Fundamentals of Artificial Intelligence
Chapter 01: **Introduction to A.I.**

Roberto Sebastiani

DISI, Università di Trento, Italy – roberto.sebastiani@unitn.it
http://disi.unitn.it/rseba/DIDATTICA/fai_2021/

Teaching assistant: Mauro Dragoni – dragoni@fbk.eu
http://www.maurodragoni.com/teaching/fai/

M.S. Course "Artificial Intelligence Systems", academic year 2021-2022

Last update: Saturday 11th September, 2021, 17:29

Outline

Outline

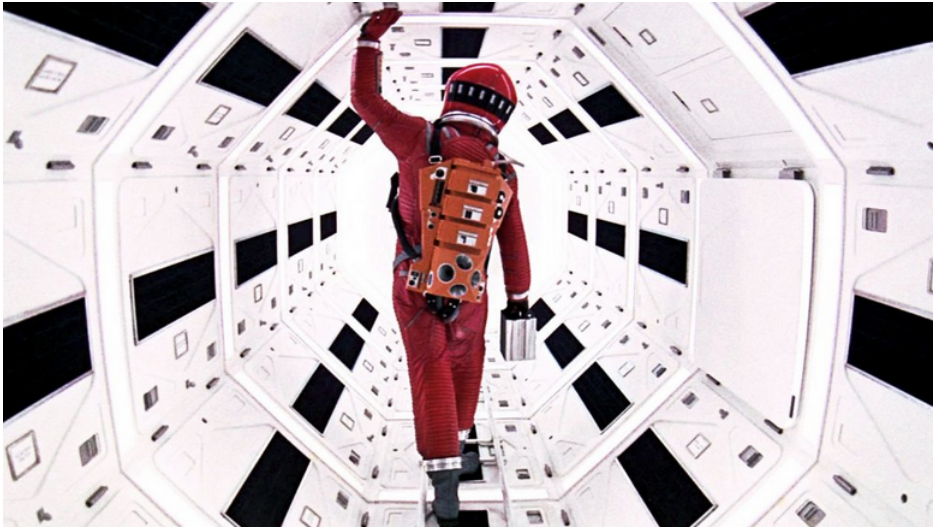
AI in Fiction

There is plenty of AI in fiction ...

