Introduction to Artificial Intelligence

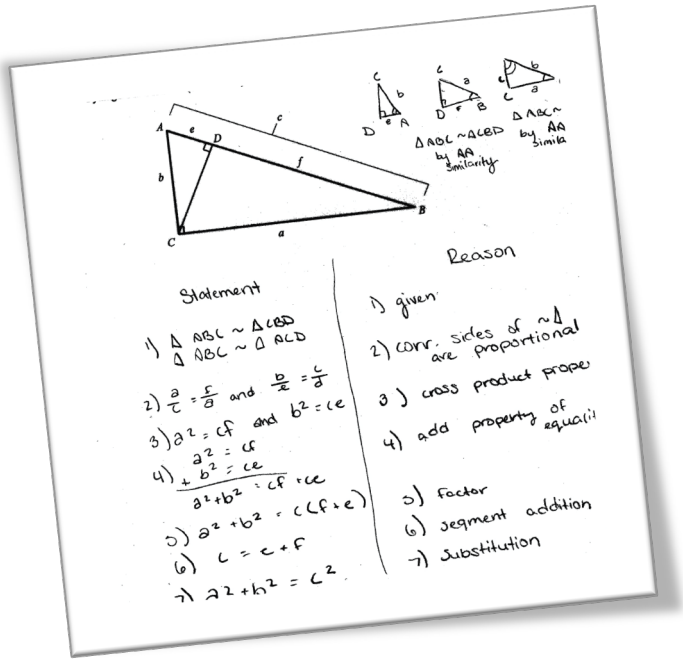
Roman Barták
Department of Theoretical Computer Science and Mathematical Logic

Artificial intelligence is the science of making machines do things that would require intelligence if done by men.

Marvin Minsky, 1967

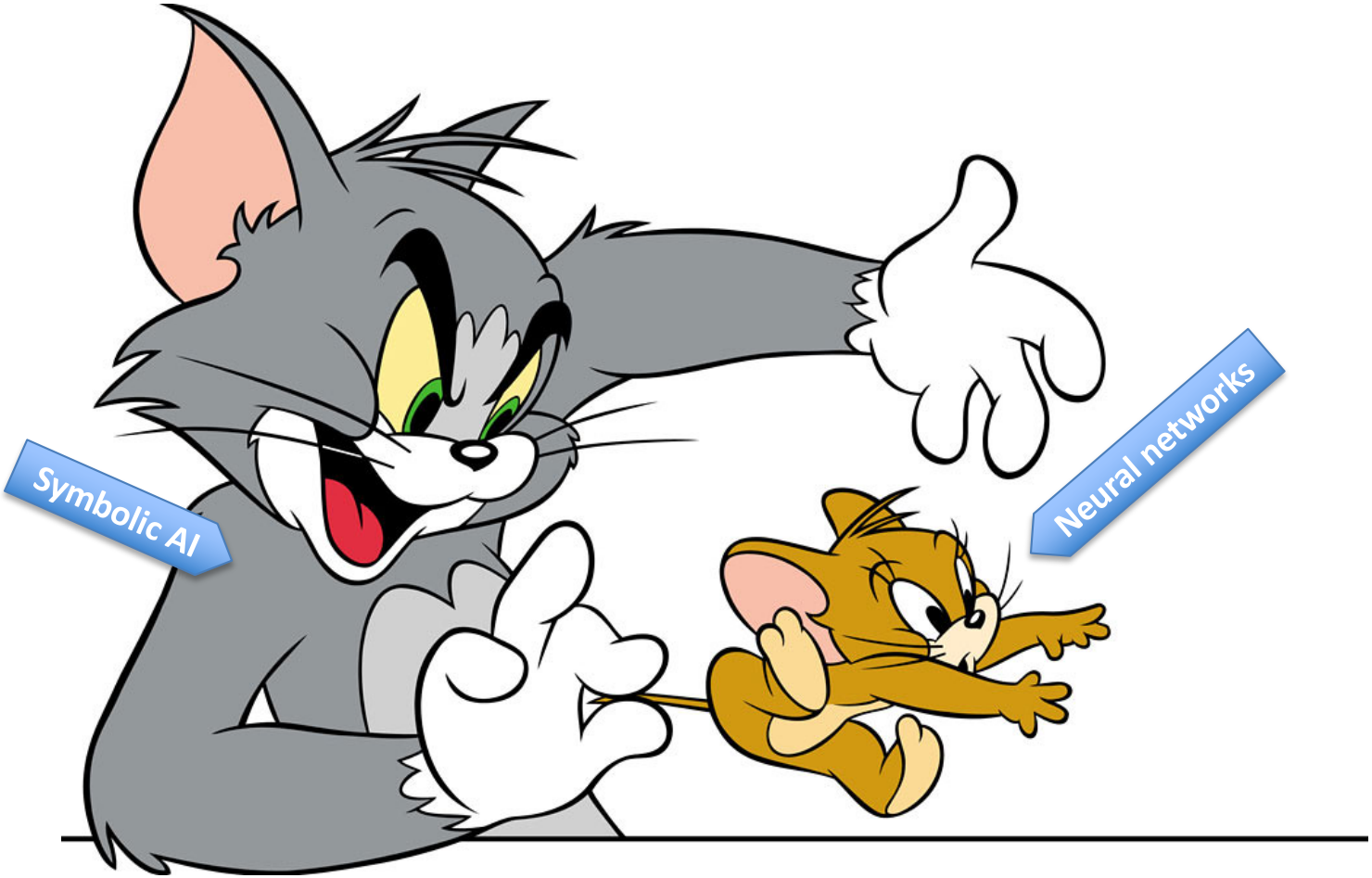
Artificial intelligence attempts to build intelligent entities.

