國立成功大學
National Cheng Kung University
1931

Introduction to Artificial Intelligence

Chapter 1
Introduction

Wei-Ta Chu (朱威達)

What is AI?

- AI started in earnest soon after World War II, and the name itself was coined in 1956.

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)	thought process
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)	behavior

Figure 1.1 Some definitions of artificial intelligence, organized into four categories.

Acting Humanly: The Turing Test

- A computer passes the Turing test if a human interrogator (訊問者), after posing some written questions, cannot tell whether the written responses come from a person or from a computer.
- The computer needs
  - Natural language processing
  - Knowledge representation
  - Automated reasoning
  - Machine learning
  - Computer vision
  - Robotics

Thinking Humanly: Cognitive Modeling

- We must have some way of determining how humans think
  - Through introspection (反思)
  - Through psychological experiments
  - Through brain imaging
- **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: Law of Thought

- By 1965, programs existed that could, in principle, solve any solvable problem described in logical notation. The so-called **logicist** tradition within artificial intelligence hopes to build on such programs to create intelligent systems.
- Two main obstacles
  - It’s not easy to take informal knowledge and state it in the formal terms required by logical notation.
  - There is a big difference between solving a problem “in principle” and solving it in practice. Even problems with just a few hundred facts can exhaust the computational resources.

Acting Rationally: Rational Agent

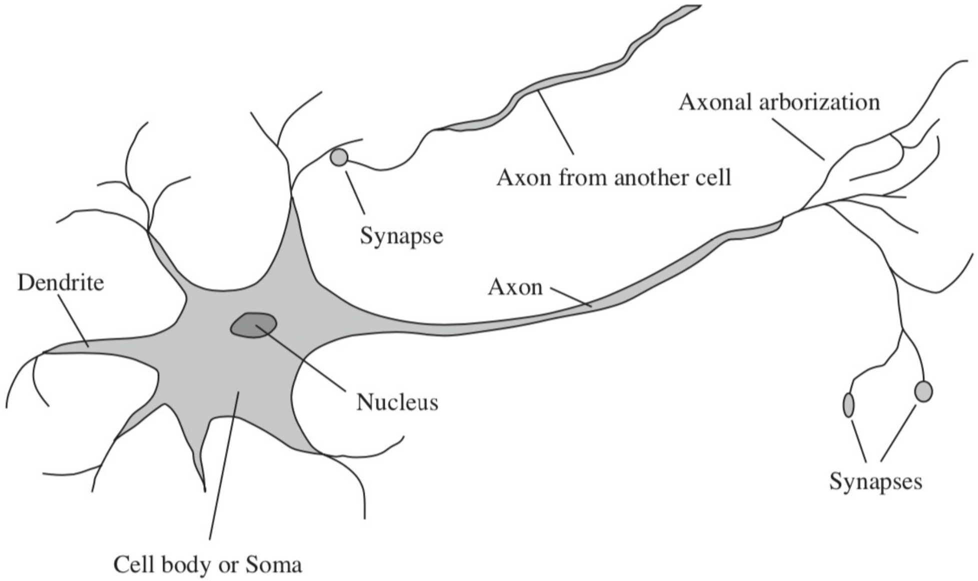
- A **rational agent** is one that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome.
- Two advantages
  - It is more general than the "laws of thought" approach because correct inference is just one of several possible mechanisms for achieving rationality.
  - It is more amenable (經得起檢驗) to scientific development than are approaches based on human behavior or human thought.

Foundations of AI

- Philosophy
  - Can formal rules be used to draw valid conclusions?
  - How does the mind arise from a physical brain?
  - What does knowledge come from?
  - How does knowledge lead to action?
- Mathematics
  - What are the formal rules to draw valid conclusions?
  - What can be computed?
  - How do we reason with uncertain information?

Foundations of AI

- Economics
  - How should we make decisions so as to maximize payoff?
  - How should we do this when others may not go along?
  - How should we do this when the payoff may be far in the future?
- Neuroscience
  - How do brains process information?



國立成功大學
National Cheng Kung University
1931

Foundations of AI

- Psychology
  - How do humans and animals think and act?
- Computer Engineering
  - How can we build an efficient computer?
- Control theory and cybernetics
  - How can artifacts operate under their own control?
- Linguistics
  - How does language relate to thoughts?

