UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Artificial Intelligence

Artificial Intelligence Overview & Course Organization

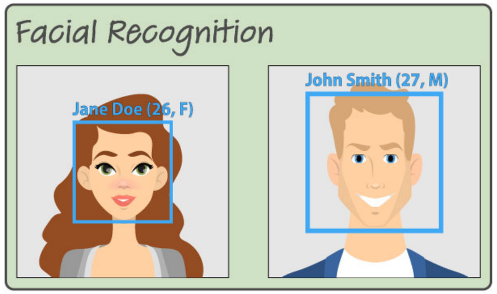
LESSON 1

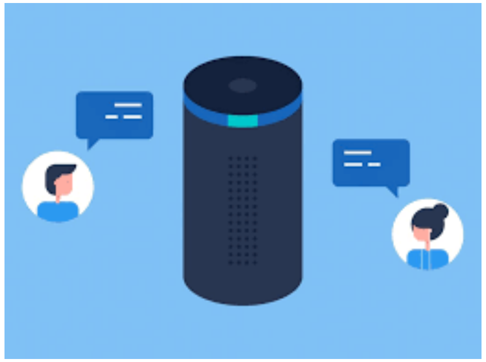
prof. Antonino Staiano

M.Sc. In ''Machine Learning e Big Data'' - University Parthenope of Naples

Artificial Intelligence

- Ideas, techniques, and algorithms at the foundation of AI





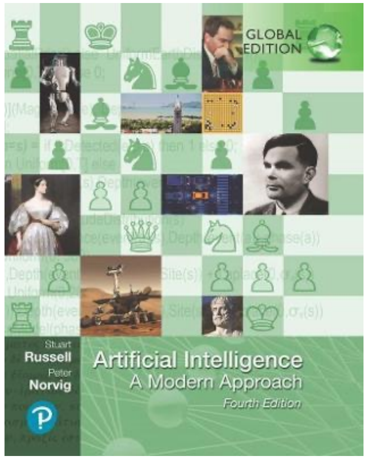
The course

Course logistics

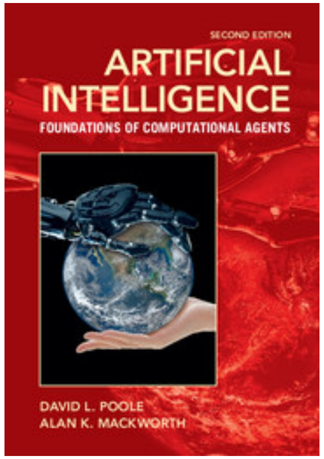
- Instructor: Antonino Staiano
- Time & location
  - Tuesday
    - 14:00 – 16:00
    - Room: Lab 2 (2nd floor, south side)
  - Friday
    - 11:00 – 13:00
    - Room: Lab 1 (2nd floor, south side)
- Office hours
  - Friday 14:00 – 15:00

Learning stuff

- Reference textbooks
  - S. Russell, P. Norvig, Artificial Intelligence – A Modern Approach, 4° ed, Pearson



  - D.L. Poole, A.K. Mackworth, Artificial Intelligence – Foundations of Rationale Agents, 2° ed, Cambridge University Press
    - Freely on line-version available at https://artint.info/2e/html/ArtInt2e.html



- Lecture slides
  - E-learning platform
- Grading
  - Oral + Project/Seminar

What this course is

- Introductory/survey course
  - Introduces many of the core activities in AI and discuss why they are challenging
  - Presents symbolic approaches and standard methods (algorithms) to AI
  - Provides sufficient background to be able to read current research papers on AI

Artificial Intelligence

- Search
  - Looking for solutions to some kind of problem
    - Get driving directions from point A to point B
    - Figure out how to play a game, e.g., figuring out what move it ought to make in tic-tac-toe
- Optimization
  - Optimize for some sort of goal when there might be multiple ways to solve a problem, i.e., looking for, potentially, the best way
- Knowledge
  - An AI must be able to know and represent information
  - Draw inferences from that information
- Uncertainty
  - Ability to deal with uncertain information, that is, deal with events certain with a given probability

Today's lecture Outline

- What is AI?
- A brief history
- The state of the art

What is Intelligence?

- Hard to uniquely define Intelligence as it can manifest in such different ways
- A general definition
  - *A set of capabilities that allows humans to learn, think, understand, communicate, be self-conscious, build abstract models of the world, plan, adapt to novel external conditions, etc.*
  - Note that some of these capabilities are also of animals
    - Associative memory, reacting to stimuli, communicating

Artificial Intelligence

- AI was born as a result of the convergence of branched efforts in pursuing two research goals
  - Understanding human intelligence
  - The construction of machines that are able to autonomously carry out complex tasks that are considered to require "intelligence"
- Branched means that different aspects of intelligence have been studied by several disciplines for a long time and with different points of view
  - Philosophy, Logic, psychology, neurophysiology, ...

