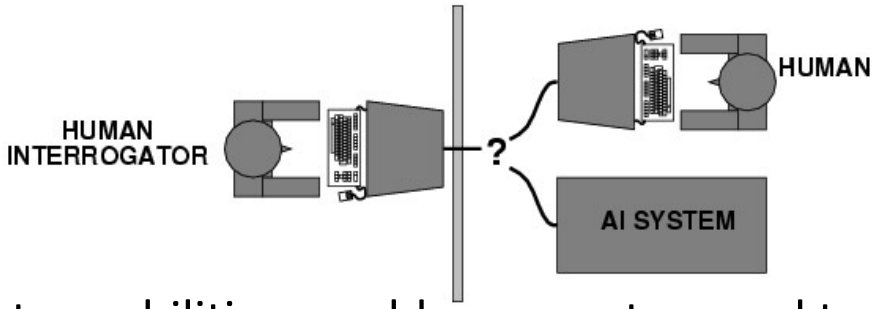
vs.

10^{9} transistors
10^{12} bits of RAM
cycle time: 10^{-9} sec

- The problem: parallel versus serial!!!!!

AI definition 2: Acting humanly

- **The Turing Test: https://www.youtube.com/watch?v=3wLqsRLvV-c**



- What capabilities would a computer need to have to pass the Turing Test?
  - Natural language processing
  - Knowledge representation
  - Automated reasoning
  - Machine learning
- Turing predicted that by the year 2000, machines would be able to fool 30% of human judges for five minutes

A. Turing, Computing machinery and intelligence, Mind 59, pp. 433-460, 1950

AI is solved?

Computer AI passes Turing test in 'world first'

Eugene Goostman simulates a 13-year-old Ukrainian boy

9 June 2014 Last updated at 08:36 ET

A computer program called Eugene Goostman, which simulates a 13-year-old Ukrainian boy, is said to have passed the Turing test at an event organised by the University of Reading.

The test investigates whether people can detect if they are talking to machines or humans.

The experiment is based on Alan Turing's question-and-answer game **Can Machines Think?**

No computer has passed the test before under these conditions, it is reported.

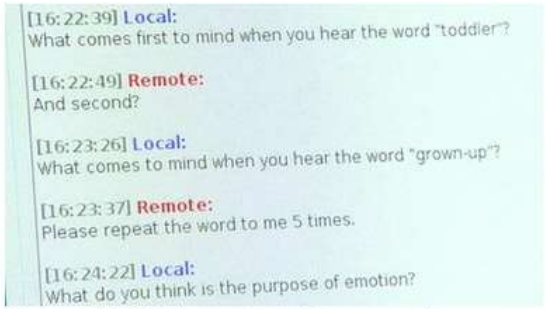
However, some artificial intelligence experts have disputed the victory, suggesting the contest had been weighted in the chatbot's favour.

The 65-year-old **Turing Test** is successfully passed if a computer is mistaken for a human more than 30% of the time during a series of five-minute keyboard conversations.

On 7 June Eugene convinced 33% of the judges at the Royal Society in London that it was human.

Other artificial intelligence (AI) systems also competed, including **Cleverbot**, **Elbot** and **Ultra Hal**.

Judges included actor Robert Llewellyn, who played an intelligent robot in BBC Two's science-fiction sitcom Red Dwarf, and Lord Sharkey, who led the successful campaign for Alan Turing's posthumous pardon, over a conviction for homosexual activity, in 2013.



Man or machine? A glimpse at one of the conversations.

Eugene was created by Vladimir Veselov, who was born in Russia and now lives in the United States, and Ukrainian-born Eugene Demchenko, who now lives in Russia.

Related Stories

How the Turing Test inspired AI

Playing solitaire with Turing

Is artificial intelligence possible?

http://www.bbc.com/news/technology-27762088

http://en.wikipedia.org/wiki/Eugene_Goostman

http://www.scottaaronson.com/blog/?p=1858

What's wrong with the Turing test?

- Variability in protocols, judges
- Success depends on deception!
- Chatbots can do well using "cheap tricks"
  - First example: ELIZA (1966)
  - Javascript implementation of ELIZA

A better Turing test?

- **Winograd schema:** Multiple choice questions that can be easily answered by people but cannot be answered by computers using “cheap tricks”

•*The trophy would not fit in the brown suitcase because it was so small.*
What was so small?

- *The trophy*
- *The brown suitcase*

H. Levesque, *On our best behaviour*, IJCAI 2013

http://www.newyorker.com/online/blogs/elements/2013/08/why-cant-my-computer-understand-me.html

A better Turing test?