AI in Fiction

“Metropolis”, 1927, by Fritz Lang

AI in Fiction



“2001, Space Odyssey”, 1968, by Stanley Kubrick

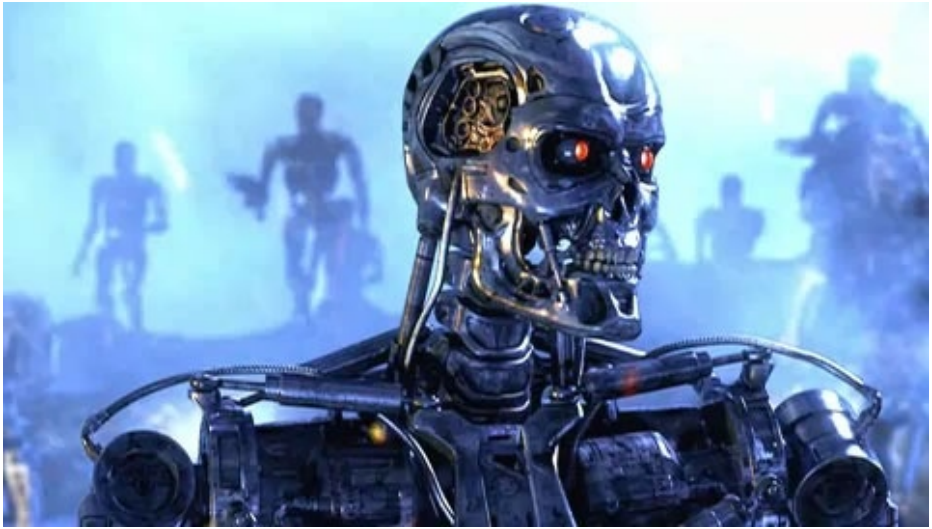
AI in Fiction

“Star Wars”, 1977, by George Lucas

AI in Fiction

“Blade Runner”, 1982, by Ridley Scott

AI in Fiction

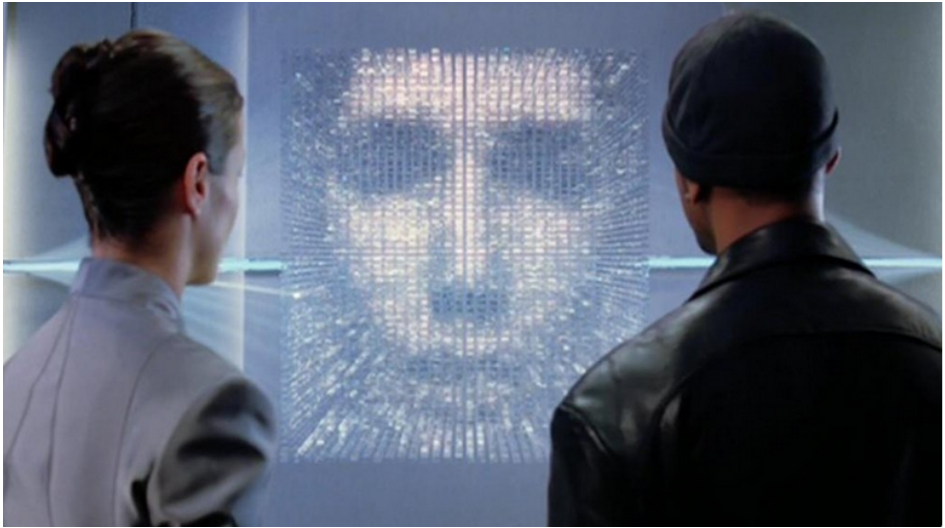


“Terminator”, 1984, by James Cameron

AI in Fiction

“A.I., Artificial Intelligence”, 2001, by Steven Spielberg

AI in Fiction

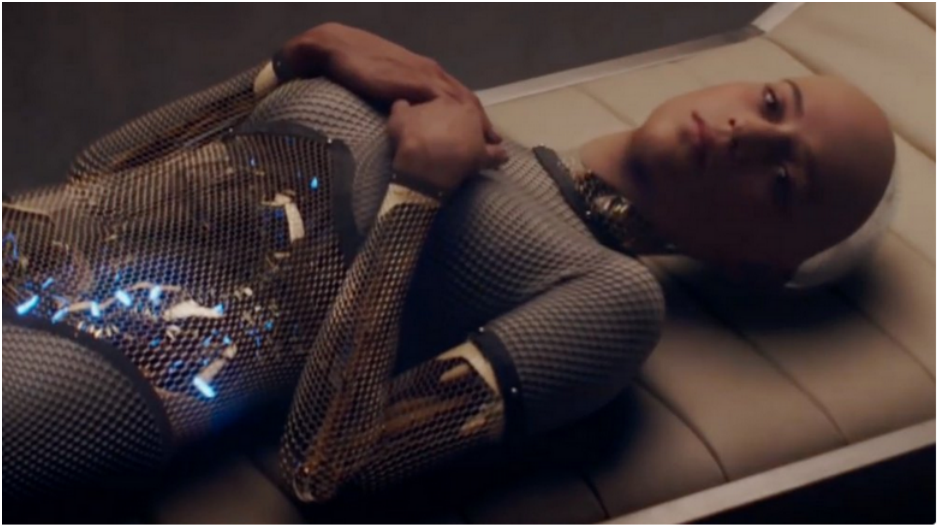


“I, Robot”, 2004, by Alex Proyas

AI in Fiction

“Wall-E”, 2008, by Andrew Stanton

AI in Fiction



“Ex Machina”, 2015, by Alex Garland

AI in Fiction

“Blade Runner, 2049”, 2017, by Denis Villeneuve

AI in Fiction

... and many others ...

(see, e.g.,
https://www.looper.com/198685/the-stunning-evolution-of-ai-in-movies/)

AI in Reality

Many AI fantasies from fiction are becoming reality ...

AI in Reality

... self-driving cars, ...

AI in Reality

... autonomous vacuum cleaners, ...

©iRobot Inc.