Defining Intelligence

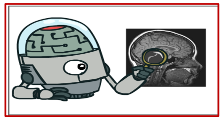
Thinking/Reasoning processes

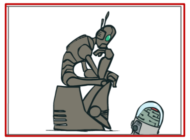
(up)

Systems that think like humans
Systems that act like humans

(down)

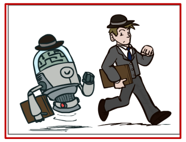
Behavior/Actions

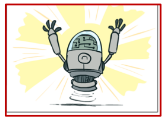




What's AI?

Systems that think rationally
Systems that act rationally





(left)

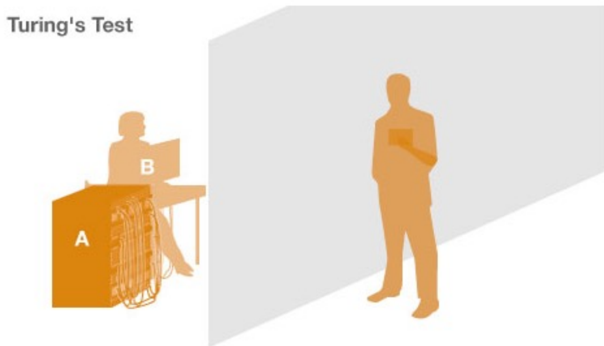
Success in terms of human reference

Success in terms of an ideal concept of intelligence: rationality

(right)

Acting humanly: The Turing test

- *Can machines think?* or *Can machines behave intelligently?*
  - Turing (1950): *Computing Machinery and Intelligence*
  - Operational test for intelligent behavior: The imitation game
    - Predicted that by 2000, a machine might have a 30% chance of fooling a layperson for 5 minutes
    - Anticipated all major arguments against AI in the following 50 years
    - Suggested major components of AI: knowledge, reasoning, language understanding, learning



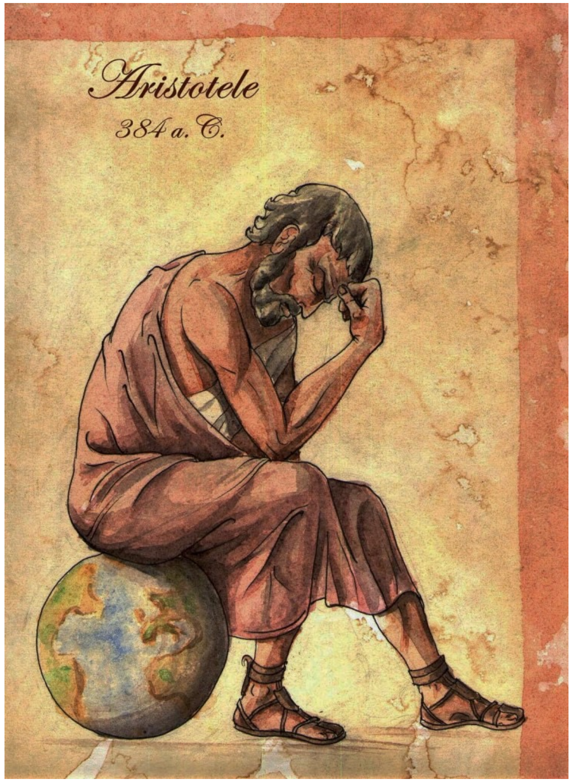
Thinking humanly: Cognitive science

- 1960s *cognitive revolution*: information-processing psychology replaced the prevailing orthodoxy of behaviorism
  - Goal: bring together computer models and experimental techniques from psychology to construct theories of mind
- Requires scientific theories of internal activities of the brain
  - How to validate? Requires
    - Predicting and testing the behavior of human subjects
- Cognitive Science and Cognitive Neuroscience are now distinct from AI
- Both share with AI the following characteristic:
  - the available theories do not explain (or engender) anything resembling human-level general intelligence

Thinking rationally: Laws of thought

- Prescriptive rather than descriptive
  - Attempt to codify the right thinking, that is, the irrefutable reasoning processes
- Several Greek schools developed various forms of logic
  - Notation and rules of derivation for thoughts through a (not necessarily) mechanized process
- Direct line through mathematics and philosophy to modern AI
- Problems:
  - Not all intelligent behavior is mediated by logical reasoning

Aristotle



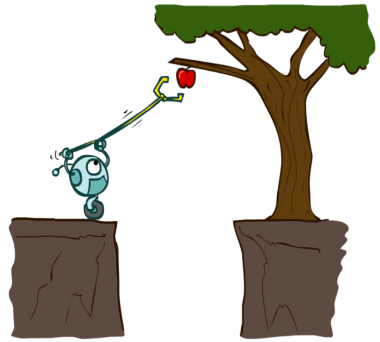
- The first man ever to define a precise set of laws trying to rule the rational component of the human mind
- An informal system of **syllogisms** for the "correct reasoning" that, in theory, allows generating conclusions, given a starting set of premises, mechanically
- Syllogism
  - From *syn*, "set" and logismòs, ''computation" -> *linked reasoning*
- Example
  - All the men (M) are mortal (A)
  - Socrates (B) is a man (M)
  - Socrates is mortal (A)

