Artificial Intelligence (CSC261 - BScCSIT)

Instructor : Bikram Acharya

bacharyalive@gmail.com
9867222415

Reference Materials

Recommended Textbook:

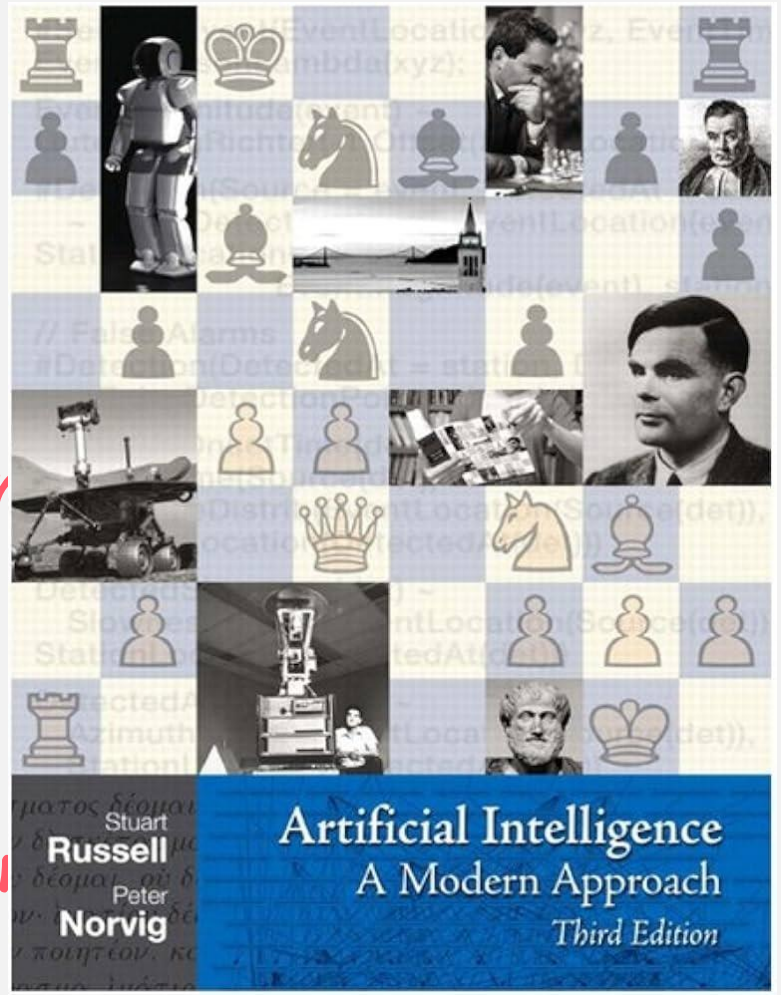
Artificial Intelligence: A Modern Approach
Stuart Russell and Peter Norvig

Alternate Book:

Computational Intelligence: A Logical Appr
Poole and Alan Mackworth

Online Course:

Artificial Intelligence by Patrick H. Winsto
https://www.youtube.com/watch?v=TjZBTDzGeGg



Evalution

	Theory	Practical	Total
Internal	20		20
Final Exam	60	20	80
Total	80	20	100

Theory

Attendance : 10%
Class Performance : 15%
Assignments : 25%
Assessments : 50%

Practical

Attendance : 20%
Class Performance : 30%
Lab reports and Program assignments : 50%

