Course: Introduction to AI

Prof. Ahmed Guessoum
The National Higher School of AI

Chapter 1

Introduction

Outline

- Intelligence
- Definitions of Artificial Intelligence
- The Foundations of AI
- A Brief History of AI
- The State of the Art

What is Intelligence?

Is it the ability to…

- Think?
- Solve problems?
- Sense the environment?
- Plan?
- Take decisions?
- Act?
- Be rational?
- Learn?
- Communicate (in Natural Language or by other means)?
- Have feelings; show emotions; show compassion?
- Reason about morality?

The Realm of AI

- AI encompasses a huge variety of subfields, ranging from
  - the general (learning and perception) to
  - the specific, such as
    - playing chess, proving mathematical theorems, writing poetry, driving a car on a crowded street, diagnosing diseases, etc.
- AI is relevant to any intellectual task.
- Truly a universal field.
- Is AI part of CS or CS part of AI? ;-)

Outline

- Intelligence
- Definitions of Artificial Intelligence
- The Foundations of AI
- A Brief History of AI
- The State of the Art

Various definitions of AI

Thinking Humanly "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)	**Thinking Rationally** "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 "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)	**Acting Rationally** "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)

Acting humanly: The Turing Test approach

Turing (1950), "Computing Machinery and Intelligence":

- "Can Machines think?" → "Can machines behave intelligently?"
- Operational test for intelligent behavior: the Imitation Game

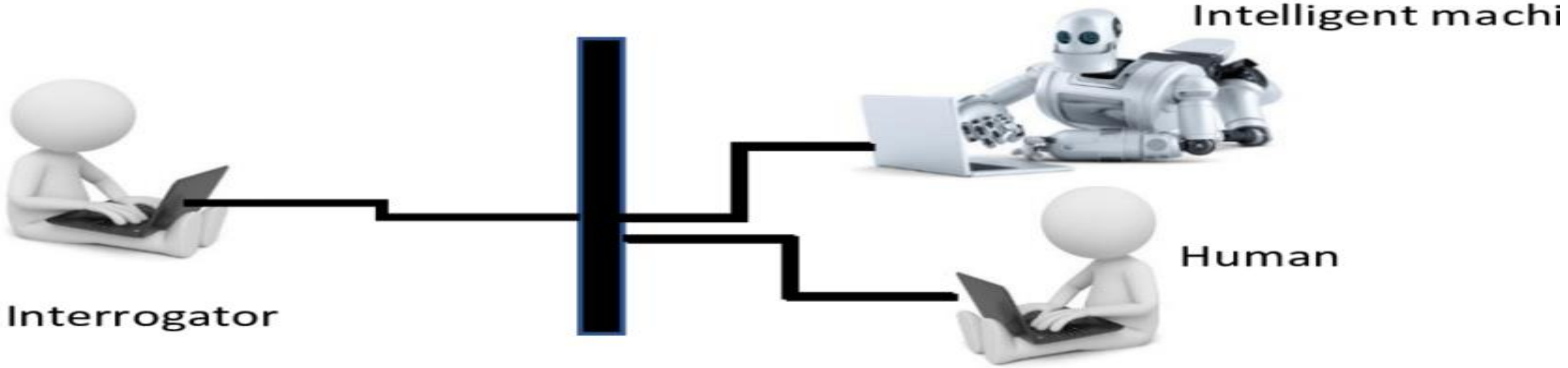


Figure: Pietikäinen, Matti & Silvén, Olli. (2022). Challenges of Artificial Intelligence -- From Machine Learning and Computer Vision to Emotional Intelligence.

The Turing Test

Turing

- Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes.
- Anticipated all major arguments against AI in the following 50 years.
- Suggested major components of AI: knowledge, reasoning, language understanding, learning

Problem: the Turing Test is not reproducible, constructive, or amenable to mathematical analysis.

What computers would need to pass the Turing Test

- **Natural language processing** to enable it to communicate successfully in English;
- **Knowledge representation** to store what it knows or hears;
- **Automated reasoning** to use the stored information to answer questions and to draw new conclusions;
- **Machine learning** to adapt to new circumstances and to detect and extrapolate patterns from data.

The Total Turing Test

- In addition to the previous, it includes
  - a <u>video signal </u>so that the interrogator can test the subject’s perceptual abilities,
  - the opportunity for the <u>interrogator to pass physical objects through the opening</u>.
- So, additionally, a computer would need:
- **Computer vision** to perceive objects, &
- **Robotics** to manipulate objects and move around.
- <u>NOTE</u>: The 6 capabilities cover most of AI!

Thinking humanly: The cognitive modeling approach

- Trying to determine how humans think; to get *inside* the actual workings of human minds.
- Three ways to do this:
  - through introspection—trying to catch our own thoughts as they go by;
  - through psychological experiments;
  - through brain imaging.
- Then develop a sufficiently precise theory of the mind, & express it as a computer program
- Compare the program's I/O behavior to that of a human.
- **Cognitive science** brings together computer models from AI and experimental techniques from psychology to construct precise and testable theories of the human mind.

