Introduction

Aditya Challa

July 28, 2024

BITS Pilani
Pilani Campus

Course: Artificial and Computational Intelligence
Lecture No. 1

Course Logistics

- Classes
  - 16 lectures (2 hours each) - Time table can be found on the course website.
  - We will stick to the handout *religiously*. However, I will try to give some other (optional) references and material which can better illustrate the idea than the book.
  - Textbook: *Artificial Intelligence: A Modern Approach* by Stuart Russell and Peter Norvig - (2010 ed.)
  - 1 Quiz and 2 Assignments (Expect some coding in the assignments)

Course Logistics

- Course Objectives - By the end of this course,
  - You should be able to identify and understand the key concepts in AI
  - More importantly, you should be able to tell which of the many approaches is right for a given problem!

Introduction to AI

- Objective of today's lecture:
  - Understand what people mean when they say AI?
  - Diassociate AI from the popular media portrayal, especially the hype!
  - Some history and context for the course.
  - (if time permits) A brief overview of the course.

Introduction to AI

- The interest in obtaining Artificial Intelligence
  - *Intelligence* is the most valued possession of humans. Homo Sapiens stands for – *man the wise*
  - But what is AI???

Introduction to AI

- Four ways to think about AI

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

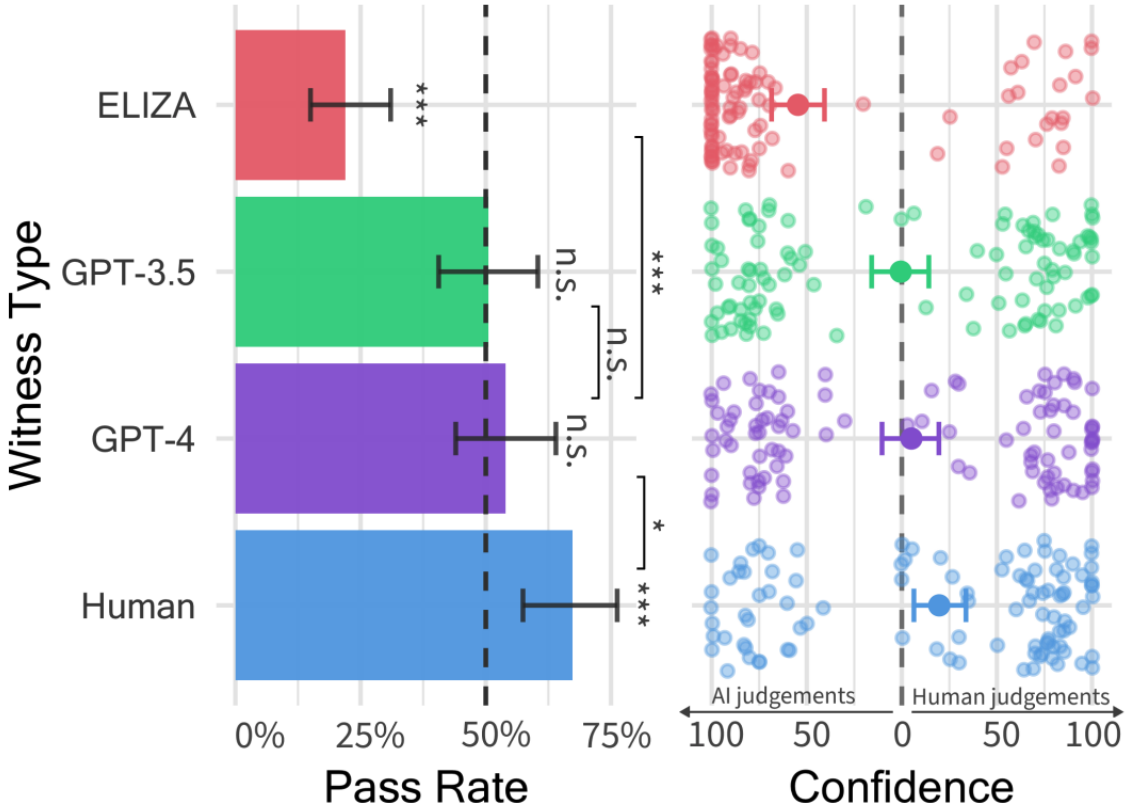
Figure 1.1 Some definitions of artificial intelligence, organized into four categories.