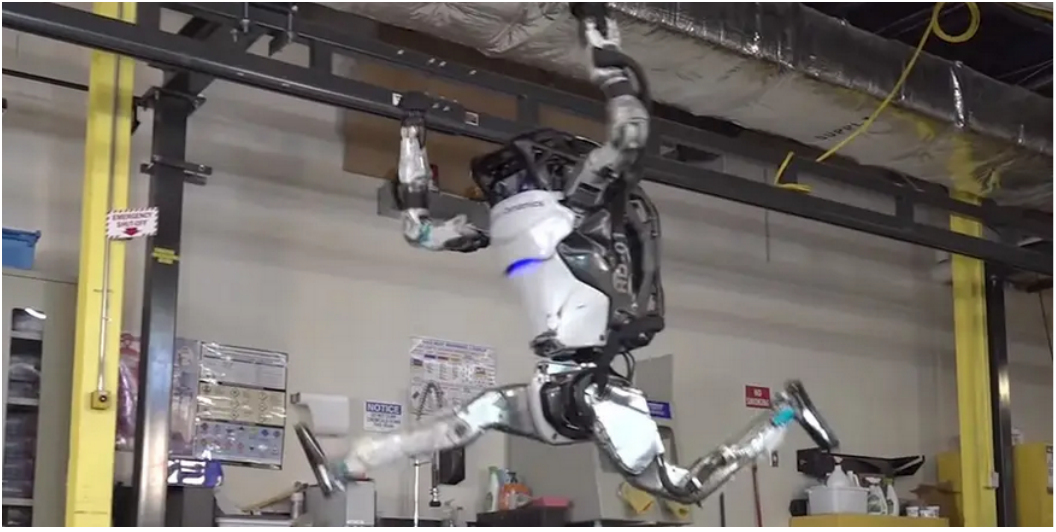
AI in Reality

... soccer-playing robots, ...

©Sony

AI in Reality

.. acrobatic humanoid robots, ...



©Boston Dynamics

AI in Reality

... autonomous trading bots, ...

AI in Reality

..., vocal assistants, ...

AI in Reality

... image/face recognition tools, ...

AI in Reality

... world-champion beating chess players, ...

AI in Reality

... world-champion beating go players, ...

AI in Reality

... AI fighter pilots, ...

AI in Reality

... and many others ...

Outline

Intelligence vs. Artificial Intelligence

Intelligence

For thousands of years, we have tried to understand how we think:

- how can a "handful of matter" perceive, understand, predict, and manipulate a world far larger and more complicated than itself?
- involves many disciplines, including logic, psychology, cognitive science, neuroscience, philosophy, ethics, linguistics, ...

Artificial Intelligence

The field of Artificial Intelligence (AI) goes further still:

- it attempts not just to understand, but also to build intelligent entities
- involves all the above disciplines, but also mathematics, computer science, engineering, economics, control theory & cybernetics, electronics, ...

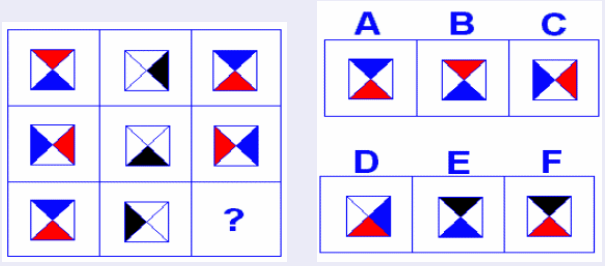
What is Intelligence?

Intelligence (from Wikipedia)

"*Intelligence has been defined in many ways: the capacity for logic, understanding, self-awareness, learning, emotional knowledge, reasoning, planning, creativity, critical thinking, and problem-solving.*
More generally, it can be described as the ability to perceive or infer information, and to retain it as knowledge to be applied towards adaptive behaviors within an environment or context. (...)"

What is Intelligence? [cont.]

Example: simple puzzle



(Courtesy of Michela Milano, UniBO)

- What is the solution of this puzzle?
  $\Longrightarrow$ (I'd say) B: result of column-by-column clockwise rotation
- What have you done for solving it?
  1. read & recognize figures $\Longrightarrow$ perceive information
  2. recognize patterns, problem and candidate solutions
     $\Longrightarrow$ retain knowledge
  3. choose solution $\Longrightarrow$ infer other knowledge

What is Artificial Intelligence?

Different definitions due to different criteria

Historically, four approaches, along two orthogonal dimensions:

- thought processes & reasoning
  vs.
  behavior & action
- Success according to human standards
  vs.
  success according to an ideal concept of intelligence: rationality.
  - human-centered approach: involves observations and hypotheses about human behavior
  - rationalist approach involves a combination of mathematics and engineering.

The four groups have both disparaged and helped each other.

What is Artificial Intelligence? [cont.]

Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Thinking Humanly: The cognitive modeling approach

Problem: How do humans think?

- Idea: develop a theory of the mind

$\Longrightarrow$ express the theory as computer programs

  - e.g. Newell & Simon's General Problem Solver (1961)

- Requires scientific theories of brain activities (cognitive model)
- Inter-disciplinary field: Cognitive Science
  - combines computer models from AI and experimental techniques from psychology
  - construct precise and testable theories of the human mind
- AI and Cognitive Science nowadays distinct
  - A.I: find an algorithm performing well on a task
  - C.S: find a good model of human performance

  although they fertilize each other (e.g. in computer vision)

Acting Humanly: The Turing Test Approach

Problem: When does a system behave intelligently?