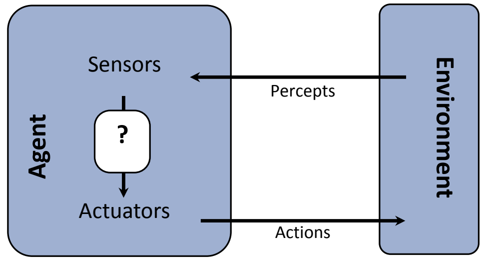
Acting rationally

- Rational behavior: doing the right thing
- The right thing: that which is expected to maximize goal achievement, given the available information
  - It does not necessarily involve thinking, e.g., blinking reflex, but thinking should be in the service of rational action
- Goals are expressed in terms of the utility of outcomes
- Being rational means maximizing your expected utility
- AI has focused on the study and construction of ***agents that do the right thing***

Designing Rational agents

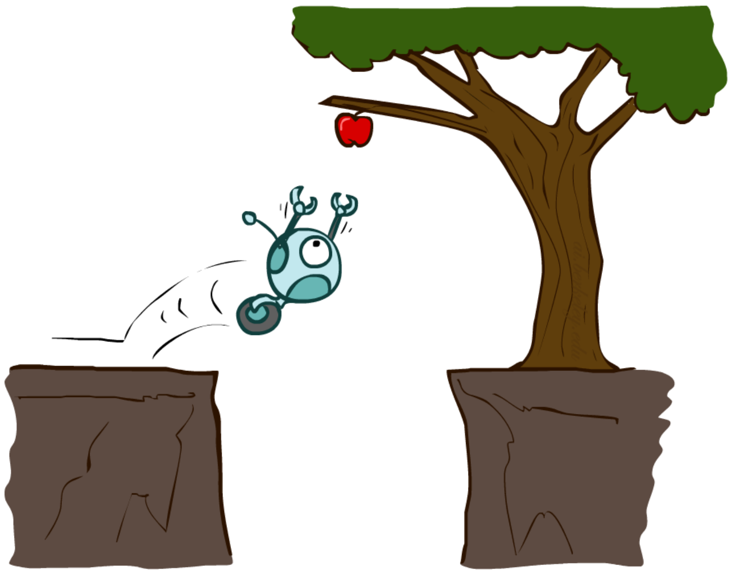
- An agent is an entity that perceives and acts
- A rational agent selects actions that maximize its (expected) utility (internal subjective value of an outcome)
- Characteristics of the percepts, environment and action space dictate techniques for selecting rational actions
- We will learn general AI techniques for a variety of problem types
- Advantages
  - More general the approach based on thinking rationally
    - Correct inference is just one option for achieving rationality
  - Scientific modeling and development
    - Rationality is mathematically well-defined and general





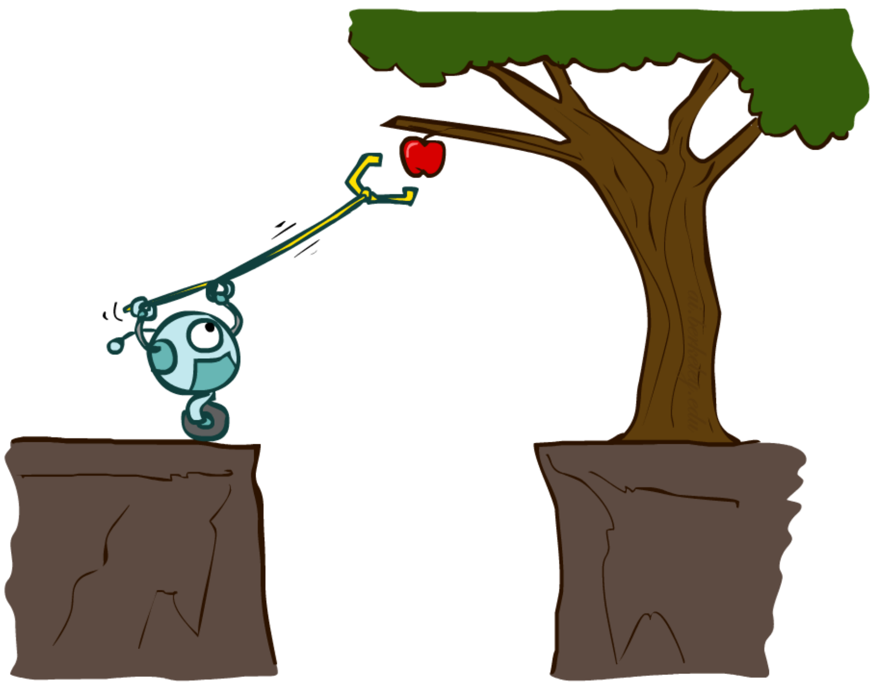
Agents that plan: Reflex agents

- Reflex agents
  - Choose action based on current percept (and maybe memory)
  - May have a memory or a model of the world's current state
  - Do not consider the future consequences of their actions
  - Consider how the world is



Agents that plan: Planning agents

- Planning agents
  - Ask “what if”
  - Decisions based on (hypothesized) consequences of actions
  - Must have a model of how the world evolves in response to actions
  - Must formulate a goal (test)
  - Consider how the world would be