Thinking rationally: The "laws of thought" approach

- *Aristotle*: one of the first to attempt to codify "right thinking," i.e. irrefutable reasoning processes.
- His **syllogisms** provided patterns for argument structures that always yielded correct conclusions for correct premises (deductive reasoning)
- These laws of thought initiated the field called **logic**
- 19th century: logicians developed a precise notation for statements about all kinds of objects in the world and the relations among them
- The so-called **logicist** tradition within AI hopes to build on such programs to create intelligent systems

Examples of Syllogisms

- All men are mortal.
  All Greeks are men.
  ∴ All Greeks are mortal.
- No reptile has fur.
  All snakes are reptiles.
  ∴ No snake has fur.
- All rabbits have fur.
  Some pets are rabbits.
  ∴ Some pets have fur.
- No homework is fun.
  Some reading is homework.
  ∴ Some reading is not fun.
- Some cats are not pets.
  All cats are mammals.
  ∴ Some mammals are not pets.
- All men are mortal.
  All Greeks are men.
  ∴ Some Greeks are mortal.

Thinking rationally: The "laws of thought" approach

Two main obstacles to this approach:

1. Not easy to take informal knowledge and state it in the formal terms required by logical notation, particularly when the knowledge is less than 100% certain.
2. Complexity: even problems with just a few hundred facts can exhaust the computational resources of any computer unless guided as to which reasoning steps to try first.

Acting rationally: The rational agent approach

- An **agent** is an entity that acts.
- **Computer agents** are expected to do more: operate autonomously, perceive their environment, persist over a prolonged time period, adapt to change, and create and pursue goals.
- A **rational agent** is one that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome.
- The <u>best outcome</u>: that which is expected to maximize goal achievement, given the available information.

Acting rationally

- Correct inference is not *all* of rationality; in some situations, there is no provably correct thing to *do*, but something must still be done.
- There are ways of acting rationally that cannot be said to involve inference: e.g. reflexes; blinking the eye, etc.
- All the skills needed for the (total) Turing Test also allow an agent to act rationally.
- Two advantages of the rational-agent approach over the other approaches.
  1. It is more general than the "laws of thought".
  2. It is more amenable to scientific development than approaches based on human behavior or thought. The standard of rationality is mathematically well defined and can lead to agent designs that provably achieve it.

Outline

- Intelligence
- Definitions of Artificial Intelligence
- The Foundations of AI
- A Brief History of AI
- The State of the Art

The Foundations of AI

- AI has emerged from research in various fields:
  - Philosophy
  - Mathematics
  - Economics
  - Neuroscience
  - Psychology
  - Computer Engineering
  - Control Theory and Cybernetics
  - Linguistics

Philosophy

- Work sparked from some fundamental questions:
  - Can formal rules be used to draw valid conclusions?
  - How does the mind arise from a physical brain?
  - Where does knowledge come from?
  - How does knowledge lead to action?

Philosophy: Reasoning

- **Aristotle (384–322 B.C.):**
  - First to formulate a precise set of laws governing the rational part of the mind.
  - Developed an informal system of syllogisms for proper reasoning, which <u>in principle</u> allowed one to <u>generate conclusions mechanically</u>, given initial premises.
- **Ramon Lull (d. 1315):** had the idea that useful reasoning could actually be carried out by a mechanical artifact.
- **Thomas Hobbes (1588–1679):** proposed that reasoning was like numerical computation, that "we add and subtract in our silent thoughts".

Philosophy: Reasoning

- **Leonardo da Vinci (1452–1519):** Around 1500 designed but did not build a mechanical calculator; recent reconstructions have shown the design to be functional.
- German scientist **Wilhelm Schickard (1592–1635)**: constructed the first known calculating machine around 1623. An adding machine preceding that of Pascal; it could perform multiplication and division.
- **Blaise Pascal (1623–1662):** built the Pascaline in 1642 (more famous than Schickard's). It could add and subtract.
- **Leibniz** built a calculator that could add, subtract, multiply, and take roots. (started in 1673 and completed in 1694)
- **Thomas Hobbes**, in his 1651 book *Leviathan*, suggested the idea of an "artificial animal".

Philosophy: Reasoning

- **René Descartes (1596–1650):** gave the first clear discussion of the distinction between mind and matter and of the problems that arise.
- *If the mind is governed entirely by physical laws, then it has no more **free will** than a rock falling on the ground*.
- Descartes was a strong advocate of the power of reasoning in understanding the world (**rationalism**), but he was also a proponent of **dualism**.
- He held that there is a part of the human mind (or soul or spirit) that is outside of nature, exempt from physical laws.
- An alternative to dualism is **materialism**, which holds that the brain's operation according to the laws of physics *constitutes* the mind.

Philosophy: The source of Knowledge?

- **Empiricism** movement, starting with Francis Bacon (1561– 1626) is characterized by a dictum of John Locke (1632–1704): "*Nothing is in the understanding, which was not first in the senses*." (reference to observation, experimentation).

