Polytech Montpellier – IG4

Artificial Intelligence & Multi-Agent Systems

Introduction

Intelligence – definition?

- Intelligence is difficult to define as it comes under different forms.
- By default human beings display an intelligent behavior.
- What about other forms of life: animals, insects, trees.
- Not yet known other superior forms, but very likely to exist.

Artificial Intelligence – definition?

- No formal commonly agreed definition, yet!
- John McCarthy first coined the term (1956):
  - The science and engineering of making intelligent machines
- Marvin Minsky:
  - AI is the science of making machines to do things that would require intelligence if done by men

AI performance

- **optimal**: it is not possible to perform better
- **strong super-human**: performs better than all humans
- **super-human**: performs better than most humans
- **sub-human**: performs worse than most humans

For example, performance at checkers is optimal, performance at chess is super-human and nearing strong super-human, and performance at many everyday tasks performed by humans is sub-human.

AI background

- Philosophy
  - logic, methods of reasoning, mind as physical system, foundations of learning, language, rationality
- Mathematics
  - logic, formal representation
- Statistics
  - probability theory, modeling uncertainty, learning from data
- Psychology
  - perception and motor control, cognitive psychology

AI background

- Economics
  - formal theory of rational decisions, utility
- Linguistics
  - grammar, syntax, semantics
- Neuroscience
  - neurons as information processing unit, synapse as learning mechanism
- Computer science
  - engineering, hardware, algorithms, complexity theory

Abridged history of AI

1943	McCulloch & Pitts: Boolean circuit model of brain
1950	Turing's "Computing Machinery and Intelligence"
1956	Dartmouth meeting: "Artificial Intelligence" adopted
1950s	Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
1957	Rosenblatt's Perceptron for training simple neural networks
1965	Robinson's complete algorithm for logical reasoning
1966 – 73	Surprise! AI discovers computational complexity Neural network research almost disappears
1969 – 79	Early development of knowledge-based systems
1980 – 90	AI becomes an industry, 5th Generation Computer Systems (parallelism and logic programming) Expert systems industry booms
1985 – 93	Expert systems industry busts: "AI Winter"
1986 – 95	Neural networks return to popularity
1995 –	The emergence of distributed AI and artificial agents
2005 – 10	AI disappoints again
2017 –	Deep learning impact

Some killer apps in AI

1991	During the Gulf War, US forces deployed an AI logistics planning and scheduling program that involved up to 50,000 vehicles, cargo, and people. Saved the US more money than spent on all AI research since 1950
1997	Deep Blue (IBM) defeated world chess champion Gerry Kasparov
2011	Watson (IBM) beat human champions on “Jeopardy”
2012	Google car obtains driver’s license in Nevada, US. By 2014, the cars have driven for 1.1 million km without accident
2017	DeepMind’s AlphaGo and AlphaGo Zero
	Elon Musk’s OpenAI Destroys Champion Gamer DOTA2! https://www.youtube.com/watch?v=XbDmxEOj9OY

AI - Russel & Norvig

Views of AI fall into four categories:

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

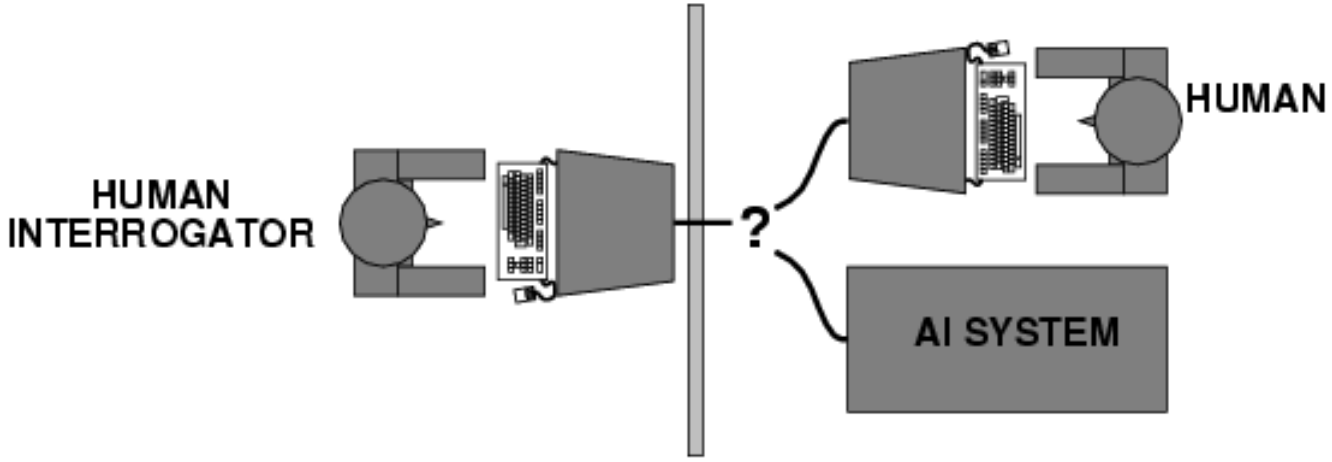
The textbook advocates "acting rationally"