History of AI

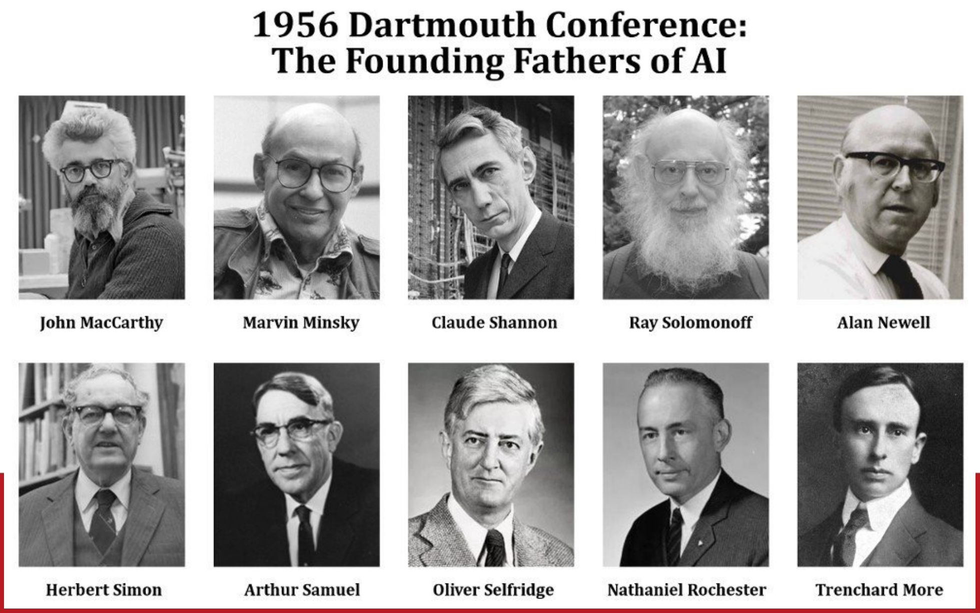
- One quick way to summarize AI history
  - Marvin Minsky (1969) and John McCarthy (1971) defined the foundations of the field based on representation and reasoning
  - Allen Newell and Herbert Simon (1975) for symbolic models
  - Ed Feigenbaum and Raj Reddy (1994) for developing expert systems
  - Judea Pearl (2011) for developing probabilistic reasoning techniques
  - Yoshua Bengio, Geoffrey Hinton, and Yann LeCun (2019) for making deep learning

History of AI

- The inception (開端) of artificial intelligence (1943–1956)
  - Warren McCulloch and Walter Pitts (1943) proposed a model of artificial neuron. They showed that any computable function could be computed by some network of connected neurons, and logical connectives could be implemented.
  - Donald Hebb (1949) demonstrated a simple updating rule for modifying the connection strengths between neurons.
  - Marvin Minsky and Dean Edmonds built the first neural network computer in 1950.
  - Alan Turing gave lectures on the topic as early as 1947.

History of AI

- The birth of artificial intelligence (1955-56)
  - John 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 in the summer of 1956.



Founding fathers of AI. Courtesy of scienceabc.com

History of AI

- Early enthusiasm, great expectations (1952–1969)
  - Newell and Simon's early success was followed up with the General Problem Solver (GPS). GPS was probably the first program to embody the "thinking humanly" approach.
  - Newell and Simon (1976) formulated the physical symbol system hypothesis -- any system (human or machine) exhibiting intelligence must operate by manipulating data structures composed of symbols.

History of AI

- Early enthusiasm, great expectations (1952–1969)
  - At IBM, Nathaniel Rochester and his colleagues produced some of the first AI programs.
  - Herbert Gelernter (1959) constructed the Geometry Theorem Prover, which was able to prove theorems.
  - Starting in 1952, Arthur Samuel wrote a series of programs for checkers (西洋棋) that eventually learned to play at a strong amateur level.

History of AI

- Early enthusiasm, great expectations (1952–1969)
  - In 1958, McCarthy defined the high-level language **Lisp**, which was to become the dominant AI programming language for the next 30 years. He described the Advice Taker, a hypothetical program that can be seen as the first complete AI system.
  - Minsky supervised a series of students who chose limited problems that appeared to require intelligence to solve. These limited domains became known as **microworlds**.
  - Early work building on the neural networks of McCulloch and Pitts also flourished.