The Turing Test

- Alan Turing “Computing Machinery and Intelligence” (1950)
- Operational test of intelligence
  (aka “The Imitation game”):
  - A human, a computer, an interrogator in a different room.
  - The interrogator should classify the human and the machine.
  - Can the computer mislead the interrogator and be classified as a human?
- ”behave intelligently” $\Longleftrightarrow$ ”behave humanly”

Acting Humanly: The Turing Test Approach

Problem: When does a system behave intelligently?

The Turing Test

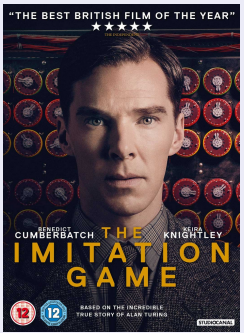
- Alan Turing “Computing Machinery and Intelligence” (1950)
- Operational test of intelligence
  (aka “The Imitation game”):
  - A human, a computer, an interrogator in a different room.
  - The interrogator should classify the human and the machine.
  - Can the computer mislead the interrogator and be classified as a human?
- ”behave intelligently” $\Longleftrightarrow$ ”behave humanly”

Acting Humanly: The Turing Test Approach

Problem: When does a system behave intelligently?

The Turing Test

- Alan Turing “Computing Machinery and Intelligence” (1950)
- Operational test of intelligence
  (aka “The Imitation game”):
  - A human, a computer, an interrogator in a different room.
  - The interrogator should classify the human and the machine.
  - Can the computer mislead the interrogator and be classified as a human?
- ”behave intelligently” $\iff$ ”behave humanly”

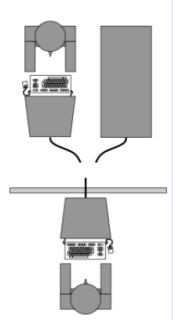


Acting Humanly: The Turing Test Approach

Problem: When does a system behave intelligently?

The Turing Test

- Alan Turing “Computing Machinery and Intelligence” (1950)
- Operational test of intelligence
  (aka “The Imitation game”):
  - A human, a computer, an interrogator in a different room.
  - The interrogator should classify the human and the machine.
  - Can the computer mislead the interrogator and be classified as a human?
- ”behave intelligently” $\Longleftrightarrow$ ”behave humanly”



Acting Humanly: The Turing Test Approach [cont.]

Capabilities for passing the Turing Test

- natural language processing to enable it to communicate successfully in English (or other)
- knowledge representation to store what it knows or hears
- automated reasoning to answer questions and to draw new conclusions
- machine learning to adapt to new circumstances and to detect and extrapolate patterns

For Total Turing test (with physical interaction wrt. interrogator):

- computer vision to perceive objects
- computer speech to communicate orally
- robotics to manipulate objects and move about

- These disciplines compose most of AI
- Turing Test is still relevant in AI (although not fundamental)

Acting Humanly: The Turing Test Approach [cont.]

Some successes with Turing test

- (2014) a chatbot by Eugene Goostman, mimicking the answer of a 13 years old boy, has succeeded the test.
  - chatbots are now frequently available
- vocal assistants are now of common use
  - e.g. Alexa (Amazon), Siri (Apple), Cortana (Microsoft), ...

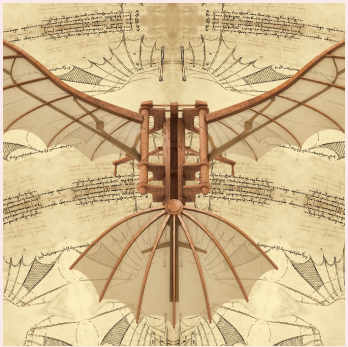
Limitations of Turing Test

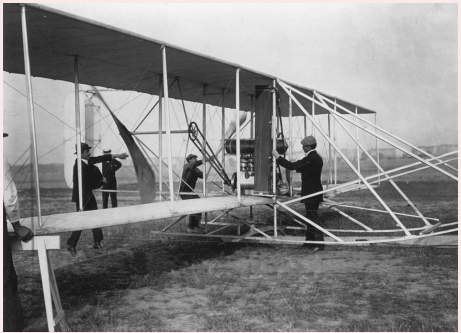
- not reproducible, constructive or amenable to mathematical analysis
- AI researchers devoted little effort to make systems pass the Turing Test
- [Do humans always pass the Turing test? (See e.g. here)]
- Should we really emulate humans to achieve intelligence?
- Shouldn't we study the underlying principles of intelligence instead?