- Playing go, chess, poker,…
- Driving a car, airplane, submarine,…
- Proving theorems, discovering molecules, …
- Writing books, poetry, newspapers,…
- Diagnosing diseases, …
- Talking and listening,…



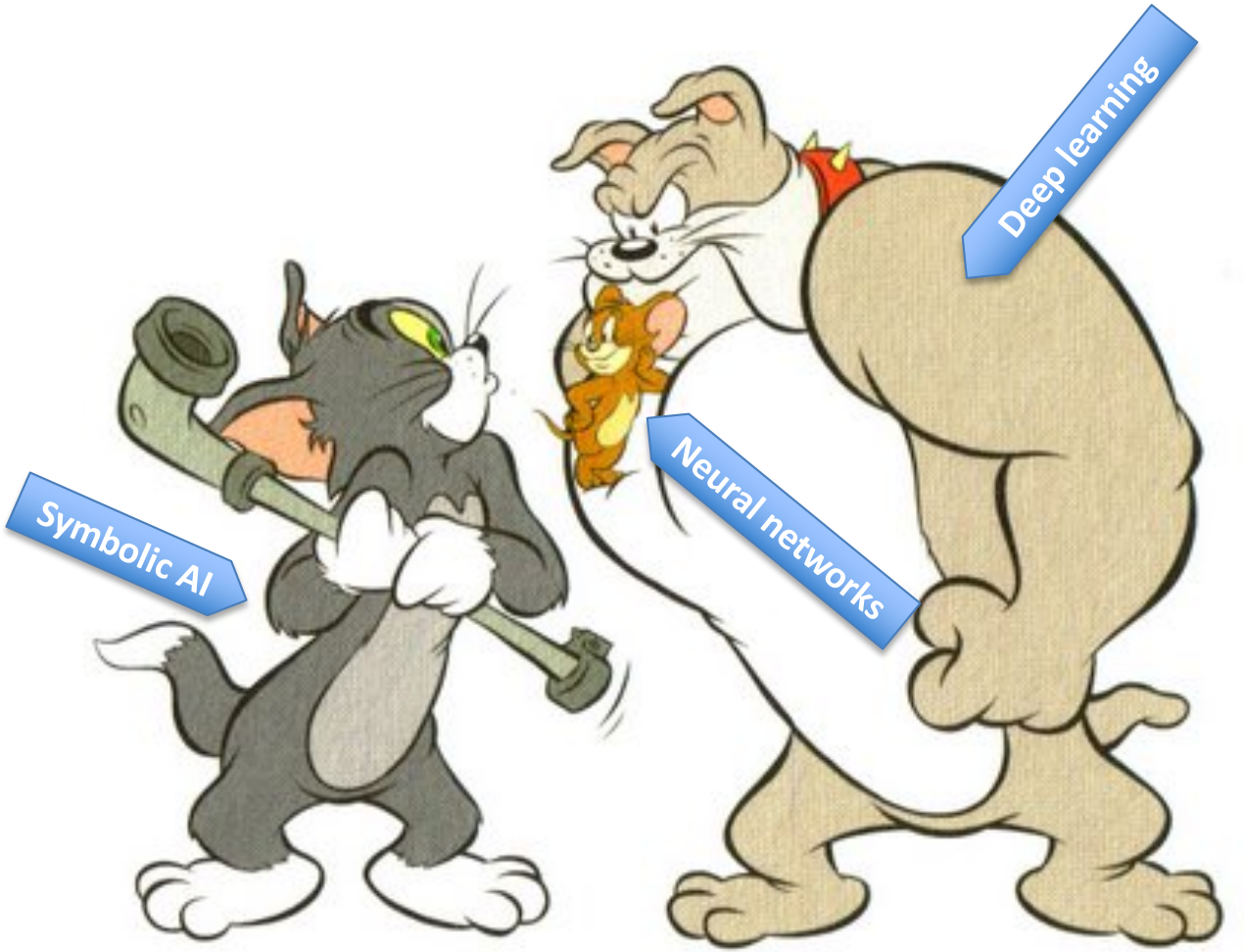
Cartoon History of AI (past)