- **David Hume (1711-1776):** proposed what is now known as the principle of **induction**: that general rules are acquired by exposure to repeated associations between their elements.

Philosophy: The source of Knowledge?

- Building on the work of **Ludwig Wittgenstein** (1889–1951) and **Bertrand Russell** (1872–1970), the famous Vienna Circle, led by **Rudolf Carnap** (1891–1970): developed the doctrine of **logical positivism**.
- **Logical positivism** holds that <u>all knowledge can be characterized by logical theories connected</u>, <u>ultimately</u>, to **observation sentences** that correspond <u>to sensory inputs</u>;

➔ logical positivism combines rationalism and empiricism

- The **confirmation theory** of Carnap and Carl Hempel (1905–1997): attempted to analyze the acquisition of knowledge from experience.

Philosophy: The source of Knowledge?

- Carnap's book *The Logical Structure of the World* (1928) defined an explicit computational procedure for extracting knowledge from elementary experiences.

- It was probably the first theory of the mind as a computational process.

Philosophy: from Knowledge to Action

- Only by understanding how actions are justified can we understand how to build an agent whose actions are justifiable (or rational).
- **Aristotle** argued that actions are justified by a logical connection between goals and knowledge of the actions outcomes.
- He described an "algorithm": "They assume the end and consider how and by what means it is attained, and if it seems easily and best produced thereby; while if it is achieved by one means only they consider *how* it will be achieved by this and by what means *this* will be achieved, till they come to the first cause, . . . and what is last in the order of analysis seems to be first in the order of becoming." (We will see *backward chaining*!)

Philosophy: from Knowledge to Action

- Aristotle's algorithm was implemented 2300 years later by **Newell and Simon** in their GPS (General Problem Solver) program. We would now call it a regression planning system.
- Goal-based analysis is useful, but does not say what to do when several actions will achieve the goal or when no action will achieve it completely.
- **Antoine Arnauld (1612–1694):** correctly described a quantitative formula for deciding what action to take in cases like this.
- **John Stuart Mill (1806–1873)** promoted in 1863 the idea of rational decision criteria in all spheres of human activity.

Mathematics

Some fundamental questions were tackled:

- What are the formal rules to draw valid conclusions?
- What can be computed?
- How do we reason with uncertain information?

- Philosophers developed some of the fundamental ideas of AI, but the leap to a formal science required a level of mathematical formalization in three fundamental areas: <u>logic</u>, <u>computation</u>, and <u>probability</u>.

Mathematics: Logic

- Logic can be traced back to Greek philosophers but its mathematical development really began with the work of **George Boole (1815–1864).**
- He worked out the details of propositional (also called Boolean) logic in 1847.
- **Gottlob Frege (1848–1925):** extended Boole's logic to include objects and relations, creating first-order logic.
- **Alfred Tarski (1902–1983):** introduced theories of truth, logical consequence and logical constants. (A theory of semantics)

Mathematics: Limits of Logic and Computation

- The first nontrivial **algorithm** is thought to be **Euclid**'s algorithm for computing greatest common divisors.
- The word *algorithm* (and the idea of studying them) comes from **al-Khawarazmi**, a Muslim mathematician of the **9th century**, whose writings also introduced Arabic numerals and algebra to Europe.
- **Kurt Gödel (1906–1978):** He proved in 1931 his **incompleteness theorem** that in any formal theory there are <u>true statements that are undecidable</u> in the sense that they have no proof <u>within the theory</u>.

Mathematics: Limits of Logic and Computation

- Godel’s incompleteness theorem is a fundamental result that can also be interpreted as showing that some functions on the integers cannot be represented by an algorithm—that is, they cannot be computed.
- **Alan Turing (1912–1954):** tried to characterize which functions *are* **computable**, i.e. capable of being computed.
- The Church–Turing thesis, which states that the Turing machine (Turing, 1936) is capable of computing any computable function, is generally accepted as providing a sufficient definition.
- Turing also showed that there were some functions that no Turing machine can compute.

Mathematics: Limits of Logic and Computation

- The notion of **tractability** has had an even greater impact.
- Roughly speaking, a problem is called intractable if the time required to solve instances of the problem grows exponentially with the size of the instances.
- Exponential growth means that even moderately large instances cannot be solved in any reasonable time
- Therefore, one should strive to divide the overall problem of generating intelligent behaviour into tractable sub-problems rather than intractable ones.

Mathematics: Limits of Logic and Computation

- **Steven Cook** (1971) and **Richard Karp** (1972), pioneered the theory of **NP-completeness** which provides a method to recognize intractable problems.
- Cook and Karp showed the existence of large classes of canonical combinatorial search and reasoning problems that are NP-complete.
- Any problem class to which the class of NP-complete problems can be reduced is likely to be intractable.