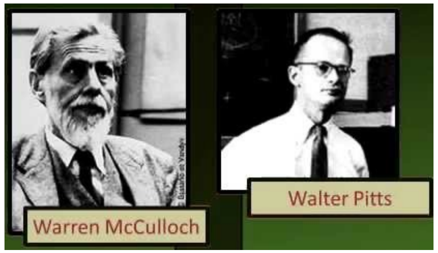


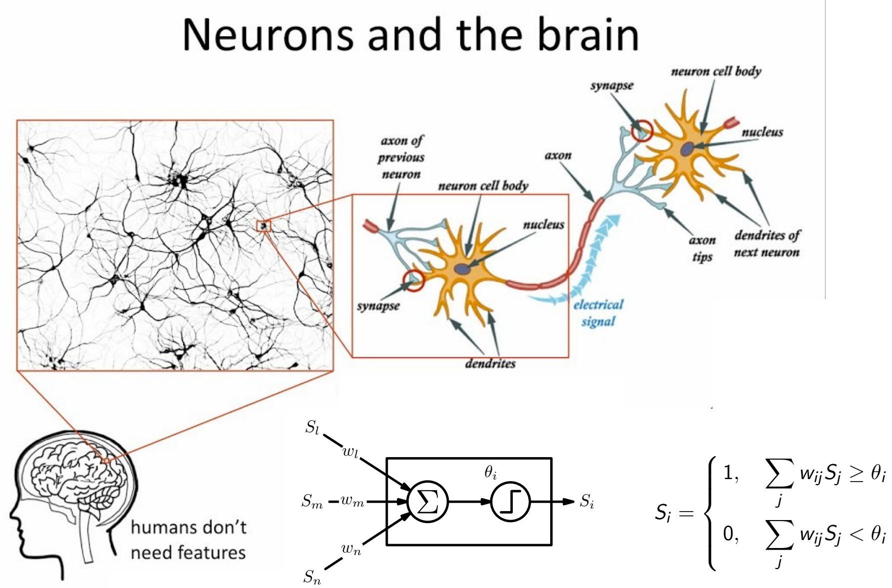
Fundamentals of AI

- Several disciplines involved in AI, each contributed as of ideas, points of view, and techniques
  - Philosophy:
    - logic, methods of reasoning, mind as a physical system, foundation of learning, language, rationality
  - Mathematics:
    - Formal representation and proof algorithms, computation, (un) decidability, (in) tractability, probability
  - Psychology:
    - adaption, phenomena of perception and motor control, experimental techniques (psychophysics)
  - Economy:
    - Formal theory of rational decisions, game theory
  - Linguistics:
    - Knowledge representation, grammar
  - Neuroscience:
    - Plastic physical substrate for mental activity
  - Control theory:
    - Homeostatic systems, stability, simple optimal agent design

A short (and incomplete) chronicle of AI

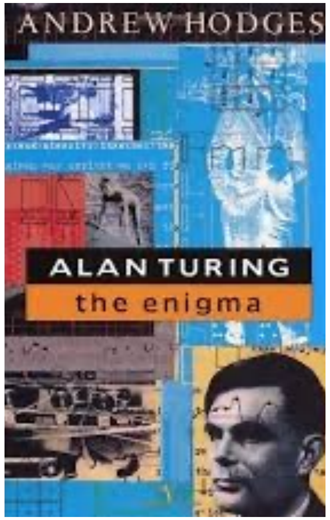
- 1940-1950: Early days
  - 1943: Warren McCulloch e Walter Pitts: an artificial Boolean neuron model for computation
    - First steps towards a connectionist computation and learning (Hebbian learning, 1949)





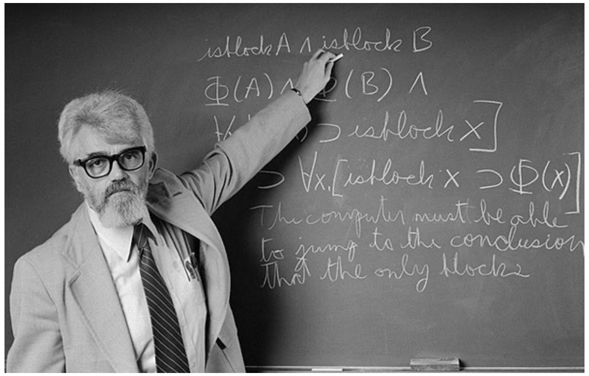
A short (and incomplete) chronicle of AI

- 1950: Alan Turing's "Computing Machinery and Intelligence" Mind, Vol. LIX, No. 263, 433–460, 1950
  - A first comprehensive vision of AI
  - Operational definition of intelligence: the Turing Test
  - Are electronic computers the right tool for building "intelligent" machines?