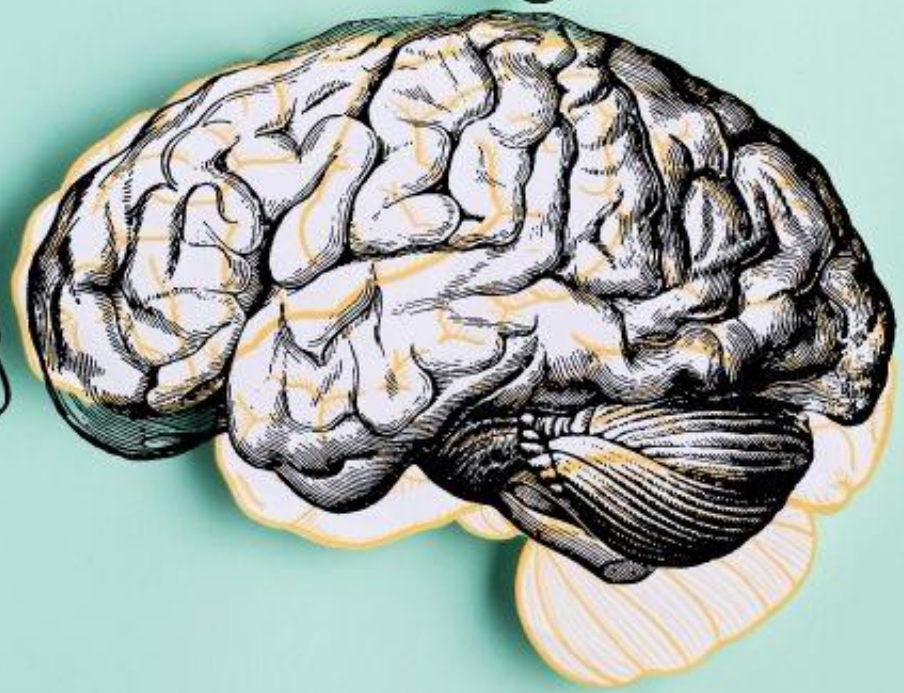
Syllabus

Unit I: Introduction (3 Hrs.)

1.1 Artificial Intelligence (AI), AI Prespectives: acting and thinking humanly, acting and thinking rationally
1.2 History of AI
1.3 Foundations of AI
1.4 Applications of AI

INTRODUCTION

What is Intelligence



What is intelligence?

- What do you think makes a person "intelligent"?

Some ideas

being able to do lots of **math** in your head

Knowing everything in the cyclopedia

being able to **memorize** lots of names and dates

having a **photographi** c memory

saying how to saywords in different languages

Can you think of any others?

Intelligence

- The study of intelligence here aims at making you able to explain how different views of intellegence influence teaching.
- Intelligence is the ability to learn from experience, solve problems, and use our knowledge to adpt to new situations.
- In summary, Intelligence is an inferred process that humans use to explain the different degrees of adaptive sucess in people's behaviour:
  - ✓ The mental abilities that enable one to adapt to, shape or select one;s environment
  - ✓ The ability to judge, comprehend, and reason
  - ✓ The ability to understand and deal with people, objects and symbols

Intelligence

- Intelligence is defined by experts as having three dimensions:
  - ✓ The ability to acquire knowledge
  - ✓ The capacity to think and reason in the abstract
  - ✓ The ability to see novel problems
- Theories of Intelligence
  - ✓ Spearman's Two-Factor Theory – suggests that people have general intelligence and specific intelligence
  - ✓ Thurstone's Theory of Primary Mental Abilities – suggests that eight separate factors make up intelligence
  - ✓ Gardner's Theory of Multiple Intelligence – believed that intelligence has a broader base and that people

Intelligence

Abstract concepts, mathematics, language, problem solving, memory, logical reasoning, planning ahead, emotions, morality, ability to learn/adapt, etc...

What Behaviors are Intelligent?

- Everyday Tasks : recognize a friend, recognize who is calling, tanslate from one language to another, interpret a photograph, talk, cook a dinner.