From Henry Kautz AAAI 2020 talk

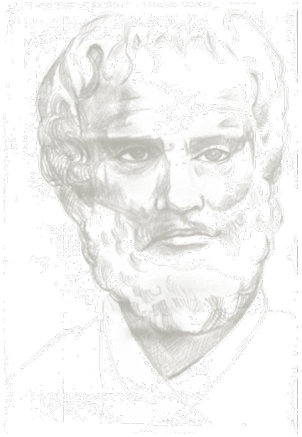


Cartoon History of AI (present)

From Henry Kautz AAAI 2020 talk



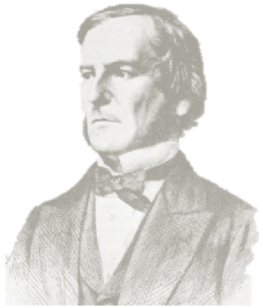
How to formalize rational reasoning?



Aristotle (384-322 B.C.): laws of rational thinking

- Syllogisms: generate conclusions mechanically
  - Patterns for argument structures that always yield correct conclusions when given correct premises
  - Socrates is a man, all men are mortal
    - $\Rightarrow$ Socrates is mortal

René Descartes (1596-1650): advocate power of reasoning in understanding the world



George Boole (1815-1864): formal logic

What if things are not black or white?

- Dealing with uncertain measurements and incomplete theories

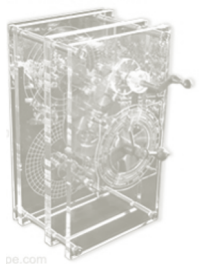
Gerolamo Cardamo (1501-1576): possible outcomes of gambling events

Thomas Bayes (1702-1761): rule for updating probabilities in the light of new evidence

How to do decisions to maximize payoff?

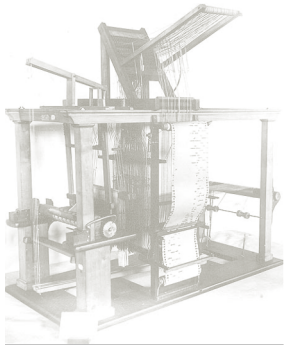
- **Adam Smith** (1723-1790): economics as a science (economies consists of individual agents maximizing their own economic well-being)
- **Economics**: study of how people make choices that lead to preferred outcomes
- **Utility theory**: a formal model for preferred outcome
- **Decision theory**: how to make decisions under uncertainty (without paying attention to other agents)
- **Game theory**: how to make decisions if they significantly affect utility of other agents
  - **John von Neumann, Oskar Morgenstern**. The Theory of Games and Economic Behavior (1944)
- **Operations research**: decisions where payoffs are not immediate but result from a sequence of actions
  - started at World War II in Britain to optimize radar installations

How to design information processing machines?