Acting Humanly: The Turing Test Approach [cont.]

Metaphorical Example

Successful flight machines have not been developed by imitating birds, rather by studying engines and aerodynamics.





(see e.g. this video).

Thinking Rationally: The “Laws of Thought” Approach

Problem: Can we capture the laws of thought?

- Aristotle: What are correct argument and thought processes?
  - codify “right thinking” i.e. irrefutable reasoning processes (syllogisms): (e.g. “all men are mortal; Socrates is a man; therefore, Socrates is mortal”)

  $\Longrightarrow$ Logic and Logical inference
- The Logicist tradition in AI hopes to create intelligent systems using logic-based inference systems
  - “algorithm = logic + control”
  - logic programming, automated-deduction systems, ...
  - logics: propositional, first-order, modal & decription, temporal, ...
- Two main limitations:
  - not easy to state informal knowledge into the formal terms of logic
  - problems undecidable or computationally very hard (NP-hard)
- Logical reasoning is currently part of many fields of AI
  - problem solving, knowledge representation & reasoning, planning,
  - does not exhaustively cover AI

Acting Rationally: The Rational-Agent Approach

Problem: Can we make systems “do the right thing”?

Rational Agents

- An agent is an entity that perceives and acts
  - persists over a prolonged time period
- A rational agent acts so as
  - to achieve the best outcome (maximize goal achievement), or
  - to achieve the best expected outcome (under uncertainty)
- Rational agents need all skills needed for the Turing Test!
- Thinking rationally is sometimes part of being a rational agent
  - e.g. planning an action
  - sometimes action without thinking (e.g. reflexes)
- Two advantages over previous approaches:
  - More general than law of thoughts approach (correct inference is just one of several possible mechanisms for achieving rationality)
  - More amenable to scientific development than human-emulation approaches (rationality mathematically well defined & general)

Acting Rationally: The Rational-Agent Approach [c.]

This course concentrates on general principles of rational agents and on the components for constructing them. (Following AIMA book.)

Remark

- achieving perfect rationality is not feasible in complex environments
  - computational demands too high
  - however, good working hypothesis and starting point for analysis

$\Longrightarrow$ dealing with limited rationality

- acting appropriately when not enough time to do all computations

AI Systems Classification

Weak vs. Strong AI

- Weak AI: Is it possible to build systems that act as if they were intelligent?
- Strong AI: Is it possible to build systems that are intelligent?
  (i.e., that have conscious minds, wills and sentiments?)

General AI vs. Narrow AI

- General AI refers to systems able to cope with any generalized task which is asked of it, much like a human.
- Narrow AI refers to systems able to handle one particular task.
  $\Longrightarrow$ AI system displays a certain degree of intelligence only in a particular narrow field to perform highly specialized tasks

AI Systems Classification [cont.]

Symbolic Approach vs. Connectionist Approach

- Top-down, or Symbolic Approach:
  - Symbolic representation of knowledge
  - Logics, ontologies, rule based systems, declarative architecture
  - Human-understandable models
- Bottom up, or Connectionist Approach:
  - Based on Neural networks.
  - Knowledge is not symbolic and it is “encoded” into connections between neurons.
  - Concepts are learned by examples
  - Non understandable by humans

Outline

The Foundations of Artificial Intelligence

Different fields have contributed to AI in the form of ideas, viewpoints and techniques

- Philosophy: Logic, reasoning, mind as a physical system, foundations of learning, language and rationality
- Mathematics: Formal representation and proof, computation, (un)decidability, (in)tractability, probability
- Economics: formal theory of rational decisions, game theory
- Neuroscience: physical substrate for mental activities
- Psychology: adaptation, phenomena of perception and motor control
- Computer Science & Engineering: algorithms, data structures, efficient implementations
- Control Theory & Cybernetics: homeostatic systems, stability, optimal agent design
- Linguistics: knowledge representation, grammar

Brief History of Artificial Intelligence

The Gestation of AI (1943-1955)

- 1943: Warren Mc Culloch and Walter Pitts: a model of artificial Boolean neurons to perform computations
  - First steps toward connectionist computation and learning
  - Marvin Minsky and Dann Edmonds (1951) constructed the first neural network computer
- 1950: Alan Turing: “Computing Machinery and Intelligence”
  - Turing Test
  - First complete vision of AI

Brief History of Artificial Intelligence [cont.]