A short chronicle of AI

- AI as a discipline (1956)
  - Dartmouth Workshop where a group of very brilliant minds met together for studying intelligence and its development in machines
    - The next 20 years were dominated by these *giants*
  - 1952-1969: Excitement: ''Look, Ma, no hands!"
    - Allen Newell and Herbert Simon: The logic theorist (the first non-numerical thinking program for theorem proving)
    - Newell & Simon General Problem Solver
      - Resembling the human problem-solving capacity
    - Arthur Samuel (1952-) checkers program
    - John McCarthy (1958-) :
      - Lisp (the second oldest high-level programming language)
      - Advice Tracker: a first complete AI system
    - Marvin Minsky (1958 -)
      - Microworlds
      - the society of mind
    - 1965: Robinson – complete algorithm for logical reasoning (first-order logic theorem proving)



J. McCarthy, M. Minsky, A. Newell, H. Simon, C. Shannon, O. Selfridge, R. Solomonoff, and others

A short chronicle of AI: 1950s and 1960s

- Goals: Identifying specific tasks that require intelligence, and figuring out how to get machines to do them
- Great interest in mimicking high-level human thought and mental abilities
  - Reasoning
  - Understanding natural language
  - Understanding images
- Some investigations also on low-level abilities
  - Recognizing speech sounds
  - Distinguishing objects in images
  - Reading cursive script
- Main problem: how do humans do that?

A short chronicle of AI: 1950s and 1960s

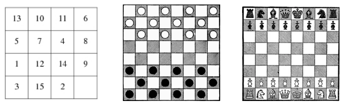
- Starting point
  - Games and toy problems (easy to formalize and investigate), and some real-world ones
    - Game playing: 15-puzzle, checkers, chess, etc.
    - Theorem proving
    - Natural language processing (NLP)
    - Recognizing objects in images

A short chronicle of AI: 1950s and 1960s

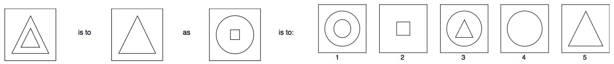
- Fundamental viewpoint
  - the essence of intelligence is deemed to be **symbol processing**
- Early AI research focused therefore on a symbolic approach, aimed at simulating **high-level** manifestations of human intelligence
- Main tools:
  - heuristic search
  - syntax analysis/generation
  - symbolic knowledge representation (symbols, lists, graphs)
  - symbolic knowledge processing: new programming languages (LISP, etc.)

Heuristic search methods

- Symbol processing approach applied to problems like
  - Game playing: 15puzzle, checkers, chess (the “Drosophila of AI)



  - Geometric analogy problems



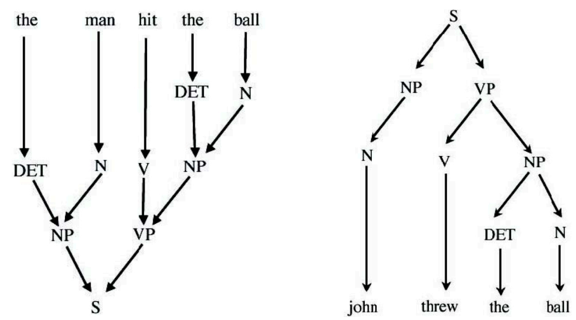
  - Theorem proving
  - Mechanizing problem solving: A. Newell and H. Simon’s General Problem Solver (1959)

Natural Language processing

- Aim: understanding, generating and translating natural language
- A difficult problem, due also to different linguistic levels:
  - Morphology: word parts (e.g., walking = walk + ing)
  - Syntax (grammar): rules that define well-formed sentences
    - *John hit the ball* and *ball the hit John*
  - Semantics: meaning of a sentence
  - Pragmatics: context and background knowledge
    - *John went to the bank*
    - *John threw the ball to the window and broke it*
    - *John threw the glass to the wall and broke it*

Natural Language Processing

- **Syntactic** level (**symbol processing**)
  - Seminal work: N. Chomsky, *Syntactic Structures*, 1957.
- Grammar definition: syntax rules for analyzing/generating sentences
  - **parse tree**



- Applications:
  - original goal: computer interfaces
  - machine translation: early optimism, but it turned out to be a very difficult task

Non-symbolic approaches

- A secondary (by then) approach was a **non-symbolic** one, aimed at simulating **low-level** manifestations/capabilities of intelligence, like perception (mainly, visual perception)
- This approach gave rise to
  - the pattern recognition discipline, which later emerged as a relevant branch of AI
  - artificial neural networks, that became one of the main AI tools (now re-flourishing as *deep learning*)

Pattern Recognition

- Goal
  - classifying different kinds of **signals** (images, sounds, electronic signals, etc.) into one of several categories
- First problem addressed: **image classification**
  - First application: optical character recognition (OCR)