Thinking humanly: Cognitive Science

- 1960s “cognitive revolution”: information-processing psychology replaced prevailing orthodoxy of behaviorism
- Tries to form computational theories of internal activities of the brain.
- How to validate? Requires
  - Predicting and testing behavior of human subjects (top-down), or
  - Direct identification from neurological data (bottom-up)
- Both approaches (roughly, Cognitive Science and Cognitive Neuroscience) are now distinct from AI
- Both share with AI the characteristics that the available theories do not explain (or engender) anything resembling human-level general intelligence

Acting Humanly: Turing Test

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

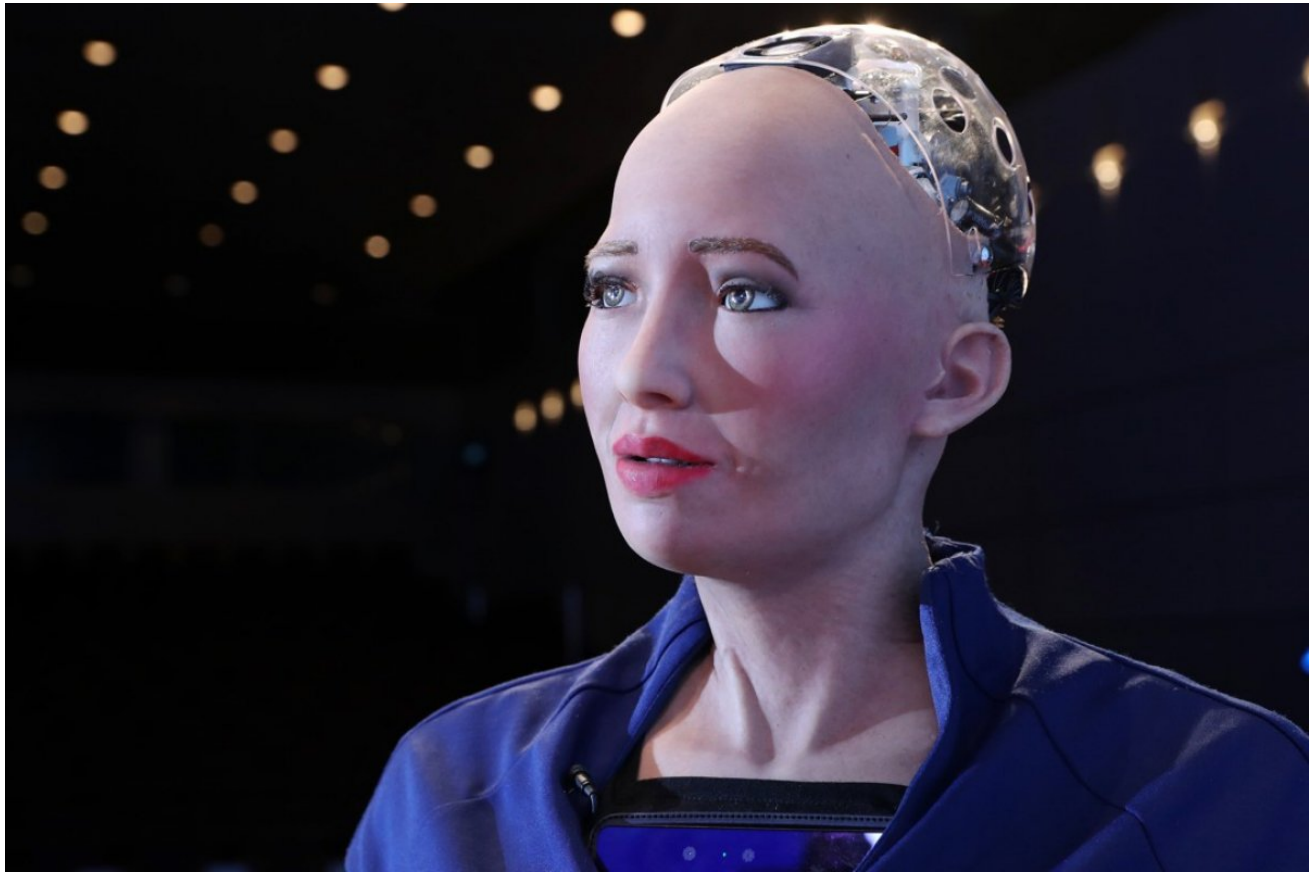


Acting Humanly - Eliza

- One of the most famous early AI programs: Eliza, the computer psychotherapist, created by Joseph Weizenbaum in 1966 at MIT.
- Eliza functions by "twisting the statements of her ’patients’ back at them in the classic manner of a non-directive psychotherapist."
- The fact that it understands the subject’s statements is an illusion. Surprisingly, many users were taking its performance quite seriously.

Acting humanly: what about Sophia?

https://www.youtube.com/watch?v=suRuQbDXcrc



Thinking rationally: Laws of Thought

- Aristotle: what are correct arguments/thought processes?
- Formalize “correct” reasoning using a mathematical model
- Several Greek schools developed various forms of logic: notation and rules of derivation for thoughts;
- Direct line through mathematics and philosophy to modern AI

Thinking rationally: Laws of Thought

Problems:

1. Not all intelligent behavior is mediated by logical deliberation
2. What is the purpose of thinking? What thoughts should I have out of all the thoughts (logical or otherwise) that I could have?
3. Formalizing (informal) common sense knowledge is difficult
4. General deductive inference is computationally intractable

Acting rationally: Rational Agents approach

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

Two main AI paradigms

- Good old fashioned AI (GOFAI)
- Situated embodied AI (SEAI)

GOFAI

- Most of the successes of AI in the 1970s and 1980s were due to research based on Newell's and Simon's Physical Symbol System Hypothesis:

  A physical symbol system has the necessary and sufficient means for general intelligent action.
- Newell and Simon viewed intelligence as symbol manipulation, and hypothesized that it didn't make difference what physical medium – brain, paper, or computer – was used to do the symbol manipulation
- Hence a special emphasis on symbolic representations, which can be interpreted as representing situations in the real world, e.g. "a block world"

Critiques of GOFAI

