Unit 1
Introduction

Syllabus

Unit 1: Introduction

1.1 What is AI?
  1.1.1 Turing test approach: Chinese room argument
  1.1.2 Cognitive approach
  1.1.3 Laws of thought approach
  1.1.4 Rational agent approach
1.2 Difference between AI and Omniscience

Intelligence

Intelligence is:

- the ability to reason
- the ability to understand
- the ability to create
- the ability to Learn from experience
- the ability to plan and execute complex tasks

Artificial Intelligence

Artificial intelligence (AI) is a wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence. Thus, artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving.

AI consists of the design of intelligent agents, which is a program that perceives its environment and takes action that maximizes its chance of success. With Ai it comes issues like deduction, reasoning, problem solving, knowledge representation, planning, learning, natural language processing, perceptron, etc.

Different definitions of AI can be divided into 2 x 2 matrix.

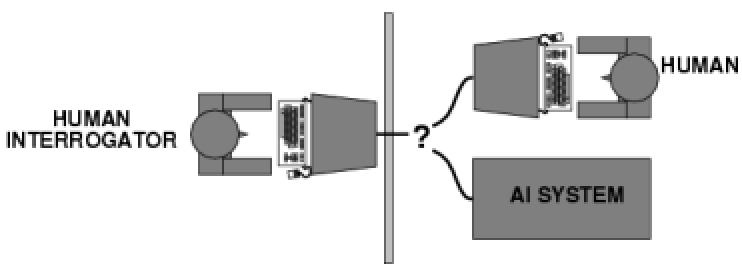
Think Humanly System that thinks like humans	**Think Rationally** System that think rationally
Acting Humanly System that acts like humans	**Acting Rationally** System that act rationally

Top dimension is concerned with **thought processes and reasoning**, whereas the bottom dimension **addresses the behavior**.
The definition on the left measures the success in terms of **fidelity of human performance**, whereas definitions on the right measure an ideal **concept of intelligence, which is called rationality**. A system is rational if it does the ***right things*** given what it knows.

Human-centered approaches must be an empirical science, involving hypothesis and experimental confirmation. A rationalist approach involves a combination of mathematics and engineering.

Acting Humanly: Turing Test Approach

The art of creating machines that perform functions that require intelligence when performed by humans. To support this, the Turing test, proposed by Alan Turing (1950) was designed to convince the people whether a particular machine can think or not. He suggested a test based on indistinguishability from undeniably intelligent entities- human beings.
The test involves an interrogator who interacts with one human and one machine. Within a given time the interrogator has to find out which of the two the human is, and which one the machine.



The computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written response comes from a human or not.
Programming a computer to pass a rigorously applied test provides plenty to work on. The computer would need to possess the following capabilities to pass the Turing Test. They are :

- ➢ **Natural Language Processing**:
  Must be able to communicate successfully in English.
- ➢ **Knowledge representation:**
  To store what it knows and hears.
- ➢ **Automated reasoning:**
  Use stored information to answer the questions and draw new conclusion.
- ➢ **Machine learning:**
  Must be able to adapt in new circumstances and to detect and extrapolate patterns.

Turing tests deliberately avoid physical interaction with human interrogators because physical simulation of human beings is not necessary for testing intelligence
However, the **total Turing test** includes video signals and manipulation capability so that the interrogator can test the subject's perceptual abilities and object manipulation ability. To pass the total Turing test computer must have following additional capabilities:

- ➢ **Computer Vision:**
  To perceive objects
- ➢ **Robotics:**
  To manipulate objects and move

Chinese room argument

The Turing Test is one of the few things that comes to our mind when we hear about reasoning and consciousness in artificial intelligence. But apart from the Turing Test, there is one more thought process called the Chinese problem presented by **John Searle**, an American philosopher, at the AI researchers popularly known as ***Chinese Room Argument***.

The Chinese Room Argument aims to refute a certain conception of the role of computation in human cognition.

In order to understand the argument, it is necessary to see the distinction between Strong and Weak versions of Artificial Intelligence. According to Strong Artificial Intelligence, any system that implements the right computer program with the right inputs and outputs thereby has cognition in exactly the same literal sense that human beings have understanding, thought, memory, etc. The implemented computer program is sufficient for, because constitutive of, human cognition. Weak or Cautious Artificial Intelligence claims only that the computer is a useful tool in studying human cognition, as it is a useful tool in studying many scientific domains. Computer programs which simulate cognition will help us to understand cognition in the same way that computer programs which simulate biological processes or economic processes will help us understand those processes. The contrast is that according to Strong AI, the correct simulation really is a mind. According to Weak AI, the correct simulation is a model of the mind.