Mathematics: Probability

- Probabilities: third major contribution of mathematics to AI.
- **Gerolamo Cardano (1501–1576)** first framed the idea of probability, describing it in terms of the possible outcomes of gambling events.
- **Blaise Pascal (1623–1662)** showed in 1654 how to predict the future of an unfinished gambling game and assign average payoffs to the gamblers.
- **James Bernoulli (1654–1705), Pierre Laplace (1749–1827),** and others advanced the theory and introduced new statistical methods.
- **Thomas Bayes (1702–1761):** proposed a rule for updating probabilities in the light of new evidence.
- Bayes' rule underlies most modern approaches to uncertain reasoning in AI systems.

Economics: maximising payoff

- **Adam Smith (1723–1790):** the first to treat Economics as a science. He thought of it as consisting of individual agents maximizing their own economic well-being.
- The concept of "preferred outcomes" or **utility** was mathematically formalised by various researchers during the 19th and 20th centuries.
- **Decision theory**, which combines probability theory with utility theory, provides a formal and complete framework for decisions (economic or otherwise) made under uncertainty.
- For "small" economies, it is more like a **game**: the actions of one player can significantly affect the utility of another (either positively or negatively).

Economics: Payoffs from sequences of actions

- Unlike decision theory, **game theory** does not offer an unambiguous prescription for selecting actions.

Question: *How to make rational decisions when payoffs from actions are not immediate but instead result from several actions taken in sequence*?

- Topic pursued in the field of **Operations Research** (from efforts in Britain to optimize radar installations; later found civilian applications in complex management decisions).
- The work of Richard Bellman (1957) formalized a class of sequential decision problems called **Markov decision processes**.

From Economics to...
... AI

- AI researcher **Herbert Simon (1916–2001)** won the Nobel Prize in economics in 1978 for his early work (1947) <u>showing that models based on **satisficing** </u>—making decisions that are "good enough," rather than laboriously calculating an optimal <u>decision—gave a better description of actual human behaviour</u>.

Neuroscience

- How do brains process information?
- **Neuroscience:** study of the nervous system, particularly the brain.
- The exact way in which the brain enables thought is one of the great mysteries of science. (Recall work from early philosophers like Aristotle.)
- **Paul Broca's (1824–1880)** study of aphasia (speech deficit) in brain-damaged patients in 1861 demonstrated the existence of localized areas of the brain responsible for specific cognitive functions.
- **Camillo Golgi (1843–1926)** developed in 1873 a staining technique allowing the observation of individual neurons in the brain.

Neuroscience

- **Santiago Ramon y Cajal (1852–1934)** used Golgi's technique in his pioneering studies of the brain's neuronal structures.
- **Nicolas Rashevsky (1936,1938):** first who applied mathematical models to the study of the nervous system.
- Some data exists on the mapping between areas of the brain and the parts of the body that they control or from which they receive sensory input.
- Not fully understood how other areas can take over functions when one area is damaged.
- Almost no theory on how an individual memory is stored.

Neuroscience

- The recent development of functional magnetic resonance imaging (fMRI) is giving neuroscientists detailed images of brain activity

➔ possible to do measurements that correspond in interesting ways to ongoing cognitive processes.

- Individual neurons can be stimulated electrically, chemically, or even optically

➔ neuronal input–output relationships can be mapped.

- However, we are still far from understanding how cognitive processes actually work.

Neuroscience

	Supercomputer	Personal Computer	Human Brain
Computational units	10^4 CPUs, 10^{12} transistors	4 CPUs, 10^9 transistors	10^{11} neurons
Storage units	10^{14} bits RAM	10^{11} bits RAM	10^{11} neurons
	10^{15} bits disk	10^{13} bits disk	10^{14} synapses
Cycle time	10^{-9} sec	10^{-9} sec	10^{-3} sec
Operations/sec	10^{15}	10^{10}	10^{17}
Memory updates/sec	10^{14}	10^{10}	10^{14}

- Futurists point to an approaching **singularity** at which computers reach a superhuman level of performance.
- But can we ever achieve the brain’s level of intelligence.

Psychology

- How do humans and animals think and act?
- Scientific psychology originated in the work of German physicist **Hermann von Helmholtz (1821–1894)** (work on vision) and his student **Wilhelm Wundt (1832–1920)** (first lab for experimental psychology in 1879)
- **Wundt**: carefully controlled experiments to study relations between perceptual or associative tasks with introspected thought processes.
- Biologists studying animal behaviour developed an objective methodology, as described by **H. S. Jennings (1906)** since introspection was not possible.

Psychology

- The **behaviorism** movement, led by **John Watson (1878–1958)**, rejected *any* theory involving mental processes: introspection provided no reliable evidence.
- Behaviorists insisted on studying only objective measures of the percepts (or *stimulus*) given to an animal and its resulting actions (or *response*).
- **Cognitive psychology**, viewing the brain as an information-processing device, can be traced back to the works of William James (1842–1910).
- **Cognitive viewpoint** eclipsed by behaviourism in the US
- **Frederic Bartlett (1886–1969):** cognitive modeling was able to flourish thanks to his work at the Applied Psychology Unit in Cambridge.