The Birth of AI (1956) and Era of Great Expectations

- Darmouth Workshop (1956) brought together top minds on automata theory, neural nets and the study of intelligence
  - Allen Newell and Herbert Simon: The Logic Theorist
    - first nonnumerical thinking program used for theorem proving
    - proved theorems from Russel&Whitehead Principia Mathematica
- The era of great expectations (1952-1969)
  - Newell and Simon introduced the General Problem Solver (GPS)
    - could handle a (limited) number of logical puzzles
    - imitation of human problem-solving: strategy to address subgoals
    - Idea: any system (human or machine) exhibiting intelligence must operate by manipulating data structures composed of symbols
  - John McCarthy
    - Invented LISP (and time-sharing)
    - Logic-oriented Advice Taker, decoupling knowledge and reasoning
  - Marvin Minsky
    - addressed microworlds, problems in limited domain that appear to require intelligence to solve (e.g. blocks-world, geometric problems)
  - S. Winograd and J.D. Cowan, et al.: early work on neural networks

Brief History of Artificial Intelligence [cont.]

Collapse in AI research (1966 - 1973)

- Progress was slower than expected.
  - enthusiast predictions turned unrealistic
- Some systems lacked scalability
  - computational intractability due to combinatorial explosion in search
- Fundamental limitations on techniques and representations
  - Minsky&Papert (1969): important limitations to neural networks

Brief History of Artificial Intelligence [cont.]

AI Revival via knowledge-based systems (1969-1970)

- General-purpose $\Longrightarrow$ domain specific systems
  - narrow domains, exploiting domain-specific knowledge
  - E.g. DENDRAL: successful in inferring molecular structure from information by mass-spectrometer (Buchanan et al. 1969)
- Expert systems applied to areas of human expertise
  - e.g., MYCIN: diagnose blood infections (Feigenbaum et al.)
  - based on 450 domain-specific rules from experts & textbooks
  - a calculus for uncertainty
- Several progresses in Natural language processing
  - incorporate domain knowledge in NLP

AI becomes an industry (1980-present)

- commercial expert system R1 at DEC (McDermott, 1982)
  - helped configure orders for computer system (saves: 40M$/year)
- followed a period of national and industry investments in AI
- followed a period of expert systems industry busts ("AI Winter")

Brief History of Artificial Intelligence [cont.]

The return of neural networks (1986-present)

- (re)invented the back-propagation learning algorithm
  - applied to many learning problems in computer science and psychology

$\Longrightarrow$ revival of connectionist models for intelligent systems
(vs. symbolic or logicist approaches)

Brief History of Artificial Intelligence [cont.]

AI adopts the scientific method (1987-present)

- A “gentle revolution” in AI content and methodology
  - build on existing theories than to propose brand-new ones
  - base claims on rigorous theorems or hard experimental evidence rather than on intuition
  - show relevance to real-world applications rather than toy example
- AI has finally come firmly under the scientific method
  - hypotheses must be subjected to rigorous empirical experiment
  - results must be analyzed statistically for their importance

  $\Longrightarrow$ general increase in technical depth
- Resurgence of probability, focus on uncertainty
  - (speech & handwriting recognition): hidden Markov models
  - neural networks benefited from statistics, pattern recognition, and machine learning $\Longrightarrow$ data mining
  - rigorous reasoning with uncertainty: Baynesian networks
  - Similar “gentle revolutions” occurred in robotics, computer vision, and knowledge representation.

Brief History of Artificial Intelligence [cont.]

The emergence of intelligent agents (1995-present)

- renewed interest in the “whole agent” problem:
  “How does an agent act/behave embedded in real environments with continuous sensory inputs?”
- Es: AI in the internet domain “-bots”
  - Decision support systems, robotic agents, natural language
- Need for interaction between sensing and reasoning
  $\Longrightarrow$ reasoning and planning systems must handle uncertainty
- AI forced into much closer contact with other fields
  - e.g. control theory, economics

Brief History of Artificial Intelligence [cont.]

The availability of very large data sets (2001-present)

Big data and massive computing power (e.g. GPUs) have enabled deep networks to be properly trained and to work properly

- Until recently: emphasis on algorithms
- Recent works in AI: emphasis on data
  (for machine learning & deep learning)

$\Longrightarrow$ learning methods rather than hand-coded knowledge engineering used to express the knowledge a system needs

- $\Longrightarrow$ Large amount and variety of AI applications
  (speech and image recognition, spam filtering, robotics, machine translation, autonomous vehicles, game playing, ...)
- many AI applications are now deeply embedded in the infrastructure of every industry

Brief History of Artificial Intelligence [cont.]