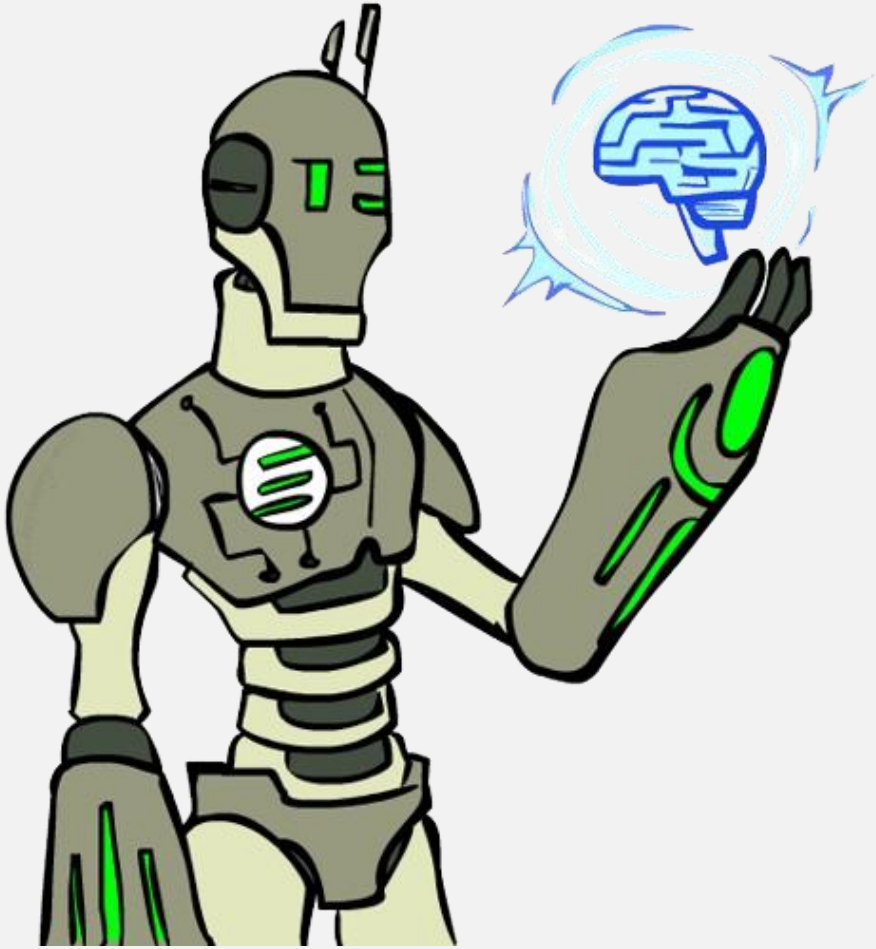
- Formal tasks : prove a logic theorem, geometry calculus, play chess, checkers or Go

- Expert tasks : engineering design, medical designers, financial analysis

ARTIFICIAL INTELLIGENCE

Artificial Intelligence

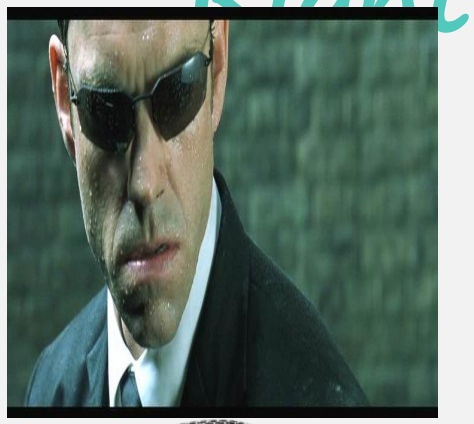
- What is Artificial Intelligence?
- Why study AI?
- What can AI do?

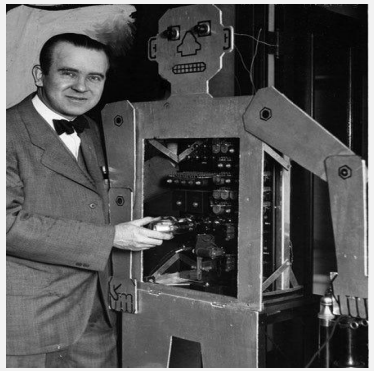


What is Artificial Intelligence?

Ans : The power of a machine to copy intelligent human behavior.

Right?





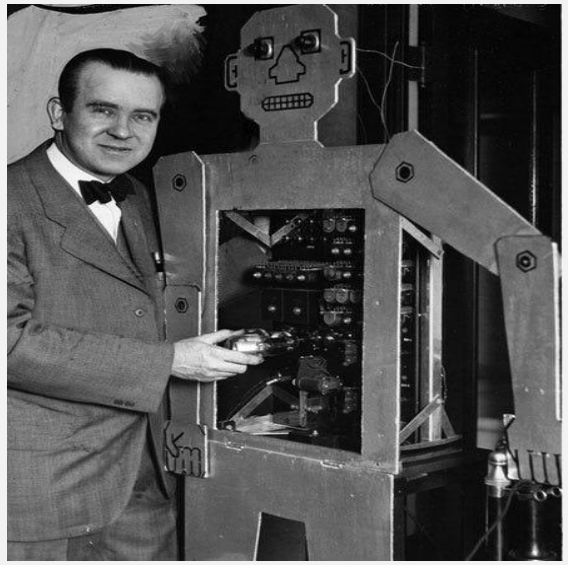
Initial Robots

SCI-FI AI

AL intergrated Robots

Are Robot really Intelligent?