Automaton (αὐτόματον)

- Antikythera mechanism

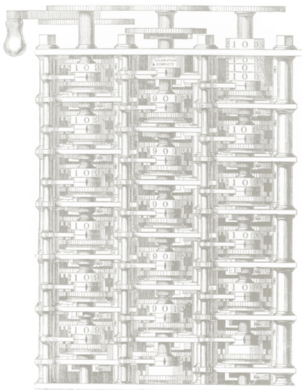


Josef Marie Jacquard (1752-1834)

first programmable machine (punched cards store instructions for the pattern of woven)

Charles Babbage (1792-1871):

- Difference Engine (compute mathematical tables)
- Analytical Engine (universal computation)



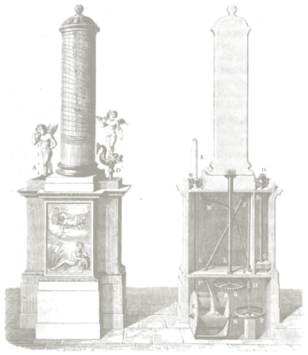
Konrad Zuse:

Z-3, first programmable computer (1941)

Alan Turing:

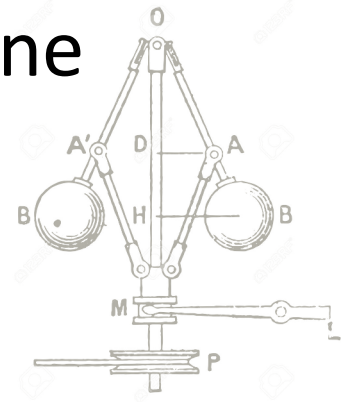
Colossus (1943)

How can artifacts operate under their own control?



Ktesibios of Alexandria (250 B.C.): a water clock with regulator that kept the flow of water running through it constant

James Watt (1736-1819): steam engine governor

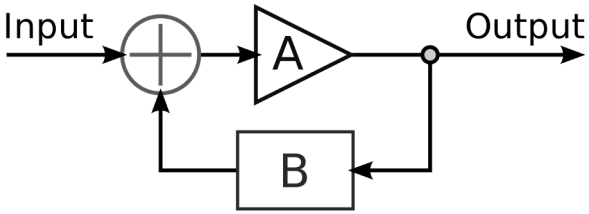


Norbert Wiener (1894-1964): cybernetics (κυβερνητική)

gives possibility of artificially intelligent agents

Control theory: Design of systems that maximize an objective function over time

using calculus and working with continuous variables



How do brains process information?

Aristotle: „Of all the animals, man has the largest brain in proportion to his size“ (comment: this is not really true)

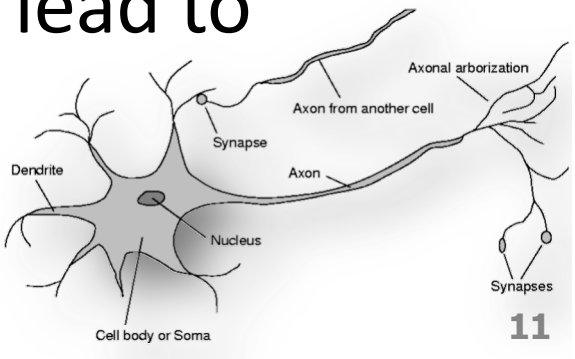
Where does consciousness sit?

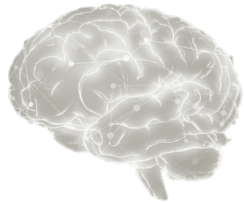
Heart? Spleen?

It was not clear until the middle of the 18th century that the brain was widely recognized as the seat of consciousness.

Paul Broca (1824-1880): study of brain-damaged patients

- Brain consists of interconnected neurons
- How does a collection of simple cells lead to thought, action, and consciousness?





How do humans and animals think and act?

- **Introspection** of own though processes (humans), subjective
- **Behaviorism** (animals): objective measures of the percepts (stimulus) given to an animal and its resulting actions (response)
- **Cognitive psychology**: brain as an information-processing device
- **Kenneth Craik** (1943): **knowledge-based agent**
  1. Stimulus is translated into an internal representation
  2. Representation is manipulated by cognitive processes to derive new internal representation
  3. This representation is translated back into an action

Linguistics: How does language relate to though?