The Deep-Learning Tsunami (2015-present)

- “Deep Learning waves have lapped at the shores of computational linguistics for several years now, but 2015 seems like the year when the full force of the tsunami hit the major Natural Language Processing (NLP) conferences.” [C. Manning]
- Previous successes in the fields of image classification and speech...
- Experts in the field (LeCun, Hinton, Bengio) agree on the fact that there will be important developments in text and video understanding, machine translation, question answering ... [Turing award]
- Google masters GO: Deep-learning software defeats human professional for the first time. AlphaGo. Nature 529, 445-446 (28 January 2016). In March 2016, Lee Sedol defeated.

Main AI Research Venues

- Major AI Journals
  - Artificial Intelligence
  - Computational Intelligence
  - Journal of Artificial Intelligence Research
  - IEEE Transactions on Pattern Analysis and Machine Intelligence
  - IEEE Intelligent Systems
  - [area-specific journals]
- Main AI Conferences
  - International Joint Conference on AI (IJCAI)
  - National Conference on AI (AAAI)
  - European Conference on AI (ECAI)
  - [area-specific conferences]
- Main professional societies for AI
  - American Association for Artificial Intelligence (AAAI)
  - ACM Special Interest Group in Artificial Intelligence (SIGART)
  - Society for Artificial Intelligence and Simulation of Behaviour (AISB)

Outline

AI is everywhere ...

- Search engines
- Route planning (e.g. maps, traffic)
- Logistics (e.g. packages, inventory, airlines)
- Medical diagnosis, machine diagnosis
- Automated help desks
- Spam/fraud detection
- Smarter devices, e.g. cameras
- Product recommendations
- Assistants, smart homes
- ... Lots more!

What can AI Systems Currently Do?

... classify incoming e-mails as spam (or not), ...

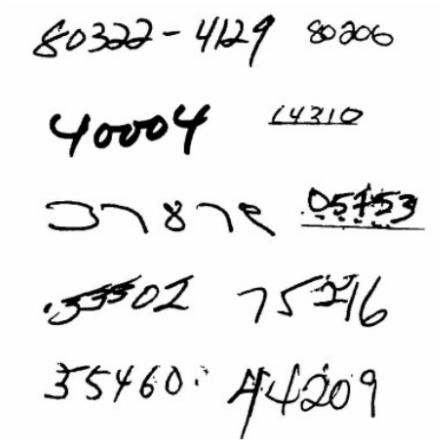
http://www.resilientsystems.co.uk/

What can AI Systems Currently Do?

... predict stock price evolution, ...

What can AI Systems Currently Do?

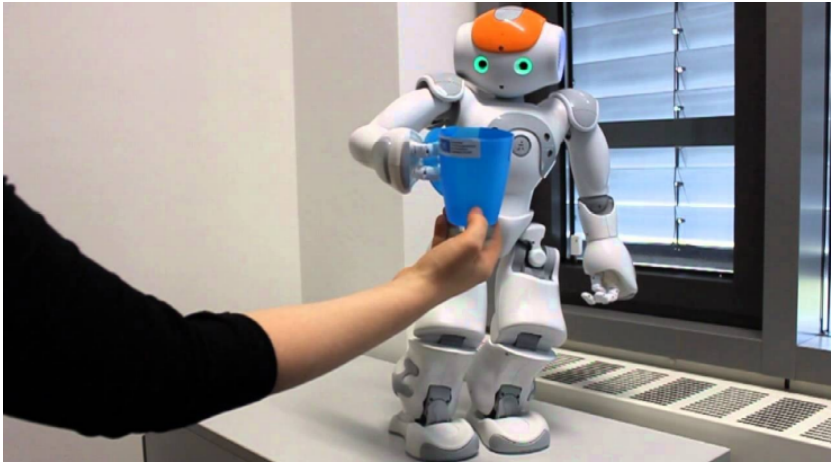
... understanding handwriting, ...



[LeCun et al. 1989]

What can AI Systems Currently Do?

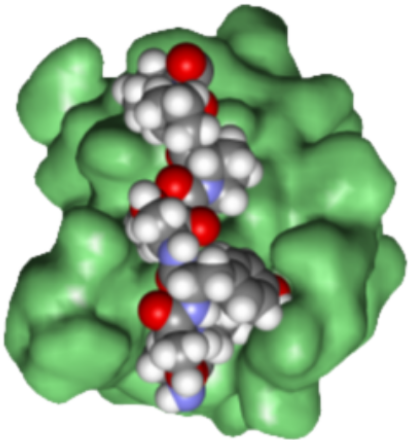
... learn to grab a cup, ...



http://www.informatik.uni-bremen.de/

What can AI Systems Currently Do?