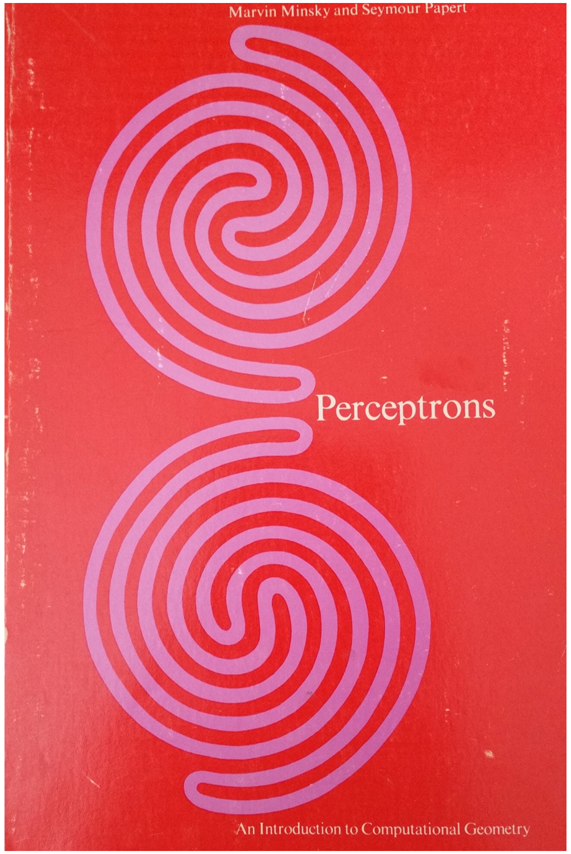
- Main approaches:
  - template matching
  - learning: image pre-processing (noise filtering, line thickening, edge enhancement, ...), feature extraction (e.g., shape), classification "rules" learnt from examples

Artificial Neural Networks

- Non-symbolic (**low-level**), connectionist approach
- The origins:
  - McCulloch and Pitts' mathematical model of neuron
  - the perceptron by F. Rosenblatt (1957): a potential model of human learning, cognition, and memory
    - network of McCulloch-Pitts' neural elements
    - learning algorithm for adjusting connection weights from examples
- First application: pattern (image) recognition
  - OCR
  - aerial images

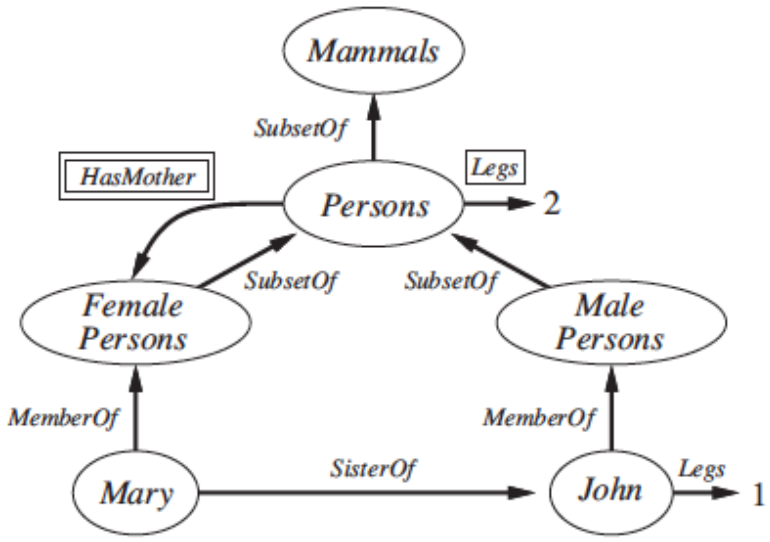
A short chronicle of AI

- The first collapse in AI research (1966 - 1973)
  - Slow advancements as expected
    - Non-realistic predictions
  - Non-scalable systems
    - Combinatorial explosion
  - Limit in fundamentals and representations
    - Minsky and Papert (1969) Perceptrons



A short chronicle of AI

- 1969-1970: AI Revival
  - Knowledge-based approaches
    - DENDRAL (Buchanan et al. 1969)
      - The first successful knowledge-based system
        - Molecular reconstruction from mass-spectrometer data
  - Expert systems
    - MYCIN for blood infection diagnosis (Feigenbaum et al.)
      - Uncertainty in reasoning
  - Knowledge representation research increase
    - Logic, frames, semantic networks, …



Knowledge-based systems

- Knowledge representation and reasoning:
  - development of consulting systems, decision support systems, expert systems
- Main idea
  - solving **domain-specific** problems by embedding **expert** knowledge in the form of IF-THEN rules
- Applications
  - chemistry, medical diagnosis, geology, military; since the 1990s: business

A Brief history of AI

- 1980: AI industry boom
  - DEC R1 (McDermott, 1982)
    - To configure new system orders
  - Japan Fifth Generation Project (1981)
    - Intelligent computer development using prolog
  - Microelectronics, the US reaction
    - Research consortium for chips and human-machine interface
- 1988-1993: Expert systems industry busts: "AI Winter"

AI Winter

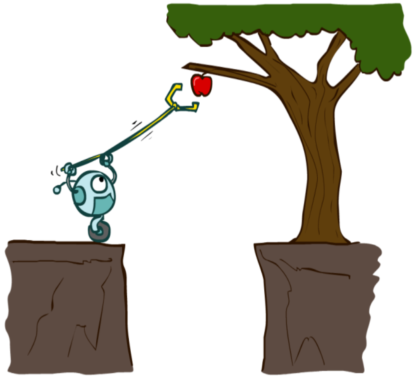
- Until the 1970s AI research is mainly based on the **symbol-processing** conception of human intelligence
  - main approach: mimicking **high-level** human abilities through heuristic search and symbolic processing ("*good, old-fashioned AI*", GOFAI)
- many successful applications through a pragmatic approach in **specific** tasks. . .
- . . . but very limited achievements with respect to early expectations for a **general** AI