Ans : A robot is only as smart as its initial program.



Artificial Intelligence VS. Robot

Artificial Intelligence	Robot
Programmed to think	Programmed to do
Social Interaction	Low level interaction
Learns	Only as smart as program

What About the Brain?

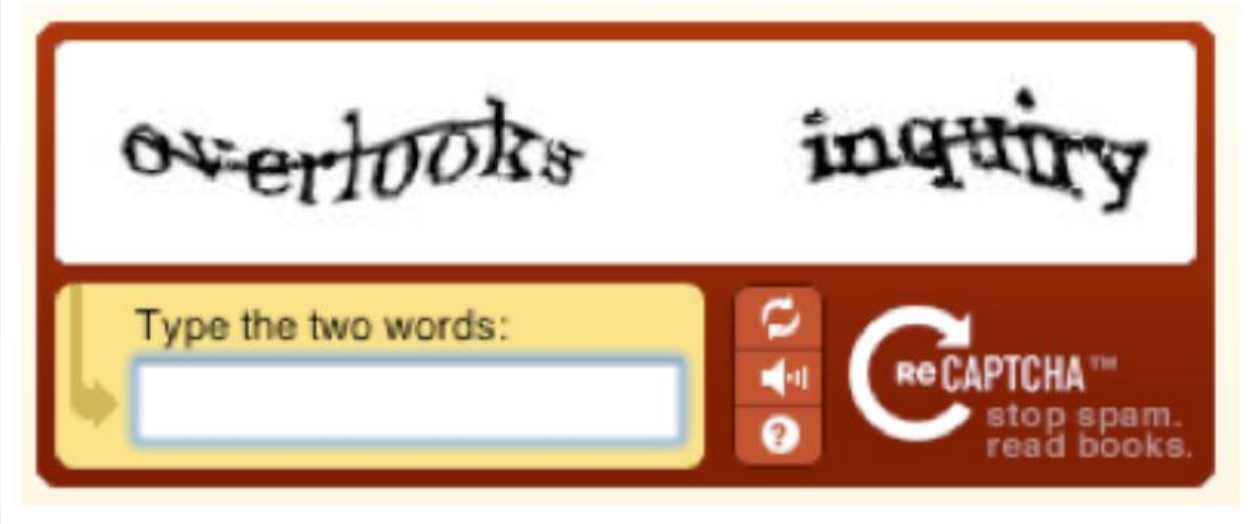
- Brains (human minds) are very good at making rational decisions, but not perfect
- Brains aren't as modular as software, so hard to reverse engineer!
- "Brains are to intelligence as wings are to flight"
- Human mind:
  - ✓ Can use common sense and past experience for problem solving.
  - ✓ Can understand Natural Language Processing.
  - ✓ Can learn from experiences and acquire knowledge.
  - ✓ Can reason about facts and deduce new facts.

Classical Test of (Human) Intelligence

- **The Turing Test:**
  - ✓ A human interrogator. Communicates with a hidden subject that is either a computer or a human.
  - ✓ If the human interrogator cannot reliably decide whether or not the subject is a computer, the computer is said to have passed the Turing Test.
  - ✓ Turing test compares the computer results to itself and not to humans.
- **Imitation Game:**
  - ✓ Involves two tests, first test involves an interrogator, a male & a female. Female pretends to be male &

Classical Test of (Human) Intelligence

- **Weak Turing type tests:**



CAPTCHA: Using Hard AI problems for Security

Human Intelligence

- Turing provided some very persuasive arguments that a system passing the Turing Test is intelligent.
  - ✓ We can only really say *it behaves like a human*
  - ✓ Nothing guarantees that it thinks lika a human
- Recent Trends:
  - ✓ Claims of development of AI system that can pass the Turing Test.
  - ✓ However, systems operate on subterfuge and were able to convince a rather naive jury that they were human like.
  - ✓ Main technique used: Obfuscation, i.e rather than answering questions the system changed the topic.