Psychology

- Bartlett's student **Kenneth Craik (1943)**, forcefully reestablished the legitimacy of such "mental" terms as beliefs and goals: being just as scientific as, using pressure and temperature to talk about gases
- **Craik** specified the **three key steps of a knowledge-based agent**:
  - (1) the stimulus must be translated into an internal representation,
  - (2) the representation is manipulated by cognitive processes to derive new internal representations, and
  - (3) these are in turn retranslated back into action.
- He explained why this was a good design for an agent.

Psychology

- **Donald Broadbent**, whose book *Perception and Communication* (1958) was one of the first works to model psychological phenomena as information processing.
- In the United States, the development of computer modeling led to the creation of the field of **cognitive science**.
- In 1956 workshop on Cognitive Science: **George Miller** presented *The Magic Number Seven*, **Noam Chomsky** presented *Three Models of Language*, and **Allen Newell and Herbert Simon** presented *The Logic Theory Machine*.
- These three influential papers showed how computer models could be used to address the psychology of memory, language, and logical thinking, respectively.

Computer engineering

- How can we build an efficient computer?
- **Charles Babbage (1792–1871)** designed two machines, neither of which he completed:
  - The Difference Engine: to compute mathematical tables for engineering and scientific projects. (finally built and shown to work in 1991 at the Science Museum in London)
  - The Analytical Engine: included addressable memory, stored programs, and conditional jumps. The first artifact capable of universal computation.
- Babbage's colleague **Ada Lovelace**, wrote programs for the unfinished Analytical Engine and speculated that the machine could play chess or compose music.
- Read the developments of computers from **Colossus** (Turing's team) to **ENIAC** (at U. of Pennsylvania)
- Since then, performance (speed and capacity) doubled every 18 months or so until around 2005 with a decrease in cost.

Control theory and cybernetics

- How can artifacts operate under their own control?
- **Ctesibius of Alexandria (c. 250 B.C.)** built the first self-controlling machine: a water clock with a regulator that maintained a constant flow rate.
- Other examples of self-regulating feedback control systems include
  - First programmable machines developed by Muslim scientists such as the programmable flute player of the well-known **Banu Musa brothers** (9th Century)
  - Ismail **Al-Jazari**'s (13th Century) "castle clock", considered the first programmable analog computer. (See *Kitab fi ma'rifat al-hiyal al-handasiya*, 'Book in knowledge of engineering tricks', where he described 50 mechanical devices, e.g. the elephant clock, along with instructions on how to construct them.)
  - **Cornelis Drebbel's (1572–1633):** the thermostat
  - **James Watt's (1736–1819):** steam engine governor

Control theory and cybernetics

- The mathematical theory of stable feedback systems was developed in the 19th century.
- **Norbert Wiener (1894–1964)**: central figure in the creation of what is now called **control theory**
- **Wiener** (and **Rosenblueth** and **Bigelow**): purposive behaviour arises from a regulatory mechanism trying to minimize "error" (i.e. difference between current state and goal state).
- **Wiener**'s book *Cybernetics* (1948) became a best seller explaining the possibility of artificially intelligent machines.
- **Ashby**'s *Design for a Brain* (1948, 1952) elaborated on his idea that intelligence could be created by the use of devices with appropriate feedback loops to achieve stable adaptive behaviour.
- Modern control theory has as its goal the design of systems that maximize an **objective function** over time.

Linguistics

- How does language relate to thought?
- Skinner's (1957) *Verbal Behavior* was a comprehensive, detailed account of the behaviourist approach to language learning.
- Noam Chomsky, *Syntactic Structures*, pointed out that the behaviorist theory did not address the notion of creativity in language—it did not explain how a child could understand and make up sentences that he/she had never heard before.
- Chomsky's theory—based on **syntactic models** going back to the Indian linguist Panini (c. 350 B.C.)—could explain this.
- Chomsky's theory was formal enough that it could in principle be programmed.
- Modern linguistics and AI, were "born" around the same time, and grew up intersecting in a hybrid field called **computational linguistics** or **Nat. Lang. Processing**

Outline

- Intelligence
- Definitions of Artificial Intelligence
- The Foundations of AI
- A not-so Brief History of AI
- The State of the Art

The gestation of AI (1943–1955)

- Warren McCulloch and Walter Pitts (1943) proposed **a model of artificial neurons**:
  - A neuron is characterized as being "on" or "off," with a switch to "on" occurring in response to stimulation by a sufficient number of neighbouring neurons.
  - They showed that any computable function (including *and*, *or*, *not*, etc.) could be computed by some network of connected neurons.
  - They suggested that suitably defined networks could learn.
- **Donald Hebb (1949)** demonstrated a **simple updating rule** for modifying the connection strengths between neurons. His rule, now called **Hebbian learning**, remains an influential model to this day.

