

# Artificial Intelligence: From NAND to Consciousness

Mark Maloof

Department of Computer Science  
Georgetown University  
Washington, DC 20057-1232  
<http://www.cs.georgetown.edu/~malooof>

IDST 010-06

Fall 2016

## Video: Elon Musk



## Video: The Great Robot Race



## Video: Self-Driving Car Test: Steve Mahan



## McCarthy et al., 1955

- ▶ “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.”

## Haugeland, 1985

- ▶ “The exciting new effort to make computers think...machines with minds, in the full and literal sense.”

## Charniak and McDermott, 1985

- ▶ “...the study of mental faculties through the use of computational models.”

- ▶ “Artificial intelligence, broadly (and somewhat circularly) defined, is concerned with intelligent behavior in artifacts. Intelligent behavior, in turn, involves perception, reasoning, learning, communicating, and acting in complex environments.”



# Disciplines Important for AI

- ▶ biology
- ▶ computer science
- ▶ electrical engineering
- ▶ linguistics
- ▶ mathematics
- ▶ mechanical engineering
- ▶ neuroscience
- ▶ philosophy
- ▶ psychology

# Russell and Norvig's Four Approaches

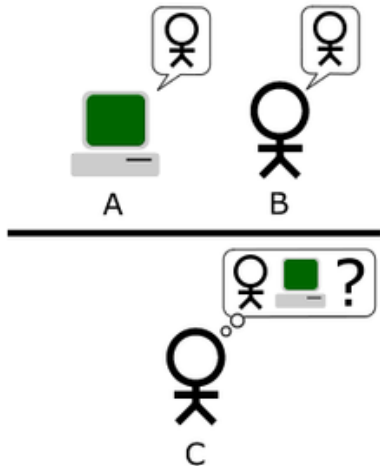
1. Think like a human
2. Act like a human
3. Think rationally
4. Act rationally

# Think Like A Human

- ▶ “...machines with minds, in the full and literal sense”
- ▶ Put simply, program computers to do what the brain does
- ▶ How do humans think?
- ▶ What is thinking, intelligence, consciousness?
- ▶ If we knew, can computers do it, think like humans?
- ▶ Does the substrate matter, silicon versus meat?
- ▶ Computers and brains have completely different architectures
- ▶ Is the brain carrying out computation?
- ▶ If not, then what is it?
- ▶ Can we know ourselves well enough to produce intelligent computers?

# Act Like A Human

## Turing Test



## Obligatory xkcd Comic



# The Brilliance of the Turing Test

- ▶ Sidesteps the hard questions:
  - ▶ What is intelligence?
  - ▶ What is thinking?
  - ▶ What is consciousness?
- ▶ If humans can't tell the difference between human intelligence and artificial intelligence, then that's it
- ▶ Proposed in 1950, Turing's Imitation Game is still relevant

# Think Rationally

- ▶ Think rationally? Think logic!
- ▶ Put simply, write computer programs that carry out logical reasoning
  - ▶ Logic: propositional, first-order, modal, temporal, ...
  - ▶ Reasoning: deduction, induction, abduction, ...
- ▶ Possible problem: Humans don't really think logically
- ▶ Do we care? Strong versus weak AI
- ▶ One problem: often difficult to establish the truth or falsity of premises
- ▶ Another: conclusions aren't strictly true or false

# Act Rationally

- ▶ Act rationally? Think probability and decision theory!
- ▶ “A rational agent is one that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome” (Russell and Norvig, 2010, p. 4)
- ▶ `<jab>` “when there is uncertainty” `</jab>`
- ▶ When *isn't* there uncertainty?
- ▶ Predominant approach to AI (for now)



# Artificial Intelligence: From NAND to Consciousness

Mark Maloof

Department of Computer Science  
Georgetown University  
Washington, DC 20057-1232  
<http://www.cs.georgetown.edu/~malooof>

IDST 010-06

Fall 2016

# References I

- E. Charniak and D. McDermott. *Introduction to Artificial Intelligence*. Addison-Wesley, Reading, MA, 1985.
- J. Haugeland. *Artificial intelligence: The very idea*. MIT Press, Cambridge, MA, 1985.
- J. McCarthy, M. I. Minsky, N. Rochester, and C. E. Shannon. A proposal for the Dartmouth summer research project on artificial intelligence, 1955. URL <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>. [Online; accessed 7 August 2014].
- N. J. Nilsson. *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann, San Francisco, CA, 1998.
- S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Upper Saddle River, NJ, 3rd edition, 2010.