Human Intelligence

- In general there are various reasons why trying to mimic humans might not be the best approach to AI.
  - ✓ Computers and humans have a very different archictecture & different abilities
  - ✓ Numerical computations
  - ✓ Visual and sensory processing

	Computer	Human Brain
Computational Units	8 CPUs, 10^{10} gates	10^{11} neurons
Storage Units	10^{10} bits RAM, 10^{13} bits disk	10^{11} synapses
Cycle Time	10^{-9} disk	10^{-3} sec
Bandwidth	10^{10} bits/sec	10^{14} bits/sec
Memory updates/sec	10^{10}	10^{14}

Human Intelligence

- **Limitation:**
  - ✓ We know very little about how the human brain performs its higher level processes. Hence, this point of view provides very little information from which a scientific understanding of these processes can be built.
  - ✓ Humans might not be best comparison?
    - ➢ Don't always make the best decisions
    - ➢ Computer intelligence can aid in our decision making

Rational Decision

- The alternative approach relies on the notion of **rational decision**, **rationality**.
- **Rational** in a very specific way, technical way:
  - ✓ *Rational:* maximally achieving pre-defined goals
  - ✓ *Rationality* only concerns what decisions are made (not the thought process behind them)
  - ✓ Goals are expressed in terms of the *utility of outcomes*
  - ✓ Being rational means *maximizing your expected utility*

Rationality

- **Rationality** is a precise formal notion of what it means to do the right thing in any particular circumstance. Provides:
  - ✓ A precise mechanism for analyzing and understanding the properties of this ideal behavior we are trying to achieve.
  - ✓ A precise benchmark against which we can measure the behavior the system we build.
  - ✓ Formal characterizations of rationality have come from diverse areas like
    - ✓ logic (laws of thought) and
    - ✓ economics (utility theory—how best to act under uncertainty, game theory how self-interested agents

Artificial Intelligence

- Based on the above, `artificial intelligence' is about the science and engineering necessary to create artifacts that can:
  - ✓ acquire knowledge, i.e., can learn and extract knowledge; and
  - ✓ reason with knowledge (leading to doing tasks such as planning, explaining, diagnosing, acting rationally, etc.)
- Studies how to achieve intelligent behavior through computational means.
- Why do we think that intelligence can be captured through computation?
  - ✓ Modeling the processing that our brains do as computation has proved to be successful. Hence, human

Computational Intelligence

- AI tries to understand and model intelligence as a computational process.
- Thus, trying to construct systems whose computation achieves or approximates the desired notion of rationality.
  - ✓ Hence, AI is part of *Computer Science.*
    - ➢ Other areas interested in the study of intelligence lie in other areas or study, e.g., cognitive science which focuses on human intelligence.
    - ➢ Such areas are very related, but their central focus tends to be different.

Artificial Intelligence – Formal Definitions

Barr and Feigenbaum:

"Artificial Intelligence is the part of computer science concerned with designing intelligence computer systems, that is, systems that exhibit the characteristics we associate with intelligence in human behavior."

Elaine Rich:

"AI is the study of how to make computers do things at which, at the moment, people are better."

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Bellman, 1978

"[The automation of] activities that we associate with human thinking, activities such as decision making, problem solving, learning"

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Charniak & McDermott, 1985

"The study of mental faculties through the use of computational models"

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Dean et al.,
1995
"The design and study of computer programs that behave intelligently. These programs are constructed to perform as would a human or an animal whose behavior we consider intelligent"

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Haugeland, 1985

"The exciting new effort to make computers think machines with minds, in the full and literal sense"

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Kurzweil,
1990
"The art of creating machines that perform functions that require intelligence when performed by people"

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Luger & Stubblefield,
1993 "The branch of computer science that is concerned with the automation of intelligent behavior"

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Nilsson, 1998

"Many human mental activities such as writing computer programs, doing mathematics, engaging in common sense reasoning, understanding language, and even driving an automobile, are said to demand intelligence. We might say

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Rich & Knight,
1991
"The study of how to make computers do things at which, at the moment, people are better"

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Schalkoff, 1990

"A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes"

Four Definition by Russell+ Norvig

	Like Humans	Not Necessarily like Human
Think	Thinking Humanly – Cognitive modeling, System should solve problems the same way humans do	Thinking Rationally – Use of logic Need to worry about modeling uncertainty and dealing with complexity
Act	Acting Humanly – The Turing Test approach	Acting Rationally – The study of rational agents, agents that maximize the expected value of their performance measure given what they cureently know