- **Noam Chomsky**: formal theory of languages (explains how to make up completely new sentences, can be programmed)
- Understanding language requires understanding of the subject matter and context, not just an understanding of the structure of sentences

Warren McCulloch, Walter Pitts (1943): model of **artificial neuron** with on/off states, any computable function can be computed by some network of connected neurons, networks could learn

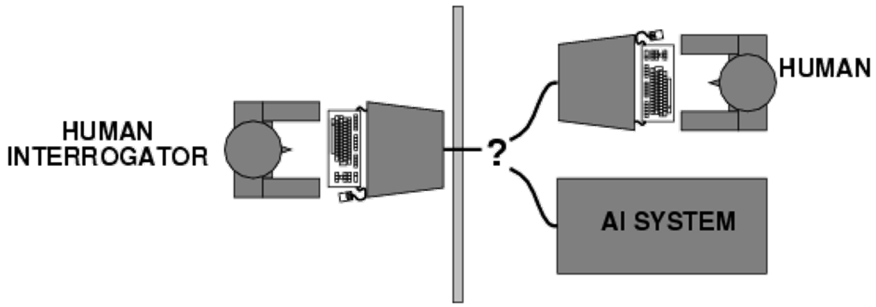
Donald Hebb (1949): updating rule for modifying the connection strengths between neurons (**Hebbian learning**)

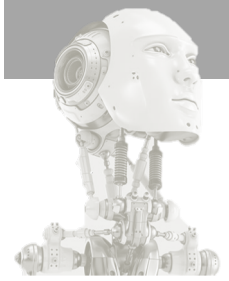
Marvin Minski and **Dean Edmonds** (1950): first neural network computer

- limitations of neural network research

Alan Turing (1950): complete vision of AI (Computing Machinery and Intelligence)

- Turing test





two-month workshop at Dartmouth College

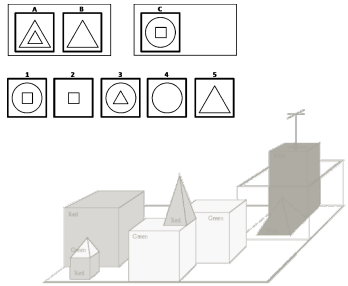
- **John McCarthy**: name **Artificial Intelligence** (computational rationality would be better)
- **Allen Newell**, **Herbert Simon**: **Logic Theorist** (program can prove theorems from Principia Mathematica)

Why AI (and not control theory/cybernetics)?

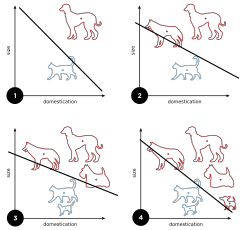
- duplicating human faculties like creativity, self-improvement, and language use
- branch of computer science
- build machines that function autonomously in complex, changing environments

demonstrating one X after another from the list "a machine can never do X"

- **Newell and Simon**: **General Problem Solver** - imitate human problem solving
- **Herbert Gelernter** (1959): **Geometry Theorem Prover**
- **John McCarthy** (1958): AI programming language **Lisp**
- Solving **microworlds**: limited problems that appeared to require intelligence to solve
  - **Tom Evans** (1968): **Analogy** (problems from IQ tests)
  - **Daniel Bobrow** (1967): **Student** (algebra story problems)
  - **Terry Winograd** (1972): **blockworld** (manipulating a set of solid blocks placed on a tabletop)



- **David Waltz** (1975): constraint propagation in computer vision
- **Frank Rosenblatt** (1962): **perceptron** (learning algorithm to adjust connection strengths to match any input data)



John McCarthy referred to this period as the **„Look, Ma, no hands!"** era.