History of AI

- Early enthusiasm, great expectations (1952–1969)
  - The most famous microworld is **blocks world**

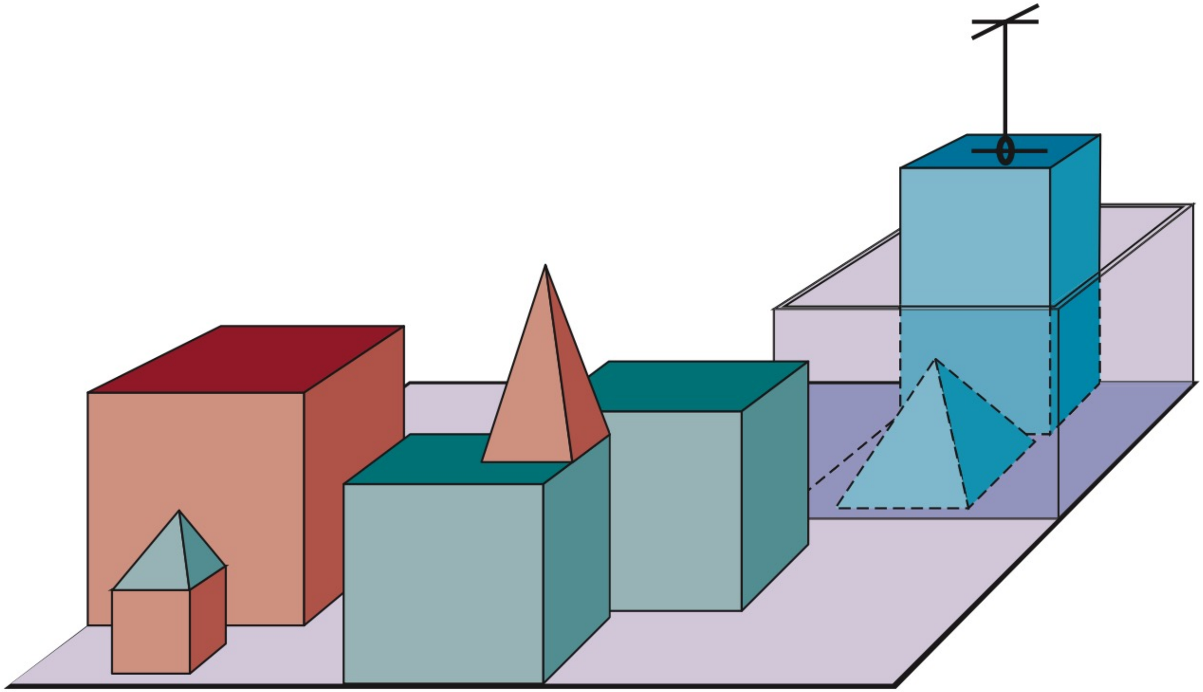


Figure 1.3 A scene from the blocks world. SHRDLU (?) has just completed the command "Find a block which is taller than the one you are holding and put it in the box."

History of AI

- A dose of reality (1966–1973)
  - AI researchers' overconfidence was due to the promising performance of early AI systems on simple examples.
  - A typical story occurred in early machine translation efforts. The famous retranslation of "the spirit is willing but the flesh is weak" (心有餘而力不足) as "the vodka is good but the meat is rotten" illustrates the difficulties.
  - Scalability: *A program can find a solution in principle does not mean that the program contains any of the mechanisms needed to find it in practice.*
  - Some fundamental limitations on the basic structures being used to generate intelligent behavior.

History of AI

- Expert systems (1969–1986)
  - Previous methods are **weak methods,** which don't scale up to large or difficult problems.
  - The alternative to weak methods is to use more powerful, domain-specific knowledge that allows larger reasoning steps and can more easily handle typically occurring cases in narrow areas of expertise.
  - The DENDRAL program (Buchanan et al., 1969) was developed to solve the problem of inferring molecular structure from the information provided by a mass spectrometer.

