Fundamentals of Artificial Intelligence

Introduction to Artificial Intelligence

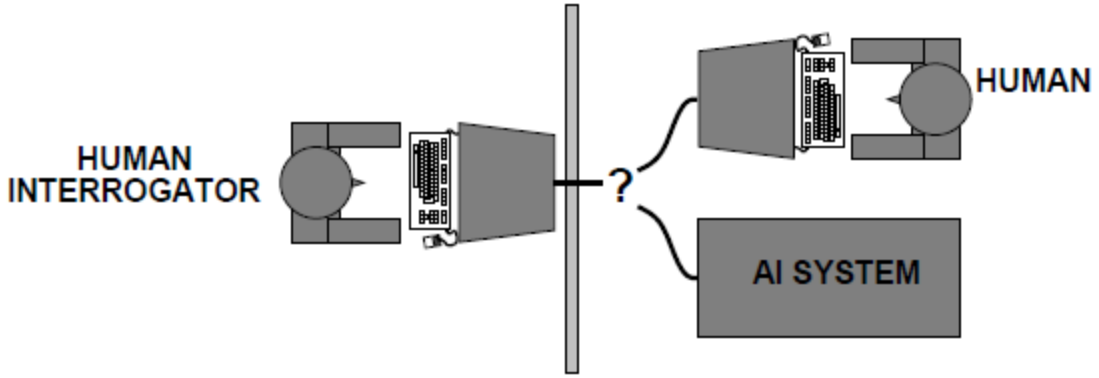
What is Artificial Intelligence?

- The art of creating machines that perform functions that require intelligence when performed by people.
- The branch of computer science that is concerned with the automation of intelligent behavior.

- Views of AI fall into four categories:
  - Systems that act like humans
  - Systems that think like humans
  - Systems that think rationally
  - Systems that act rationally

Acting humanly: The Turing Test

- Turing Test, proposed by Alan Turing (1950), was designed to provide a satisfactory operational definition of **intelligence**
- Can machines think? ➔ Can machines behave intelligently?
- Operational test for intelligent behavior: the **Imitation Game**



- Suggested major components of AI system: knowledge representation, automated reasoning, natural language understanding, machine learning.

Thinking humanly: Cognitive Science

- The field of **cognitive science** brings together computer models from AI and experimental techniques from psychology to construct precise and testable theories of the human mind.
- Requires scientific theories of internal activities of the brain
  - What level of abstraction? Knowledge or Circuits?
  - How to validate? Requires:

    1) Predicting and testing behavior of human subjects (top-down), or

    2) Direct identification from neurological data (bottom-up)
- Both approaches (**Cognitive Science** and **Cognitive Neuroscience**) are now distinct from AI.
- Both share with AI the following characteristic: the available theories do not explain anything resembling human-level general intelligence

Thinking rationally: Laws of Thought

- Aristotle: what are correct arguments/thought processes?
- Development of various forms of **logic**:
  - notation and rules of derivation for thoughts.
  - may or may not have proceeded to the idea of mechanization
- Direct line through mathematics and philosophy to modern AI
- Main obstacles to this approach.
  - First, it is 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.
  - Second, there is a big difference between solving a problem “in principle” and solving it in practice.

Acting rationally: The rational agent approach

- An **agent** is just something that acts.
- A **rational agent** is one that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome.
- In the “laws of thought” approach to AI, the emphasis was on *correct inferences*.
  - Making correct inferences is sometimes part of being a rational agent, and
  - There are also ways of acting rationally that cannot involve inference.
- **Rational behavior**: doing the right thing
  - The right thing: that which is expected to maximize goal achievement, given the available information
  - Doing the right thing doesn't necessarily always involve thinking but thinking should be in the service of rational action

Acting rationally: Rational agents

- An **agent** is an entity that perceives and acts
- **This AI course concentrates on general principles of rational agents and on components for constructing them.**
- Abstractly, an agent is a function from percept histories to actions:

  $$\mathbf{f : P^* \rightarrow A}$$

- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance.

- Computational limitations can make perfect rationality unachievable

  ➔ Design best program for given machine resources

Foundations of AI

- Many disciplines contribute to a foundation for artificial intelligence.
  - **Philosophy**: logic, methods of reasoning, mind as physical system
  - **Mathematics**: formal representation and proof, algorithms, computation, decidability, probability
  - **Psychology**: phenomena of perception and motor control
  - **Economics**: formal theory of rational decisions
  - **Linguistics**: knowledge representation, grammar
  - **Neuroscience**: how the brain works
  - **Computer engineers** provided powerful machines that make AI applications possible.
  - **Control theory**: designing devices that act optimally on the basis of feedback from the environment

History of Artificial Intelligence

1943 McCulloch & Pitts: Boolean circuit model of brain

- The first work that is now generally recognized as AI.

1950 Turing's “Computing Machinery and Intelligence"

- Turing introduced the Turing Test, machine learning, and reinforcement learning.

1950s Early AI programs:

- Samuel's checkers program, Newell & Simon's Logic Theorist

1956 Dartmouth meeting: Birth of “Artificial Intelligence"

- McCarthy, Minsky and others come together to study artificial intelligence.

1958 McCarthy defined the high-level language Lisp,

- Lisp was to become the dominant AI programming language

1965 Robinson’s discovery of the resolution method

- a complete theorem-proving algorithm for first-order logic

History of Artificial Intelligence

1966-74 AI discovers computational complexity

- Neural network research almost disappears

1969-79 Early development of knowledge-based systems

1980-88 Expert systems industry booms

1985-95 Neural networks return to popularity

1987- AI adopts the scientific method

- Hidden Markov Models (HMMs), Bayesian network formalism for uncertainty, data mining

2001- Availability of very large data sets

The State of The Art

- What can AI do today?
  - **Robotic vehicles**: driverless robotic cars
  - **Speech recognition**
  - **Autonomous planning and scheduling**: NASA's Remote Agent program became the first on-board autonomous planning program to control the scheduling of operations for a spacecraft
  - **Game playing**: IBM's DEEP BLUE became the first computer program to defeat the world champion in a chess match
  - **Spam fighting**
  - **Robotics**
  - **Machine Translation**