Winston, 1992

"The study of the computations that make it possible to perceive, reason, and act"

Approach 1 – Thinking Humanly

The Turing Test Approach

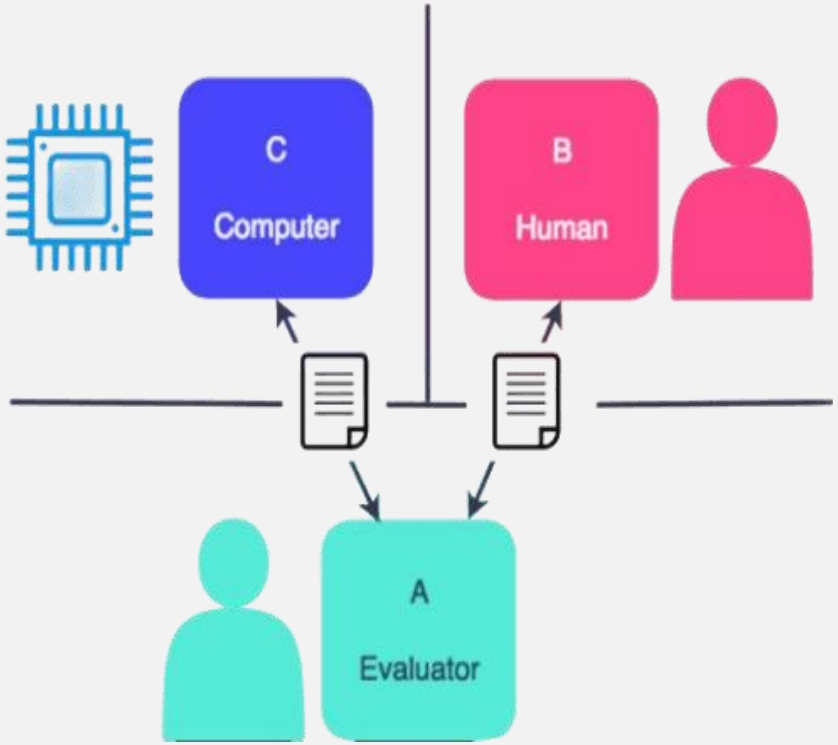
- Proposed by Turing (1950) "Computing machinery and Intelligence"
- Designed to provide a satisfactory operational definition of intelligence, ultimate test for acting humanly

"A computer passes the test if a human interrogator, afterposing some written questions, cannot tell whether the written responses come from a person or from a computer."

The Turing Test Approach

Setup

- Consider the following setup:
  - ✓ There are two rooms: "Room C" contains Computer and "Room B" contains Human
  - ✓ The interrogator/evaluator is outside on "Room A" and does not know which one is a computer.
  - ✓ Interrogator asks questions through a teletype and receives answers from both A and B.
  - ✓ The interrogator needs to identify whether A or B are humans.
  - ✓ To pass the Turing test, the machine



Turing Test Setup

The Turing Test Approach

Effectiveness of the Test

- Agents must:
  - ✓ Have strong command of language(Natural Language Processing)
  - ✓ Have wide range of knowledge(Knowledge Presentation)
  - ✓ Be able to reason(automated reasoning)
  - ✓ Be able to learn and adapt to new circumstances(machine learning)
  - ✓ Demonstrate human traits (humor, emotion)
- Turing test deliberately avoided direct physical interaction between the interrogator and the computer, because physical simulation of a person is unnecessary for intelligence.
- **Total Turing Test:**
  - ✓ Includes a video signal so that the interrogator can test the subject's perceptual abilities
  - ✓ As well as the opportunity for the interrogator to pass physical objects "through the hatch."
  - ✓ Agents must:
    - ➢ Be able to perceive objects(computer vision)
    - ➢ Be able to manipulate objects and move about(robotics)