... design a molecule with given properties, ...



http://pande.stanford.edu/

What can AI Systems Currently Do?

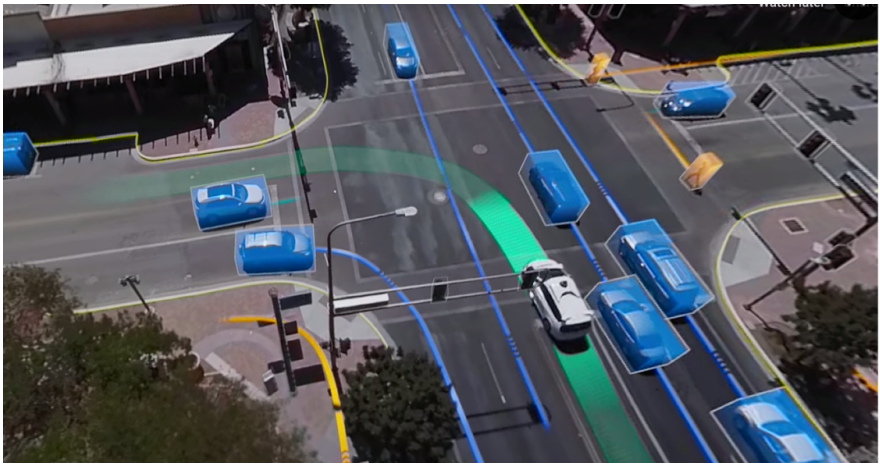
... translate text from Chinese to English, ...

What can AI Systems Currently Do?

... convert a voice into text, ...

What can AI Systems Currently Do?

... predict traffic trajectories, ...



What can AI Systems Currently Do?

... automatically writing the caption of a figure, ...

"man in black shirt is playing guitar."

"construction worker in orange safety vest is working on road."

"two young girls are playing with lego toy."

"boy is doing backflip on wakeboard."

"girl in pink dress is jumping in air."

"black and white dog jumps over bar."

"young girl in pink shirt is swinging on swing."

"man in blue wetsuit is surfing on wave."

[Karpathy & Fei-Fei, 2015; Donahue et al., 2015; Xu et al, 2015;...]

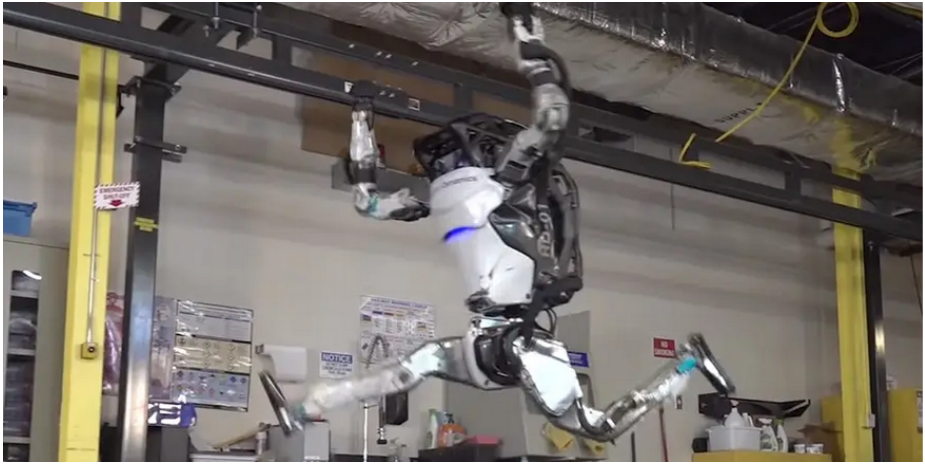
What can AI Systems Currently Do?

... driving autonomously, ...

©Google Inc.

What can AI Systems Currently Do?

... run & jump on two legs, ...



©Boston Dynamics

What can AI Systems Currently Do?

... beat a top-gun pilot in a simulated F16 dogfight, ...

Quiz: What can AI Systems Currently Do?

- Play a decent game of Jeopardy? YES
- Win against any human at chess? YES
- Win against the best humans at Go? YES
- Play a decent game of tennis? YES
- Grab a particular cup and put it on a shelf? YES
- Unload any dishwasher in any home? NO
- Drive safely along the highway? YES
- Drive safely in Naples' center on rush hour? NO
- Buy groceries on the web? YES
- Buy groceries at next corner shop? NO
- Discover and prove a new mathematical theorem? NO
- Perform a surgical operation? NO
- Translate spoken Chinese into spoken English in real time? YES
- Write an intentionally funny story? NO