Strong AI is answered by a simple thought experiment. Searle explained the concept eloquently by drawing an analogy using Mandarin. The definition hinges on the thin line between actually having a mind and simulating a mind.

Searle's thought experiment goes like this:

Suppose a closed room has a non-Chinese speaker with a list of Mandarin characters and an instruction book. This book explains in detail the rules according to which the strings (sequences) of characters may be formed — but without giving the meaning of the characters. Suppose now that we pass to this man through a hole in the wall a sequence of Mandarin characters which he is to complete by following the rules he has learned. We may call the sequence passed to him from the outside a "question" and the completion an "answer." Now, this non-Chinese speaker masters this sequencing game so much that even a native Chinese person will not be able to spot any difference in the answers given by this man in an enclosed room.

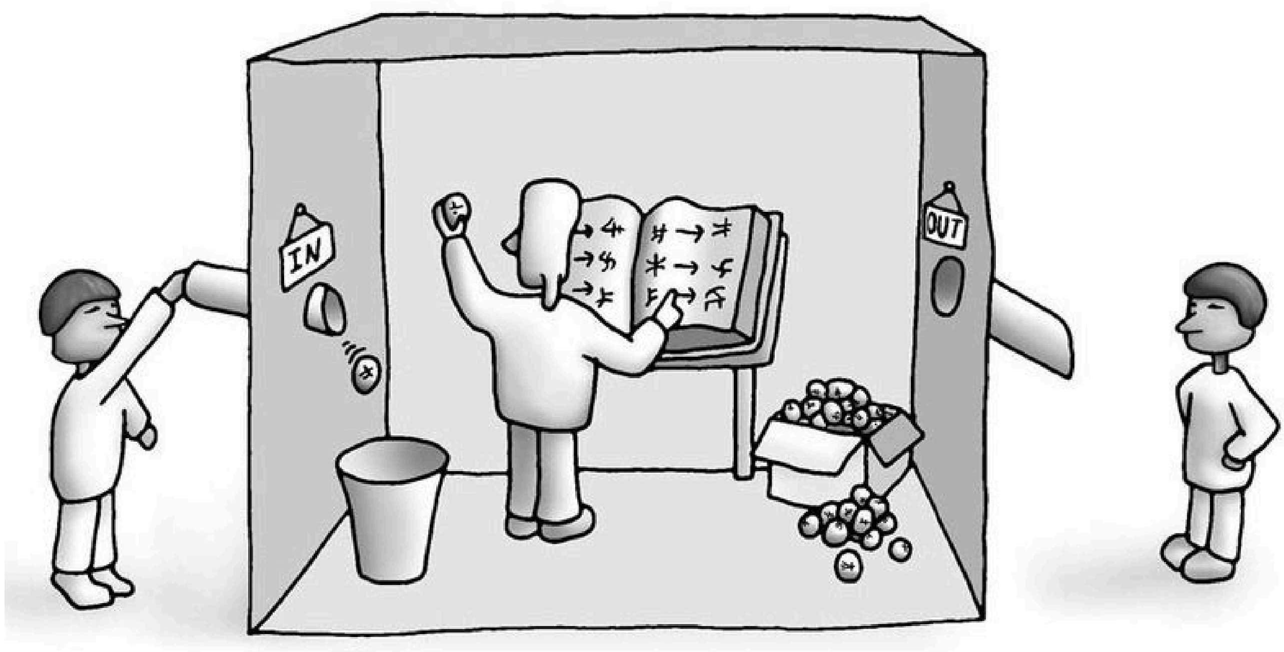
But the fact remains that not only is he not Chinese, but he does not even understand Chinese, far less think about it.

Now, the argument goes, a machine, even a Turing machine, is just like this man, in that it does nothing more than follow the rules given in an instruction book (the program). It does not understand the meaning of the questions given to it nor its own answers, and thus cannot be said to be thinking. Making a case for Searle, if we accept that a book has no mind of its own, we cannot then endow a computer with intelligence and remain consistent.

The whole point of Searle's experiment is to make a non-Chinese man simulate a native Chinese speaker in such a way that there wouldn't be any distinction between these two individuals.

If we ask the computer in our language if it understands us, it will say that it does, since it is imitating a clever student. This corresponds to talking to the man in the closed room in Chinese, and we cannot communicate with a computer in a way that would correspond to our talking to the man in English.