- **Winograd schema:** Multiple choice questions that can be easily answered by people but cannot be answered by computers using "cheap tricks"

•*The trophy would not fit in the brown suitcase because it was so **large**.*
*What was so **large**?*

- *The trophy*
- *The brown suitcase*

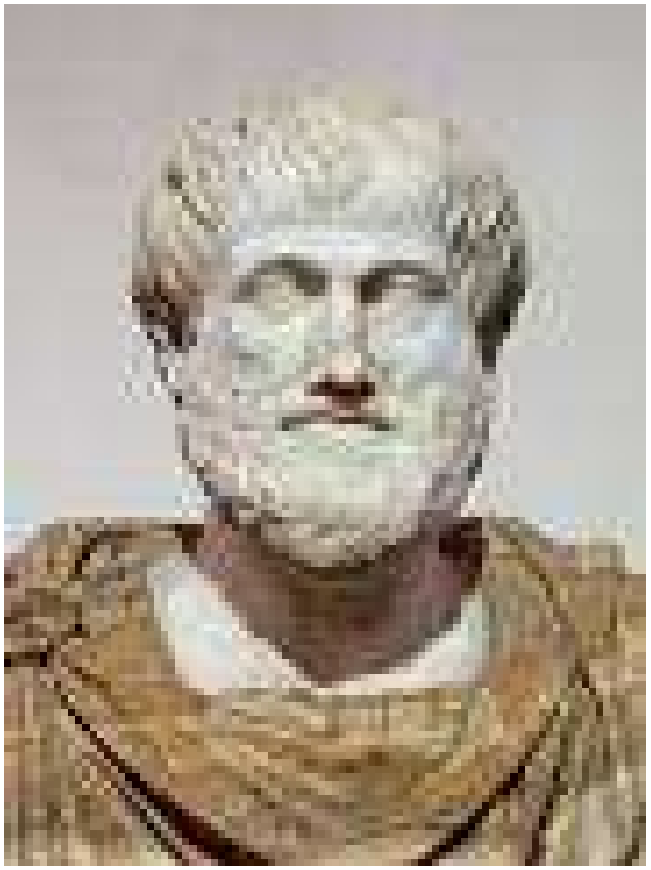
H. Levesque, *On our best behaviour*, IJCAI 2013

http://www.newyorker.com/online/blogs/elements/2013/08/why-cant-my-computer-understand-me.html

Winograd schema

- Advantages over standard Turing test
  - Test can be administered and graded by machine
  - Does not depend on human subjectivity
  - Does not require ability to generate English sentences
  - Questions cannot be evaded using verbal dodges
  - Questions can be made “Google-proof” (at least for now…)
- Winograd schema challenge
  - Held at IJCAI conference in July 2016
  - Six entries, best system got 58% of 60 questions correct (humans get 90% correct)