natural language processing

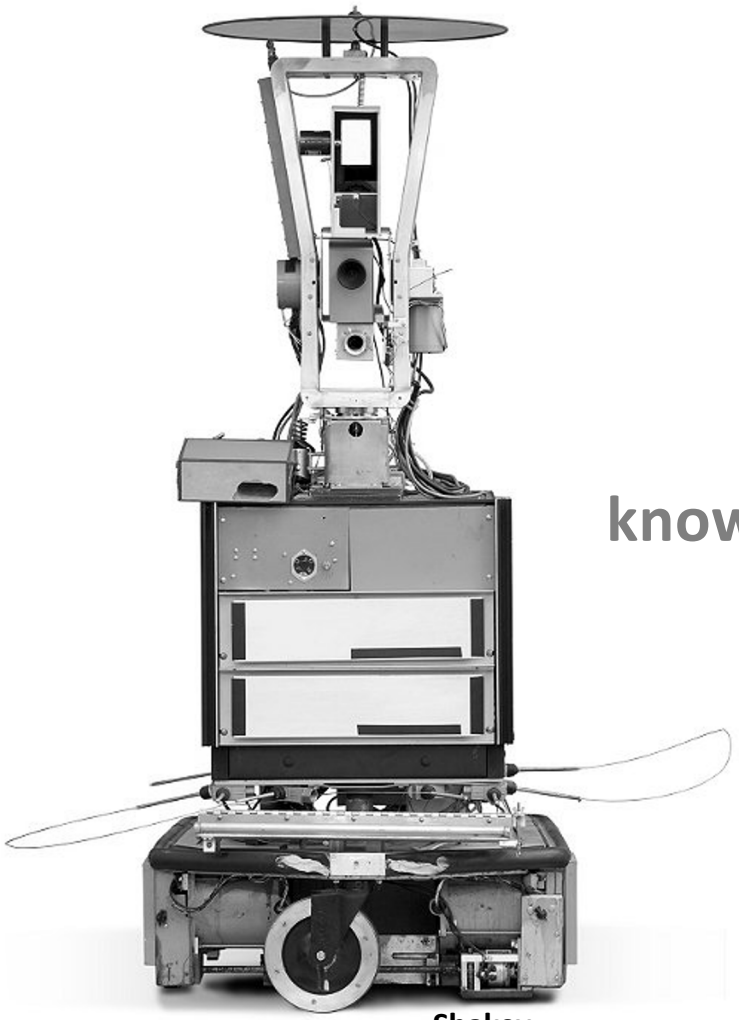
planning

machine learning

STRIPS

knowledge representation

computer vision



Shakey

robotics

algorithm A*

1957, Herbert Simon:
„there are now machines that think, that learn and that create“

1958, H. A. Simon and Allen Newell:
"within ten years a digital computer will be the world's chess champion" and "within ten years a digital computer will discover and prove an important new mathematical theorem."

1965, H. A. Simon:
"machines will be capable, within twenty years, of doing any work a man can do."

1967, Marvin Minsky:
"Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved."

1970, Marvin Minsky:
"In from three to eight years we will have a machine with the general intelligence of an average human being.“

DARPA (US, 1969) stopped funding basic undirected research

- failure of language translation based on syntactic transformation of grammars
- intractability of problems solved (trying out different combinations of steps do not scale up)

Lighthill report (UK, 1973): end support for AI research in all but two universities in UK

- failure of AI outside toy micro-worlds
- AI could never handle combinatorial explosion of real-world domains

Minski & Papert's book **Perceptron** (1969)

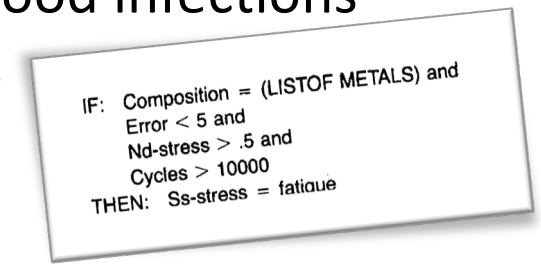
- perceptrons can learn anything they can represent; but they can represent very little (for example XOR)

Knowledge-based systems

- The alternative to „weak“ general methods is to use more powerful, domain-specific knowledge.

Expert systems:

- **Dendral** (Feigenbaum, 1968): inferring molecular structure from information provided by a mass spectrometer
  - Using rules that based on peaks in the spectrum suggested common substructures in the molecule
- **MYCIN** (Shortliffe & Buchannan, 1975): diagnose blood infections
  - deduction rules acquired from experts, textbooks, …
  - certainty factors (to work with uncertainty)
- **R1/XCON** (Dermott, 1975): configure orders of new computer systems (DEC), $40 mil./year saving



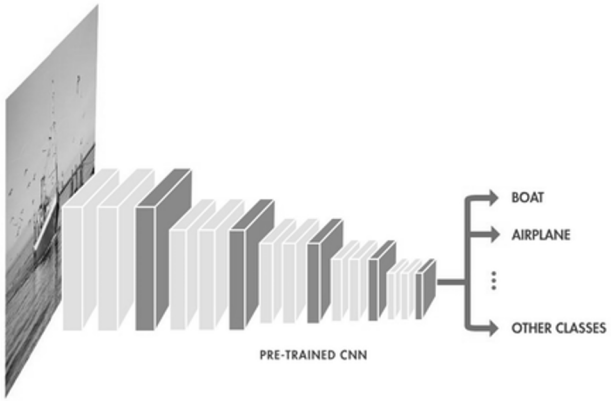
Boom of AI industry (billions of dollars in 1988): all major US companies had AI groups investigating expert systems

