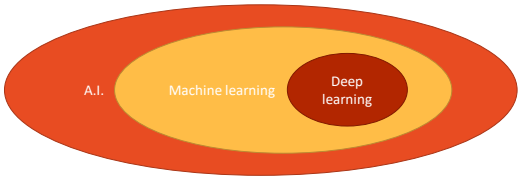
Hierarchy



A Venn diagram of A.I. – Goodfellow, I., Bengio, Y., & Courville, A. (2016).

9

9

What is AI?

- Systems that think like humans
- Systems that think rationally
- Systems that act like humans
- Systems that act rationally

10

10

Acting humanly: The Turing test

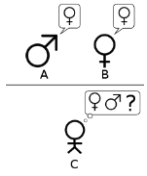
- Turing (1950) "Computing machinery and intelligence"
  - "Can machines think?"
  - "Can machines behave intelligently?"
  - Operational test for intelligent behavior?

11

11

The (Original) Imitation Game

- Player C, through a series of written questions, attempts to determine which of the other two players is a man, and which of the two is the woman.
- Player A, the man, tries to trick player C into making the wrong decision
- Player B tries to help player C

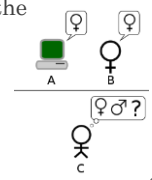


12

12

The Imitation Game

- The original imitation game test, in which the player A is replaced with a computer.
- The computer is now charged with the role of the man,
- Player B continues to attempt to assist the interrogator.



13

13

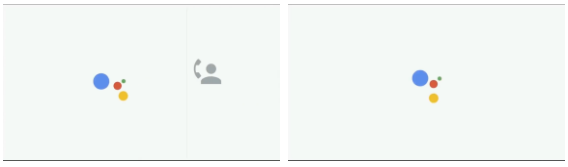
The Turing test

- Suggested major components of AI
  - Natural language understanding
  - Knowledge representation
  - Automated reasoning
- Predicted that by 2000, a machine might have a 30% chance of fooling a person for 5 minutes
- Problem: Turing test is not reproducible or not amenable to mathematical analysis

14

14

Google passes Turing Test?



15

15

Thinking humanly: Cognitive Science

- 1960s "cognitive revolution"
  - Information-processing psychology replaced prevailing orthodoxy of behaviorism
- Requires scientific theories of internal activities of the brain
- Once we have a sufficiently precise theory of the mind, it becomes possible to express the theory as a computer program.

16

16

Thinking rationally: Laws of Thought

- Aristotle
  - What are correct arguments/thought processes?
- Several Greek schools developed various forms of logic
  - Notation and rules of derivation for thoughts

17

17

Propositional Logic

- We have the two premises:
  - "All men are mortal."
  - "Socrates is a man."
- And the conclusion:
  - "Socrates is mortal."

$$\frac{\forall x(Man(x) \rightarrow Mortal(x)) \quad Man(Socrates)}{\therefore \quad Mortal(Socrates)}$$

- Problems:
  - Not all intelligent behavior is mediated by logical deliberation

18

18

Acting rationally

- Rational behavior: doing the right thing
- The right thing
  - that which is expected to maximize goal achievement, given the available information

19

19

Rational agents

- An agent is an entity that perceives and acts
- This course is about designing rational agents
- Abstractly, an agent is a function from percept histories to actions

20

20

Rational agents

- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- Caveat
  - Computational limitations make perfect rationality unachievable
  - design best program for given machine resources

21