AI definition 3: Thinking rationally

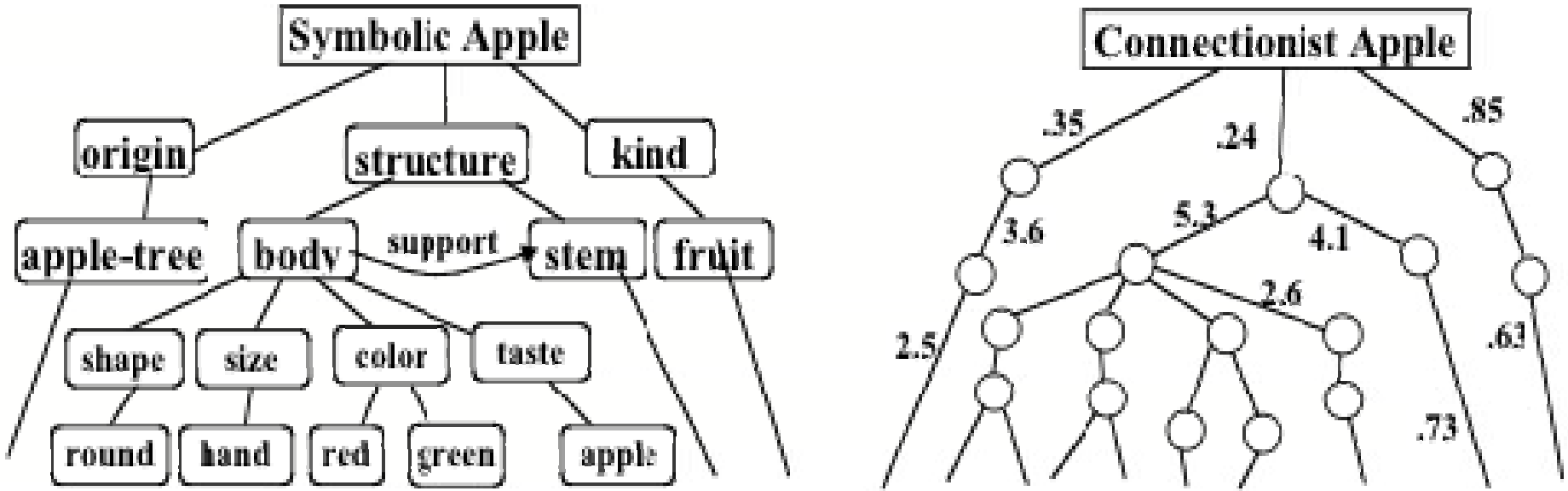


Aristotle, 384-322 BC

AI definition 3: Thinking rationally

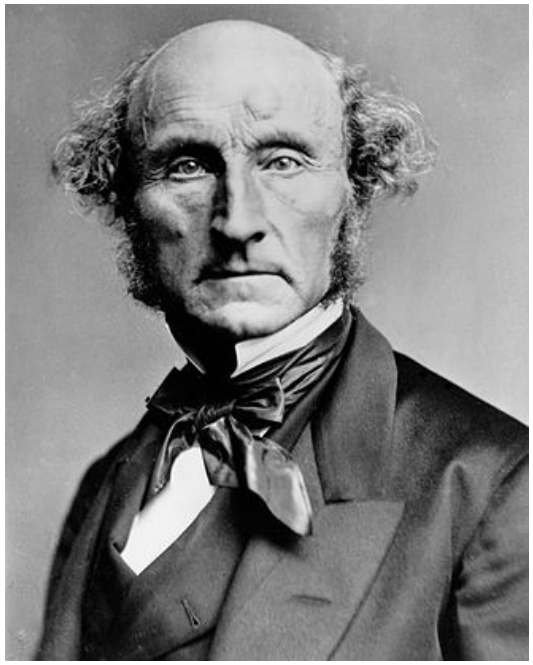
- Idealized or “right” way of thinking
- **Logic:** patterns of argument that always yield correct conclusions when supplied with correct premises
  - *“Socrates is a man; all men are mortal; therefore Socrates is mortal.”*
- **Logicist approach to AI:** describe problem in formal logical notation and apply general deduction procedures to solve it
- Problems with the logicist approach
  - Computational complexity of finding the solution
  - Describing real-world problems and knowledge in logical notation
  - Dealing with uncertainty
  - A lot of “rational” behavior has nothing to do with logic

AI definition 3: Thinking rationally



- Most of the approaches in this class (search … Bayesian nets) are capable of incorporating a logist approach (left), though they might go beyond it
- Perceptrons and deep learnings usually give up the symbolic representation in favor of better end-to-end performance (right)

AI definition 4: Acting rationally



John Stuart Mill, 1806-1873

AI definition 4: Acting rationally

- A **rational agent** acts to optimally achieve its goals
  - Goals are application-dependent and are expressed in terms of the **utility of outcomes**
  - Being rational means **maximizing your (expected) utility**
- This definition of rationality only concerns the decisions/actions that are made, not the cognitive process behind them
- An unexpected step: rational agent theory was originally developed in the field of economics
  - Norvik and Russell: "most people think Economists study money. Economists think that what they study is the behavior of rational actors seeking to maximize their own happiness."
- By standard economic definitions: a rational agent is any agent that acts rationally, e.g., human 3-year-olds:

•https://www.youtube.com/watch?v=EjzadtXR-zs

Utility maximization formulation

- Advantages
  - Generality: goes beyond explicit reasoning, and even human cognition altogether
  - Practicality: can be adapted to many real-world problems
  - Naturally accommodates uncertainty
  - Amenable to good scientific and engineering methodology
  - Avoids philosophy and psychology
- Disadvantages
  - Practical disadvantage: In practice, utility optimization is subject to the agent's computational constraints (**bounded rationality** or **bounded optimality**)
  - Theoretical disadvantage: does being human involve anything other than rationality? Are the other aspects of being human important? Would it be useful if a machine had them, whatever they are?
  - What is utility?

What is AI?

1. What is the advantage of defining AI in terms of action, rather than thought? What is its disadvantage?
2. What is the advantage of defining AI in terms of being rational, instead of being human-like? What is its disadvantage?
3. What is the Turing test? What are some problems with the Turing test? How might they be solved? What is the Winograd schema? Why is it better than the Turing test? Why might it fail?
4. What is the logicist approach to AI?
5. What is a rational agent? What is utility?