Approach 1 – Thinking Humanly

Chinese Room Argument

- Searle (1999) summarized the Chinese Room Argument
- A thought experiment (gedanken experiment): main premises rest upon our intitutions about the experiment.
- Scenerio: For functionalism, and AI: The brain is a computer, the mind is a program.
  - ✓ Someone could write a program to enable a computer to simulate understanding of Chinese. It passes the Turing Test.
  - ✓ Does it understand chinese?
- Conclusion:
  - ✓ Main: Strong AI is false.
  - ✓ Also, Turing Test is faulty.(Functionalism is false)

Functionalism: mental states are defined as dispositional-causal relations between other mental states and input output.(Contrast with

Chinese Room Argument

Setup

Premise 1

- Assume there exists a program that understands Chinese, that is, passes the Turing Test for comprehension of Chinese:
  - ✓ The program's first inputs are a script and a story in chinese.
  - ✓ Then when questions in chinese about the story are input, the output is answers in chinese that indicate understanding of the story.

Premise 2

- A man in a room performs all of the operations of the program following an English translation of the program form 1, above:
  - ✓ He gives precisely the same output as the program for any input.

  [He is slower but that is supposed to be irrelevant]

Premise 3

- The man does not understand chinese. but by following the program, he convinces people in the room that someone in the room understands Chinese but he does not.

Chinese Room Argument

Conclusion:

- Since "the man is the appropriately programmed computer", Strong AI is false.
- Functionalism is false, since functionalism implies Strong AI.

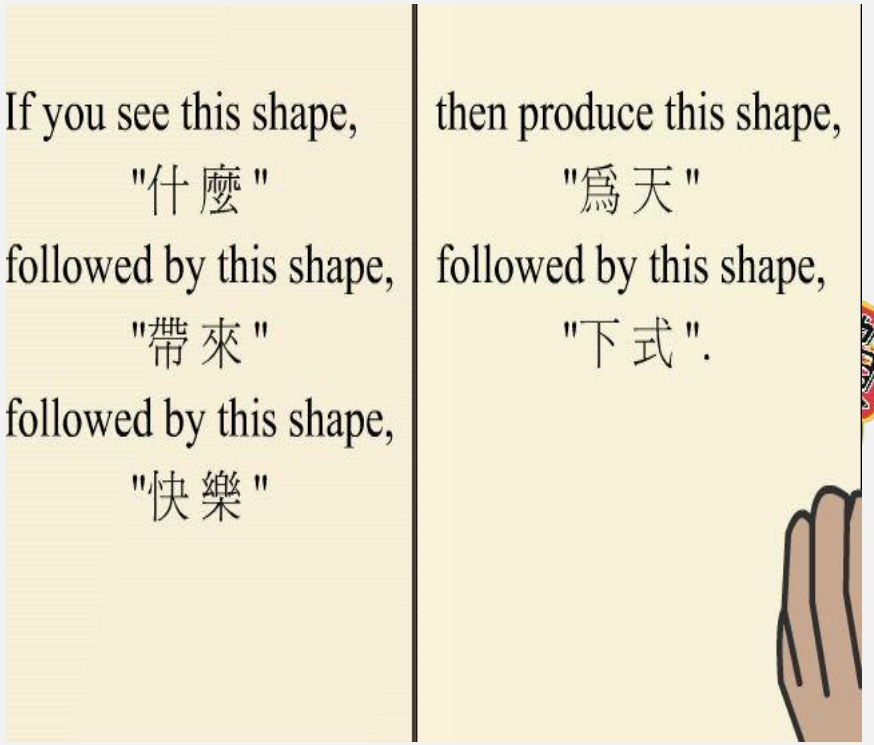
Searle's Explanation

- ✓ The computer has no intentionality, because semantics cannot be reduced to syntax.
- ✓ Intentionality is a causla power of brain matter.

Chinese Room Argument

Example

- Imagine that the brain is a computer and you cna get a printout of its program.
- What would such a program look like?
- Suppose you are given a printout of a Chinese person's brain.
- Now, suppose that room has two holes, one for people to drop in written questions(written in Chinese), and one for you to send out written answers to the questions(in Chinese).
- You send the new Chinese text out of the Answer slot. This text is an answer to the original question that had been put into the question



Chinese Room Argument

Example

- Common sense suggest that you do NOT understand the meaning of the questions or the answers.
- You don't even know that they were questions and answers.
  - It was just symbols that you transformed by hand working the brain's program.
- Searle says, the room has syntax, not semantics