History of AI

- Expert systems (1969–1986)
    - Feigenbaum and others at Stanford developed MYCIN to diagnose blood infections. With about 450 rules, MYCIN was able to perform as well as some experts, and considerably better than junior doctors.
    - The importance of domain knowledge was also apparent in the area of understanding natural language.

History of AI

- Expert systems (1969–1986)
  - The AI industry boomed from a few million dollars in 1980 to billions of dollars in 1988, including companies building expert systems, vision systems, robots, and software and hardware specialized for these purposes.
  - Soon after that came a period called the “AI Winter,” in which many companies fell by the wayside as they failed to deliver on extravagant promises.

History of AI

- 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.
  - These so-called **connectionist** models of intelligent systems were seen by some as direct competitors both to the **symbolic** models promoted by Newell and Simon and to the **logicist** approach of McCarthy and others (Smolensky, 1988).

History of AI

- Probabilistic reasoning and machine learning (1987–present)
  - Approaches incorporating probability rather than Boolean logic, machine learning rather than hand-coding, and experimental results rather than philosophical claims.
  - Shared benchmark problem sets became the norm for demonstrating progress.
  - In the 1980s, approaches using **hidden Markov models** came to dominate the area of speech recognition. They are based on a rigorous math theory. They are generated by a process of training on a large corpus of real speech data.

History of AI

- Probabilistic reasoning and machine learning (1987–present)
  - Note that there was no scientific claim that humans use HMMs to recognize speech.
  - In 1988 Judea Peral's development of **Bayesian networks** yielded a rigorous and efficient formalism for representing uncertain knowledge as well as practical algorithms for probabilistic reasoning.
  - In 1988 Rich Sutton's work connecting reinforcement learning to the theory of Markov decision processes developed in the field of operations research.

History of AI

- Big data (2001–present)
  - Remarkable advances in computing power and the creation of WWW have facilitated the creation of very large data sets – **big data**. This has led to the development of learning algorithms specially designed to take advantage of very large data sets.
  - The availability of tens of millions of images in the ImageNet database sparked a revolution in the field of computer vision.
  - Big data was a crucial factor in the 2011 victory of IBM's Watson system over human champions in the Jeopardy! Quiz gam.

History of AI

- Deep learning (2011–present)
  - The term deep learning refers to machine learning using multiple layers of simple, adjustable computing elements. **Convolutional neural networks** are found some success in hand-written digit recognition in the 1990s.
  - In the 2012 ImageNet competition, a deep learning system created in Geoffrey Hinton's group showed a dramatic improvement over previous systems. Since then, deep learning systems have exceeded human performance on some vision tasks.
  - AlphaGo

The State of the Art

- Publications: AI papers increased 20-fold between 2010 and 1019.
- Sentiment: About 70% of news articles on AI are neutral, but articles with positive tone increased from 12% in 2016 to 30% in 2018.
- Students: Course enrollment increased 5-fold in the U.S. and 16-fold internationally from a 2010 baseline.
- Diversity: AI Professors worldwide are about 80% male, and 20% female.
- Conferences: Attendance at NeurIPS increased 800% since 2012.
- Industry: AI startups in the U.S. increased 20-fold to over 800.
- Internationalization: China publishes more paper per year than the U.S. and about as many as all of Europe.

The State of the Art

- Vision: Error rates for object detection improved from 28% in 2010 to 2% in 2017. Accuracy on open-ended VQA improved from 55% to 68% since 2015.
- Speed: Training time for the image recognition task dropped by a factor of 100 in just the past two years.
- Language: Accuracy on question answering, as measured by F1 score on the Stanford Question Answering Dataset, increased from 6- to 95 from 2015 to 2019.
- Human benchmarks: By 2019, AI systems had reportedly met or exceeded human-level performance in chess, Go, poker, Pac-Man, Jeopardy!, ImageNet object detection, speech recognition in a limited domain, …

Risks and Benefits of AI

- Benefits: the potential for AI and robotics to free humanity from repetitive work and to dramatically increase the production of goods and services could presage an era of peace and plenty.
- Risks:
  - Lethal autonomous weapons
  - Surveillance and persuasion
  - Biased decision making
  - Impact on employment
  - Safety-critical applications
  - Cybersecurity