- Searle: a program (or any physical symbol system) could not be said to understand the symbols that it uses; the symbols have no meaning for the machine (Chineese Room Argument against weak AI)
  - Intelligence is substrate dependent.
  - The material humans are made of is fundamental for our intelligence.
  - Thinking is possible only in special machines - living ones made of proteins.

Critiques of GOFAI

- Brooks: our most basic skills of motion, survival, perception, balance etc. do not seem to require high level symbols at all; the use of high level symbols was more complicated and less successful
- Harnad: the symbol grounding problem: an agent does not perceive symbols, instead the brain converts sensory inputs into higher level abstractions, e.g. symbols

Emergence of SEAI

- The GOFAI approaches turned out to be brittle and very little robust when deployed on real-world problems
- Trying to define a model of the world turned out to be quite hard - this led to Brooks' statement that "the world is its own best model"
- Situated and embodied AI focuses on having a body (i.e. motor skills) in a physical environment.

Approaches to AI

- Cybernetics and brain simulation
- Traditional symbolic AI
- Sub-symbolic AI
- Intelligent agent paradigm
- Integrating the approaches

Problems of AI

- Deduction, reasoning, problem solving
- Knowledge representation
- Planning
- Learning
- Natural language processing
- Motion and manipulation
- Perception
- Social intelligence
- Creativity
- General intelligence

Tools of AI research

- Search and optimization
- Logic
- Probabilistic methods for uncertain reasoning
- Classifiers and statistical learning methods
- Neural networks
- Control theory
- Specialized languages

Ethics, movies

Ethics of artificial intelligence: if a machine can be created that has intelligence, could it also *feel*? If it can feel, does it have the same rights as a human being?

Movies: A.I. (ethics, savior), a servant (R2D2 in *Star Wars*), a comrade (Lt. Commander Data in *Star Trek*), an extension to human abilities (*Ghost in the Shell*), a conqueror (*The Matrix*), a dictator (*With Folded Hands*), an exterminator (*Terminator*, *Battlestar Galactica*) and a race (Asurans in *Stargate Atlantis*)

Future?

Ray Kurzweil has used Moore's law to calculate that computers will have the same processing power as human brains by the year 2029, and that by 2045 artificial intelligence will reach a point where it is able to improve *itself* at a rate that far exceeds anything conceivable in the past, a scenario that science fiction writer Vernor Vinge named the **technological singularity** (1993)

Edward Fredkin argues that artificial intelligence is the next stage in evolution.

Cyborgs = merging humans and machines.