The gestation of AI (1943–1955)

- **Marvin Minsky** and **Dean Edmonds** (then undergraduates at Harvard) built in 1950 SNARC, the first neural network computer. It used 3000 vacuum tubes to simulate a network of 40 neurons.
- Later, **Minsky** proved influential theorems showing the limitations of neural network research.
- In 1950, **Turing** introduced the Turing Test, machine learning, genetic algorithms, and reinforcement learning.
- **Turing** suggested the idea that *computers should start like a child* and then learn/improve instead of starting with an *adult mind*.

The Birth of AI (1956)

- **McCarthy** convinced **Minsky**, **Claude Shannon**, and **Nathaniel Rochester** to help him bring together U.S. researchers interested in automata theory, neural nets, and the study of intelligence.
- They organized a two-month workshop at **Dartmouth College** in the summer of 1956 for 10 big names.
- Aim: study/discuss how to make machines use language, form abstractions and concepts, solve kinds of problems considered reserved for humans, and improve themselves.
- **Newell and Simon** showed, the *Logic Theorist* (LT), about which Simon claimed a computer program capable of thinking non-numerically.

The Birth of AI (1956)

- Soon after the workshop, *Logic Theorist* was able to prove most of the theorems in Chapter 2 of Russell and Whitehead's *Principia Mathematica*.
- Logic Theorist came up with a proof for one theorem that was shorter than the one in *Principia*. The editors of the *Journal of Symbolic Logic* were less impressed; they rejected a paper coauthored by Newell, Simon, and Logic Theorist.
- Outcomes of the Dartmouth workshop:
  - Did not lead to any new breakthroughs, **BUT**
  - It did introduce all the major figures to each other. These (and their students at MIT, CMU, Stanford, and IBM) dominated the field for the next 20 years.

Early enthusiasm, great expectations (1952–1969)

- The early years of AI were full of (relative) successes, i.e. given the computational limitations.
- **Newell & Simon** introduced the *General Problem Solver*.
- Unlike Logic Theorist, GPS was designed to imitate human problem-solving protocols. The order in which the program considered sub-goals and possible actions was similar to that in which humans approached the same problems. (Thinking Humanly approach)
- In 1976, they formulated the **Physical Symbol System hypothesis**, which states that "a P.S.S. has the necessary and sufficient means for general intelligent action." That is, any system (human or machine) exhibiting intelligence must operate by manipulating data structures composed of symbols.

Early enthusiasm, great expectations (1952–1969)

- At IBM, **Rochester** (**et al.**) produced some of the first AI programs.
- **Gelernter** (1959) constructed the *Geometry Theorem Prover*, which was able to prove theorems that many students of mathematics would find quite tricky.
- Starting in 1952, **Samuel** wrote a series of programs for checkers.
- His program quickly learned to play a better game than its creator. Demonstrated on TV in 1956.

Early enthusiasm, great expectations (1952–1969)

- In 1958, McCarthy (who moved to MIT) defined the high-level language **Lisp**, which dominated AI as a programming language for the next 30 years (especially in the US).
- To handle computation resources limitations, he and others at MIT invented time sharing.
- McCarthy (1958) published a paper "*Programs with Common Sense*", in which he described the Advice Taker:
  - A hypothetical program that can be seen as the first complete AI system.
  - It was designed to solve problems by also taking into consideration World Knowledge.

Early enthusiasm, great expectations (1952–1969)

- Advice Taker:
  - It was designed to accept new axioms in the normal course of operation, ➔ ability to achieve competence in new areas *without being reprogrammed*.
  - It embodied the central principles : to have a formal, explicit representation of the world and its workings and manipulate it with deductive processes.
- J. A. Robinson's discovery (1965) of **the resolution method** was a complete theorem-proving algorithm for first-order logic.
- Work at Stanford (where McCarthy was) advanced general-purpose methods for logical reasoning, included question-answering and planning systems, and Robotics.

Early enthusiasm (1952–1969)

- Minsky and colleagues at MIT (and their students) worked on the **microworlds** (limited-domain) problems which led to developing/solving:
  - Closed-form calculus integration problems
  - Geometric analogy problems (in IQ tests)
  - Algebra story problems
  - Vision and constraint-propagation
  - The learning theory
  - A natural-language-understanding program
  - A planner program
- Hebb's learning methods were enhanced by Bernie Widrow (Widrow and Hoff,1960; Widrow, 1962), and by Rosenblatt (1962) with his **perceptrons**.
- The **perceptron convergence theorem** was proved.

A dose of reality (1966–1973)

- AI researchers were making predictions of their coming successes
- <u>1st problem</u>: In almost all cases, however, these <u>early systems</u> turned out to <u>fail miserably when tried out on wider selections of problems and on more difficult problems</u>.
- Same for **Machine Translation** (too simplistic then) generously funded by the U.S. National Research Council in an attempt to <u>speed up the translation of Russian scientific papers</u> in the wake of the Sputnik launch in 1957.
  - E.g. famous retranslation of "the spirit is willing but the flesh is weak" as "the vodka is good but the meat is rotten".
  - In 1966, upon an advisory report on MT, <u>All U.S. government funding for academic translation projects was canceled.</u>

A dose of reality (1966–1973)

- 2nd problem: **intractability of many of the problems** AI tackled:
- Originally, AI solved limited problems that had very few objects and hence very few possible actions and very short solution sequences.
- Scaling up was a totally different problem requiring faster hardware and larger memories.
- Theorem provers failed to prove theorems involving more than a few dozen facts.
- Early experiments in **machine evolution** (now called **genetic algorithms**) using random mutations and combinations: though the principle was correct, despite thousands of hours of CPU time, almost no progress was demonstrated.
- Due to the **failure to handle the combinatorial explosion**, the British government ended support for AI research in all but two universities.