The texts or the set of instructions cannot be dissociated from the man in the experiment because this instruction, in turn, is prepared by some native Chinese person. So, when the Chinese expert on the other end of the room is verifying the answers, he actually is communicating with another mind which thinks in Chinese. So, when a computer responds to some tricky questions by a human, it can be concluded, in accordance with Searle, that we are communicating with the programmer, the person who gave the computer, a certain set of instructions to perform.



Thinking Humanly: The Cognitive Modeling Approach

If we are going to say that a given program thinks like a human, we must have some way of determining how humans think. We need to get inside the actual workings of human minds. There are three ways to do this:

- through introspection: catch our thoughts while they go by
- through psychological experiments.: observing the person in action
- Through brain imaging : observing the brain in action

Once we have a precise theory of mind, it is possible to express the theory as a computer program. If the program's input-output behavior matches the corresponding human behavior, that is the evidence that some of the program's mechanisms could also be operating in humans. For example, Allen Newell and Herbert Simon, who developed GPS(General Problem Solver), were not content merely to have their program solve problems correctly.They were more concerned with comparing the trace of its reasoning steps to traces of human subjects solving the same problem.

The field of cognitive science brings together computer models from AI and experimental techniques from psychology to try to construct precise and testable theories of the workings of the human mind. Therefore it is not possible to make machines that think like the human brain.

Thinking Rationally: The "Law of thought" Approach

The Greek philosopher Aristotle was one of the first to attempt to codify "Right Thinking", that is, irrefutable reasoning processes. His syllogisms provided patterns for argument structures that always yield correct conclusion when given correct premise

For example:

Ram is a man.
All men are mortal.
————————————————
Ram is mortal.

These laws of thought were supposed to govern the operation of mind: This study initiated the field of logic. The logicist tradition in AI hopes to create intelligent systems using logic programming. However there are two 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 knowledge is not 100% certain. Second, solving problems principally is different from doing it in practice. Even problems with certain dozens of facts may exhaust the computational resources of any computer unless it has some guidance as to which reasoning step to try first.

Acting Rationally: The rational Agent approach

Agent is something that acts. Computer programs do something, but Computer agent is expected to have following attributes:

- Autonomous control
- Perceiving their environment
- Persisting over a prolonged period of time
- Adapting to change
- Capable of taking on another's goal

Rational behavior: doing the right thing. Here, the right thing is the thing that is expected to maximize goal achievement, given the available information.
So, **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 given to correct inferences. Making correct inferences is sometimes part of being a rational agent, because one way to act rationally is to reason logically to the conclusion and act on that conclusion.
On the other hand, correct inference is not all of rationality, in some situations, there are probably no correct things to do, but something must still be done. There are also some ways of acting rationally that cannot be said to involve inference. For Example, recoiling from a hot stove is a reflex action that is usually more successful than a slower action taken after careful deliberation. Advantages:

- It is more general than the **laws of thought** approach, because correct inference is just one of several mechanisms for achieving rationality.
- It is more amenable to scientific development than are approaches based on human behavior or human thought because the standard of rationality is clearly defined and completely general.

Foundation of AI

Philosophy:
Logic, reasoning, mind as a physical system, foundations of learning, language and rationality.

- ➔ Where does knowledge come from?
- ➔ How does knowledge lead to action?
- ➔ How does mental mind arise from physical brain?
- ➔ Can formal rules be used to draw valid conclusions?

Mathematics:
Formal representation and proof algorithms, computation, undecidability, intractability, probability.

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

Psychology:
Adaptation, phenomena of perception and motor control.
➔ How do humans and animals think and act?

Economics:
Formal theory of rational decisions, game theory, operation research.
➔ 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 future?

Linguistics:
Knowledge representation, grammar
➔ How does language relate to thought?

Neuroscience:
Physical substrate for mental activities
➔ How do brains process information?

Control theory:
Homeostatic systems, stability, optimal agent design
➔ How can artifacts operate under their own control?

Characteristics of AI

Some of the characteristics of AI are:

- **Symbolic Reasoning:** reasoning about objects represented by symbols, and their properties and relationships, not just numerical calculations.
- **Knowledge:** General principles are stored in the program and used for reasoning about noble situations.
- **Search:** A ``weak method" for finding a solution to a problem when no direct method exists. Problem: combinatoric explosion of possibilities.
- **Flexible Control:** Direction of processing can be changed by changing facts in the environment.