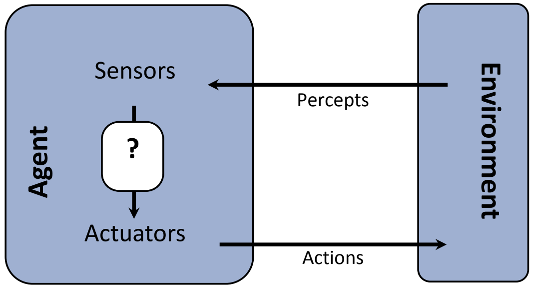
AI Winter

- Real-world tasks turned out to require much more “intelligence” than that achievable by heuristic search and symbolic processing (GOFAI)
- Two main issues emerge:
  - computational complexity: combinatorial explosion
  - human problem-solving relies on a large body of implicit background knowledge (including common sense)
- The non-symbolic, connectionist approach (artificial neural networks) exhibits limitations as well
- Main consequences (“AI winter”):
  - drop in interest in AI
  - scaling back AI’s goals
  - reduction of research funding

A short chronicle of AI

- 1986: The connectionist revival
  - Parallel distributed computation (Rumelhart and McClelland, 1986); backpropagation
- 1990: Statistical approaches
  - The resurgence of probability, focus on uncertainty
  - A general increase in technical depth
  - Agent and learning systems… "AI Spring"?
  - 1996: Kasparov defeats Deep Blue at chess
  - 1997: Deep Blue defeats Kasparov at chess





Towards the AI Spring

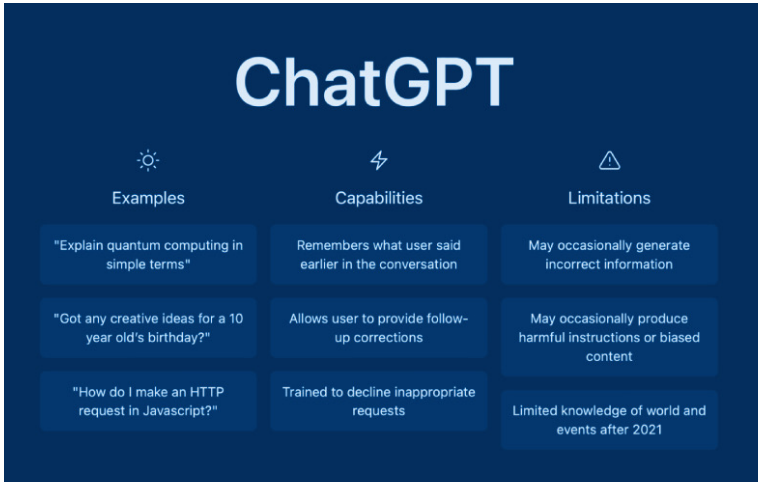
- The AI winter was overcome thanks to new results in several fields, based on solid theoretical foundations from:
  - Mathematics
  - statistics and probability theory
  - control engineering
- This enabled concrete progress in real-world tasks, albeit still **far** from early expectations:
  - knowledge representation and reasoning
  - machine learning
  - computer vision
  - Intelligent Agent architectures

Towards the AI Spring

- The rise of machine learning:
  - availability of large amounts of data in digital form, often manually annotated by users (e.g., user preferences, text translated into different languages by humans)
- Main idea
  - automatically inferring knowledge (patterns, rules, etc.) from data instead of eliciting it from domain experts
  - data analysis methods: data mining, etc.
- theoretical foundations: statistics
- Novel techniques:
  - inductive logic programming, decision trees, the resurgence of ANNs (1986: back-propagation algorithm), support vector machines, ensemble methods, etc.
- many application fields: computer vision, natural language processing, etc.