A dose of reality (1966–1973)

- Third difficulty: some fundamental limitations on the basic structures being used to generate intelligent behaviour.
- **Minsky and Papert's** book *Perceptrons* (1969) proved that, although perceptrons could be shown to learn anything they were capable of representing, they could represent very little. (e.g. inability of a perceptron to recognize when its two inputs were different)
- As a consequence, funding to NN research went to almost zero.

Knowledge-based systems: The key to power? (1969–1979)

- Original approaches to problem solving (a general-purpose search mechanism trying to do elementary reasoning steps to find complete solutions) were called **weak methods** because <u>they did not scale up</u> to large/difficult problem instances.
- <u>Alternative to weak methods</u>: use more powerful, domain-specific knowledge that allows larger reasoning steps and can more easily handle typically occurring cases in narrow areas of expertise.
- Example: The <u>DENDRAL</u> program (**Buchanan, Feigenbaum, Lederberg**): inferring molecular structure from the information provided by a mass spectrometer.
- DENDRAL: first successful *knowledge-intensive* system; its expertise derived from large numbers of specialised rules.

Knowledge-based systems: The key to power? (1969–1979)

- MYCIN (**Shortliffe, Buchanan**, and others) to diagnose blood infections:
  - About 450 rules; MYCIN was able to perform as well as some experts, and considerably better than junior doctors
  - No general theoretical model existed from which the MYCIN rules could be deduced. Rules acquired from extensive interviewing of experts.
  - The rules had to reflect the uncertainty associated with medical knowledge (Certainty Factors)

Knowledge-based systems: The key to power? (1969–1979)

- The need for domain knowledge became apparent in the area of understanding natural language.
- E.g. **Winograd's** SHRDLU NLU system engendered a good deal of excitement but had limitations (reliance on syntactic structures, limited pronoun inference and ambiguity resolution).
- Several researchers suggested that **robust language understanding** would require general knowledge about the world and a general method for using that knowledge.
- **Schank** and his students built a series of NLU programs with emphasis less on language *per se* and more on K. representation and reasoning.
- **Prolog** became popular in Europe
- **Minsky's** idea of **frames** (1975) adopted for K. rep.

AI becomes an industry (1980–present)

- **R1** (Digital Equipment Corporation, McDermott, 1982): first successful commercial expert system.
  - R1 helped configure orders for new computer systems; by 1986, it was saving the company an estimated $40 million a year.
  - By 1988, DEC's AI group had 40 expert systems deployed, with more on the way
- In the **80s**, nearly every major U.S. corporation had its own AI group and was either using or investigating expert systems.
- 1981, the Japanese announced the "**Fifth Generation Computer Systems" project**, a 10-year plan to build intelligent computers running Prolog. $400 million.

AI becomes an industry (1980–present)

- The United States formed the Microelectronics and Computer Technology Corporation (MCTC) as a research consortium designed to assure national competitiveness.
- In Britain, <u>the Alvey report reinstated the funding that was cut by the Lighthill report</u>.
- In all three countries (Japan, USA, UK), the projects never met their ambitious goals
- Overall, the <u>AI industry boomed</u> from a few million dollars in 1980 <u>to billions of dollars in 1988</u>.
- The "AI winter" soon followed!

The return of neural networks (1986–present)

- In the mid-1980s at least four different groups reinvented the **back-propagation** learning algorithm first found in 1969 by Bryson and Ho.
- Backpropagation was applied to many learning problems in computer science and psychology.

  ➔ The results caused excitement.
- **Connectionist** models of intelligent systems were seen by some as competitors both to the symbolic models promoted by Newell and Simon and to the logicist approach of McCarthy and others.
- Current view: connectionist and symbolic approaches are complementary, not competing

AI adopts the scientific method (1987–present)

- More recent AI approach:
  - Build on existing theories (e.g. Information theory, stochastic modelling, classical optimization and control, etc.) rather than propose brand-new ones,
  - Develop rigorous theorems or hard experimental evidence rather than intuition, and
  - Work on real-world applications rather than toy examples
- Rigorous empirical experiments, and the results must be analyzed statistically for their importance.
- It is now possible to replicate experiments by using shared repositories of test data and code.
- In speech recognition, mathematically-founded **hidden Markov models** (HMMs), trained on large datasets, have replaced ad-hoc methods of the 1970s.