History of AI

- **1943: Warren Mc Culloch and Walter Pitts**: a model of artificial boolean neurons to perform computations.
  - First steps toward connectionist computation and learning (Hebbian learning).
  - Marvin Minsky and Dann Edmonds (1951) constructed the first neural network computer
- **1950: Alan Turing's** "Computing Machinery and Intelligence" – First complete vision of AI.

- **The birth of AI (1956)**:
  - Dartmouth Workshop bringing together top minds on automata theory, neural nets and the study of intelligence. – Allen Newell and Herbert Simon: The logic theorist (first nonnumeric thinking program used for theorem proving).
  - For the next 20 years the field was dominated by these participants.

- **Great expectations (1952-1969):**
  - Newell and Simon introduced the General Problem Solver.
    - Imitation of human problem-solving
  - Arthur Samuel (1952-) investigated game playing (checkers) with great success
  - John McCarthy(1958-) :
    - Inventor of Lisp (second-oldest high-level language)
    - Logic oriented, Advice Taker (separation between knowledge and reasoning)
  - Marvin Minsky (1958 -)
    - Introduction of microworlds that appear to require intelligence to solve: e.g. blocksworld.
    - Anti-logic orientation, society of the mind

- **Collapse in AI research (1966 - 1973):**
  - Progress was slower than expected.
    - Unrealistic predictions.
  - Some systems lacked scalability.
    - Combinatorial explosion in search.
  - Fundamental limitations on techniques and representations.

        - Minsky and Papert (1969) Perceptrons.

- **AI revival through knowledge-based systems (1969-1970):**
    - General-purpose vs. domain specific
        - E.g. the DENDRAL project (Buchanan et al. 1969) First successful knowledge intensive system.
    - Expert systems
        - MYCIN to diagnose blood infections (Feigenbaum et al.) - Introduction of uncertainty in reasoning.
    - Increase in knowledge representation research.
        - Logic, frames, semantic nets, …

- **AI becomes an industry (1980 - present):**
    - R1 at DEC (McDermott, 1982)
    - Fifth generation project in Japan (1981)
    - American response

  Puts an end to the AI winter.

- **Connectionist revival (1986 - present): (Return of Neural Network):**
    - Parallel distributed processing (RumelHart and McClelland, 1986); backprop.

- **AI becomes a science (1987 - present):**
    - In speech recognition: hidden markov models
    - In neural networks
    - In uncertain reasoning and expert systems: Bayesian network formalism

- **The emergence of intelligent agents (1995 - present):**
    - The whole agent problem:

      "How does an agent act/behave embedded in real environments with continuous sensory inputs"

Application of AI

- Autonomous planning and scheduling
- Game playing
- Autonomous Control
- Expert Systems

- Logistics Planning
- Robotics
- Language understanding and problem solving
- Speech Recognition
- Computer Vision

Omniscience

Difference between AI and Omniscience

In Latin, Omnis means ‘all’ and Science means ”Knowing”. Omniscience is the state of possessing unlimited knowledge about all things possible. The religious aspects suggest omniscience is an attribute of God. For those who do not believe in God, the discussion might be applied to hypothetical powerful beings. It is supposed to mean complete knowledge. When we say a being knows everything, there are many kinds of knowledge the being could have. For example, knowledge may be factual or propositional. An omniscient agent knows the actual effects of its actions. An omniscient agent takes its decision based on an infinite knowledge base. There is nothing the agent does not know. Thus, it “knows” that it is making the correct decision.

AI, sometimes called machine intelligence, is a simulation of human intelligence demonstrated by machines, particularly computer systems. It is a branch of computer science that is concerned with the creation of intelligent machines that can think and learn like the way humans do. The idea is to get the machines to think for themselves and make decisions based on the data being fed. So the agent is called a rational agent. A rational system tries to get the best possible outcome given limited knowledge. A rational agent takes its action according to its percept sequence and instruction from the agent function. Thus, it “thinks” that it is making the correct decision.

For example: a rational agent may try to cross the street after perceiving that the road is empty. But while he is in the middle of the road, a metal box can fall from a plane flying overhead. He won’t have any idea about this. The omniscient agent would have perceived that too.

AI vs **Omniscience**

Comparison Chart

AI	Omniscience
AI is simulation of human intelligence demonstrated by machines, particularly computer systems.	Omniscience refers to the capacity of knowing unlimited knowledge of all things that can be known.
The idea is to get the machines to think for themselves and make decisions based on the data being fed.	Omniscience is the state of possessing unlimited or complete knowledge about all things possible.
AI is based on algorithms created by humans to help the machines think and learn.	It is an attribute given to the God alone because in reality, omniscience is impossible.