Fifth Generation of computers (Japan, 1981): a 10-year plan to build intelligent computers running Prolog

- **fail to deliver** on extravagant promises
- knowledge-engineered experts systems proved **costly to maintain** and being **brittle**
- **failure** of Japan's Fifth-generation AI project
- **collapse of market** for specialized AI workstations (Lisp machines substituted by general-purpose PCs)

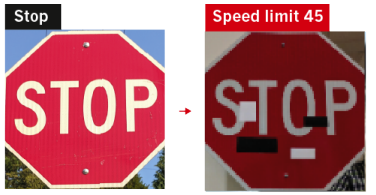
Neural networks strike back: Deep Learning, Big Data, GPU

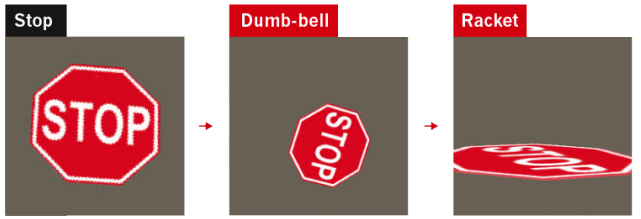
- **ImageNet** Challenge (2012)
  recognizing objects in pictures
- **AlphaGo** (DeepMind, Google)
  - beats human best players in Go
  - learns to play chess and Atari games
- **Watson** (IBM)
  - beats best human players in Jeopardy (open Q&A)
  - tries applications in law and medicine
- **DeepStack** and **Libratus**
  - beat best human players in Poker
- **Self-driving cars**
  - Grand Challenges (CMU, Stanford)
  - Google, Uber, Tesla,…



Superhuman performance (in narrow tasks)

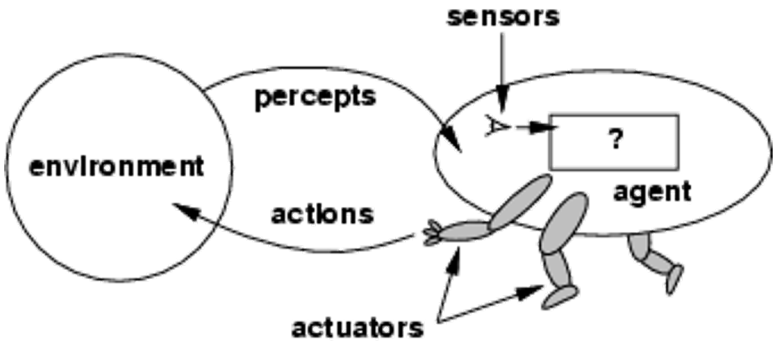
Gartner Hype Cycle



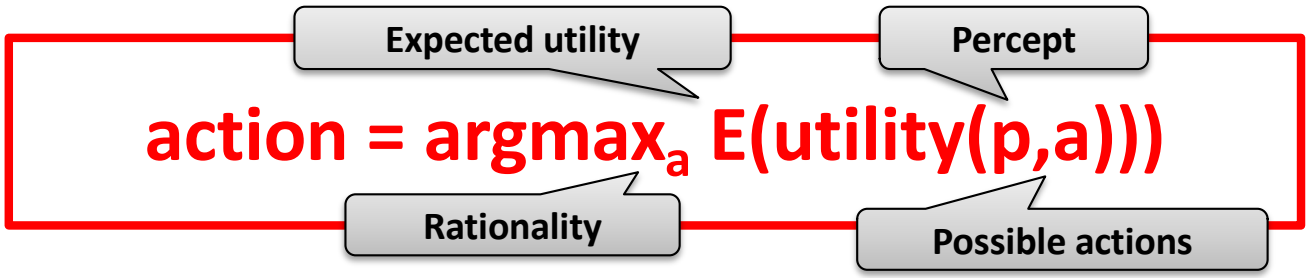


AI is about construction of **rational agents**.

An **agent** is anything that can be viewed as perceiving its **environment** through **sensors** and acting upon that environment through **actuators**.



A rational agent should select an action that is expected to maximize its performance measure.