AI adopts the scientific method (1987–present)

- Speech technology and handwritten character recognition are already making the transition to widespread industrial and consumer applications.
- Using improved methodology and theoretical frameworks, neural nets and Deep Learning have revolutionized AI areas (vision, NLP, speech, etc.) ➔ Data Mining technology!
- The **Bayesian network** formalism was invented to allow efficient representation of, and rigorous reasoning with, uncertain knowledge ➔ Learning from experience.
- In robotics, computer vision, and knowledge representation, a better understanding of the problems and their complexity properties, combined with increased mathematical sophistication, has led to robust methods.

Emergence of intelligent agents

- **Newell, Laird, and Rosenbloom** (Newell, 1990; Laird *et al.*, 1987), designed SOAR: a complete agent architecture.
- **Internet**:
  - excellent environment for agents (the "-bot" suffix in everyday language).
  - tools such as search engines, recommender systems, and web site aggregators
- Agent perspective ➔ AI has been drawn into much closer contact with other fields (control theory, economics, etc.) that also deal with agents.
  - Example: progress in the control of self-driving cars (sensors, planning, vision, etc.)
- Current targets: **Human-level AI** HLAI (machines that think, learn, and create) and related idea of **Artificial General Intelligence** or AGI (a universal algorithm for learning and acting in any environment)

Availability of very large data sets (2001–present)

- During the **1960s**, emphasis on the *algorithms*.
- Common understanding today that, for many problems, it makes more sense to worry about the *data* and be less picky about which algorithm to apply.
- Increased availability of web data and genomic sequences (think of time-record development of COVID vaccine).
- Programmes exclusively trained on data can have excellent performances (compared to other algorithms) e.g. word-sense disambiguation, masking objects in images.
- The K. bottleneck (how much K. needs to be represented in K. engineering) may be solved in many applications using learning from data.

The State of the Art

- Autonomous cars
- Speech recognition (using bots)
- Autonomous planning and scheduling, logistics planning (NASA, the US Air Force)
- Game playing (AlphaGo)
- Spam detection
- Robotics (iRobot's Roomba)
- Machine Translation (Google Translate)

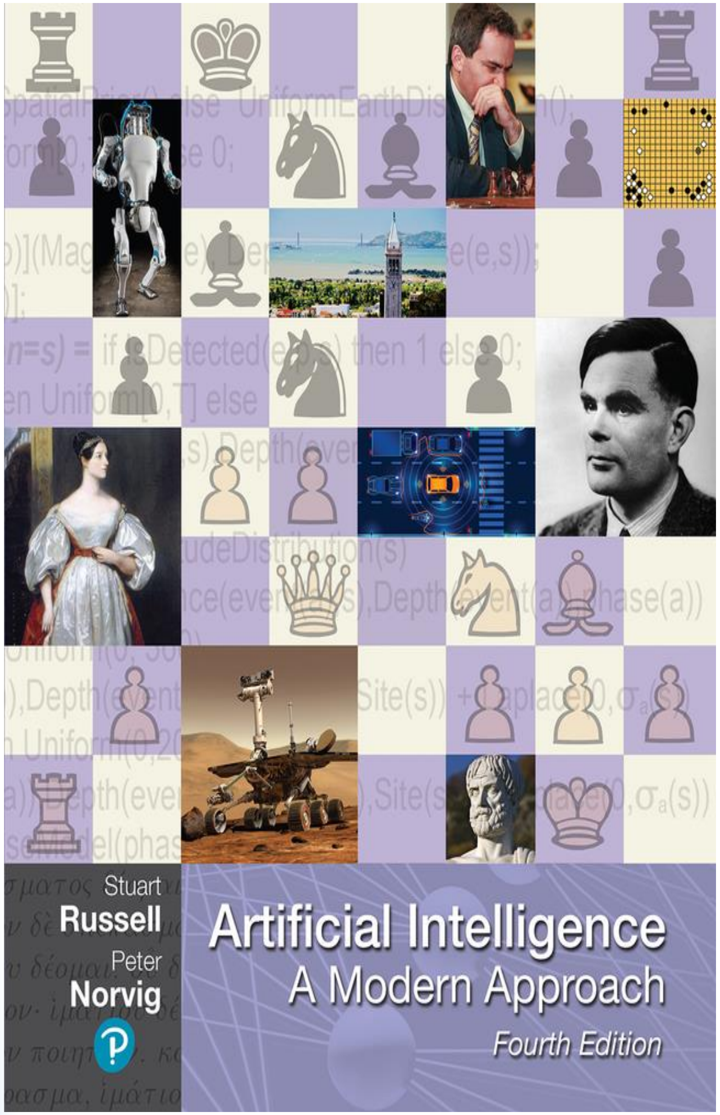
and ….

- ChatGPT!

More readings

- http://www-formal.stanford.edu/jmc/slides/dartmouth/dartmouth/node1.html

Slides based on the textbook



- Russel, S. and Norvig, P. (2020) Artificial Intelligence,

A Modern Approach (4th Edition), Pearson Education Limited.