A Brief History of AI

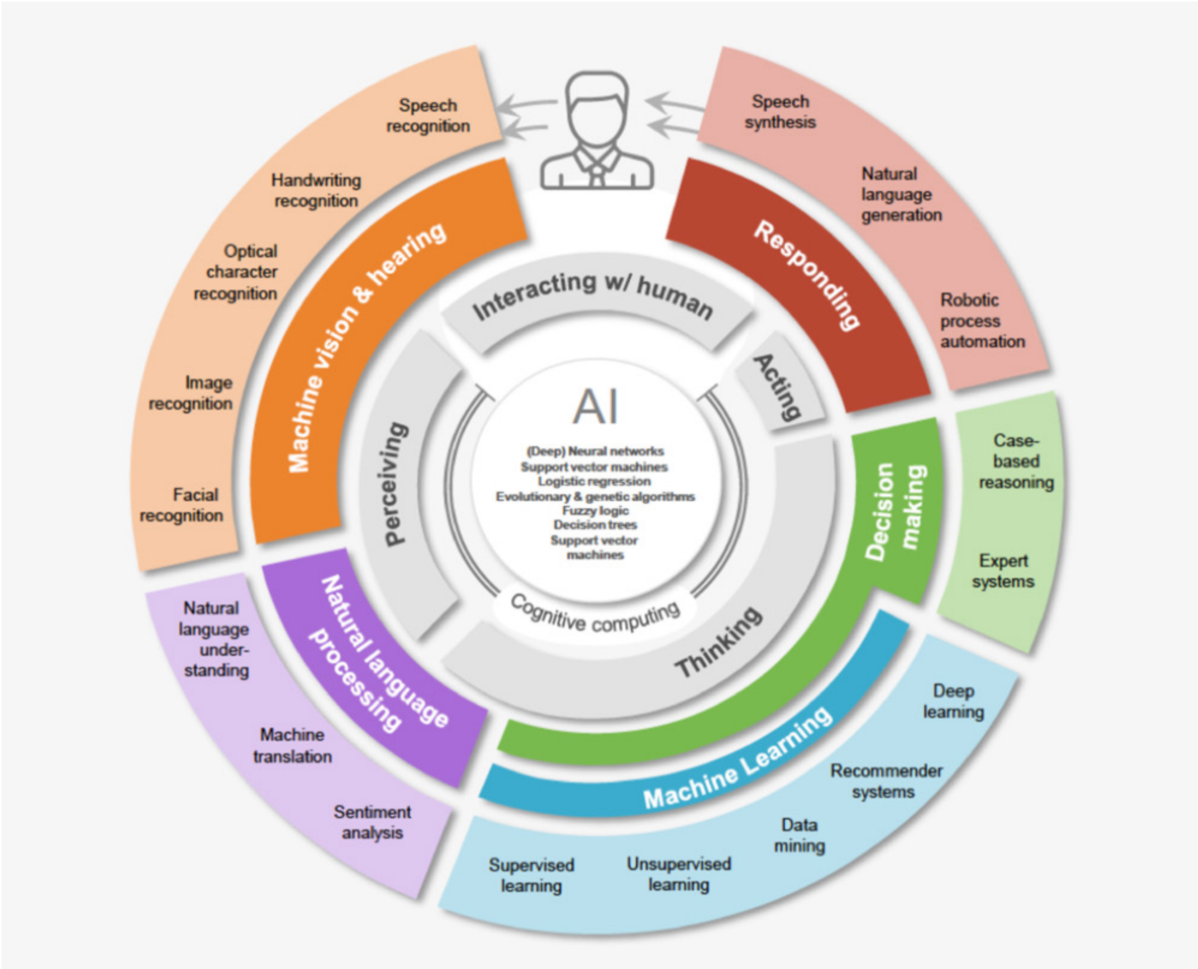
- 2000: Where are we now?
  - Big Data, big computing, neural networks
  - Some re-unification of sub-fields
  - AI used in many industries
  - Chess engines running on ordinary laptops can defeat the world's best chess player
  - 2011: IBM's Watson defeats the Jeopardy champion
  - 2016: Google's AlphaGo beats Lee Sedol at Go
  - 2023: OpenAI's ChatGPT an amazing (with its own limits) conversational Agent



What Can AI Do?

- Play a decent game of table tennis
- Drive safely along a curving mountain road
- Drive safely along Telegraph Avenue
- Buy a week's worth of groceries on the web
- Buy a week's worth of groceries at Berkeley Bowl
- Play a decent game of bridge
- Discover and prove a new mathematical theorem
- Design and execute a research program in molecular biology
- Write an intentionally funny story
- Give competent legal advice in a specialized area of law
- Translate spoken English into spoken Swedish in real time
- Converse successfully with another person for an hour
- Perform a complex surgical operation
- Unload any dishwasher and put everything away

Today AI research fields and applications



Risks and Benefits of AI

- *First solve AI, then use AI to solve everything else*
  - Demis Hassabis, CEO of Google DeepMind
- Benefits
  - Decrease repetitive works
  - Increase production of goods and services
  - Accelerate research (disease cures, climate change and resource shortages solutions)
- Risks
  - Lethal autonomous weapons
  - Surveillance and persuasion
  - Biased decision making
  - Impact on employment
  - Safety-critical applications
  - Cybersecurity threats

Risks and Benefits of AI

- Development of an artificial superintelligence that surpasses human intelligence may pose a significant risk
- Gorilla problem
  - Humans and gorillas evolved from the same species, but humans have more control than other primates
- Thus, we should design AI systems in such a way that they do not end up taking control in the way Turing suggests they might

Ethical issues

- Some ethical issues against AI
  - Even if we could build intelligent machines, should we?
  - Consequences on humans
    - loss of jobs, loss of the sense of being unique, end of humanity, etc
  - Accountability (e.g., driverless cars)
  - …

Philosophical issues

- A long-standing question: Can machines be “intelligent”?
- Two main hypotheses:
  - Weak AI: machines can **emulate** intelligence (act intelligently)
  - Strong AI: machines can **be** intelligent
    (if they act intelligently, they are intelligent, e.g.: Turing test)
- Another long-standing question: Is the human mind a machine?