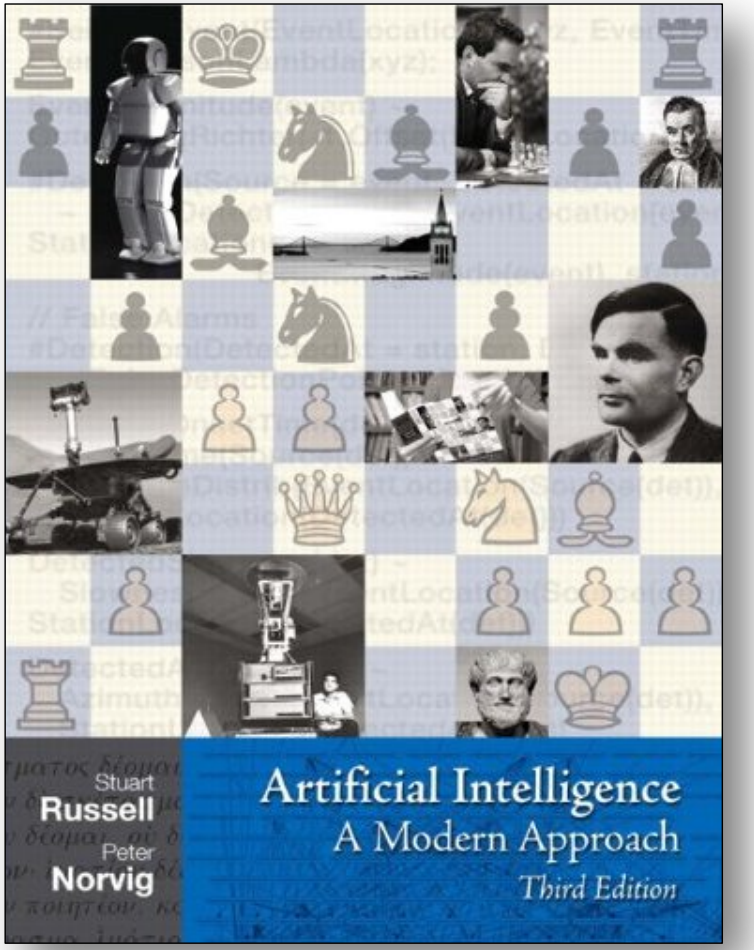


Agents operate in some environment with some properties:

- **Fully observable / partially observable**
  - agent‘s sensors give access to the complete state of environment
- **Deterministic / stochastic**
  - the next state of environment is fully determined by the current state and the action executed
  - strategic = only (other) agents can modify the environment
- **Episodic / sequential**
  - the agent’s experience is divided into atomic episodes (the next episode does not depend on actions taken in previous episodes)
- **Static / dynamic**
  - environment is not changing while an agent is deliberating
  - semidynamic = environment does not change, but the performance score does
- **Discrete / continuous**
  - depends of the state of the environment, the way time is handled, and the percepts and actions of the agent
- **Single agent / multi-agent**
  - Which entities must be viewed as agents?
    If their behavior is best described as maximizing performance measure.
  - competitive vs. co-operative multi-agent environments

- Basic terminology, history, background
- **Problem solving** via **search** (A* and others)
- **Constraint satisfaction**
- **Logical reasoning** (forward and backward chaining, resolution, SAT)
- **Probabilistic reasoning** (Bayesian networks)
- **Knowledge representation** (situation calculus, Markovian models)
- **Automated planning**
- **Markov decision processes**
- **Games** and theory of games
- **Machine learning** (decision trees, regression, reinforcement learning)
- **Philosophical** and **ethical** aspects

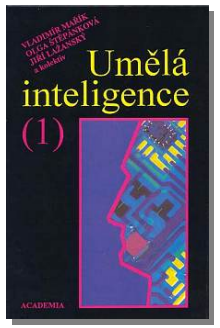
Artificial Intelligence: A Modern Approach

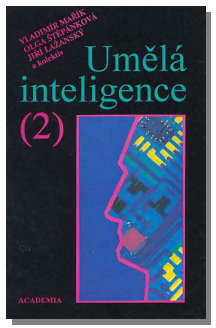


- S. Russell and P. Norvig
- Prentice Hall, 2010 (third ed.)
- http://aima.cs.berkeley.edu/

Umělá inteligence 1-6

- Vladimír Mařík, Olga Štěpánková, Jiří Lažanský a kol.
- Academia





Department of Theoretical Computer Science and Mathematical Logic
bartak@ktiml.mff.cuni.cz