Introduction to AI

- Acting Humanly
  - Turing said - *If an observer cannot differentiate between a machine and person, then the machine can be considered Artificial Intelligence!*
  - This was quite difficult till recently! Deep learning, especially LLMs changed this!. `https://arxiv.org/pdf/2405.08007v1`

Introduction to AI

- Acting Humanly



Introduction to AI

- Acting Humanly - What does it take to pass the Turing Test:
  - Natural Language Processing - Understanding and generating human language (english, hindi, etc.)
  - Knowledge Representation - Representing knowledge in a way that machines can understand and importantly *store* - Question: How do you think humans store knowledge?
  - Automated Reasoning - Given a set of facts, can the machine derive new facts? - Reasoning is essentially using knowledge to derive new knowledge!
  - Machine Learning - Identifying patterns within the signals (data) and using them to make decisions.
- Total Turing Test - Interact physically!
  - Computer Vision - Understanding and interpreting visual data
  - Robotics - Interacting with the physical world (manipulating objects, etc.)

Introduction to AI

- Acting Humanly - Caveats of Turing Test
  - As the (in)famous joke goes - If a machine passes the turing test, it's might not be because it's intelligent, but because the humans are not!
  - But is *human intelligence* the only kind of intelligence? – What about animals, aliens, etc.?
  - Do we "flap" the wings of an airplane to make it fly like a birds?

Introduction to AI

- Thinking Humanly - Cognitive Science
  - This is the *cognitive modeling* approach - trying to model the human brain.
  - Consider reading - *Tell Tale Brain* by V.S. Ramachandran
  - We will not be focusing on this in this course.

Introduction to AI

- Thinking Rationally - Logic
  - Another way to understand intelligence is through *logical reasoning* a.k.a mathematics.
  - Symbolic AI - Representing knowledge in a logical form and using logical reasoning to derive new knowledge.
  - Our education system actually conflates intelligence with logical reasoning! But they are different!
  - Think birds/dogs - they are intelligent but not logical!

Introduction to AI

- Acting Rationally - Rational Agent
  - Rational Agents - Agents that act to achieve the best outcome or, when there is uncertainty, the best expected outcome - encompass most of the techniques in AI.
  - But then again rationality is not something which even humans exhibit all the time! - Consider reading *Thinking Fast and Slow* by Daniel Kahneman
  - This course concentrates on general principles of rational agents and on components for constructing them!

Foundations of AI

• Foundations of AI - Several disciplines contribute to our understanding of AI. Although, these disciplines are not *aimed* at solving the problem of AI, they provide the necessary tools and techniques.

- Philosophy - What is intelligence? What is knowledge? What is reasoning?
- Mathematics - Logic, Probability, Optimization, etc.
- Economics - Game Theory, Decision Theory, etc.
- Neuroscience - How does the brain work?
- Psycohology - How do humans think?
- Control Theory - How do we control systems?
- Computer Science - Algorithms, Data Structures, etc.
- Linguistics - How do we understand language?

Foundations of AI

- Philosophy - Mind vs Matter:
  - A brain in it's essence is a physical entity - a collection of atoms and molecules -¿ Neurons -¿ Synapses etc.
  - Philosophers have debated for centuries - Is the mind a separate entity from the body? - Dualism
  - (Subjective Opinion): The subject of chaos and emergence is the way I understand the mind-body problem.

Foundations of AI

- Philosophy - Game of Life
  - A simple cellular automata game - a grid of cells that can be in one of two states - alive or dead.
  - The rules are simple - a cell is born if it has exactly 3 neighbors, a cell dies if it has less than 2 or more than 3 neighbors.
  - The game is deterministic - the next state of the grid is completely determined by the current state.
  - But the game exhibits complex behavior - patterns that move, patterns that replicate, etc.

Foundations of AI

- Mathematics - Logic:
  - Logic, while not exactly AI itself, is a foundational tool for AI.
  - You have rules and inference, and you perform inference on specific objects – Example: All men are mortal, Socrates is a man, therefore Socrates is mortal.
  - *Inception effect:* But you can have rules that apply to rules themselves - Example: If A then apply rule 1, if B then apply rule 2, etc.

Foundations of AI

- Mathematics - Limits of Logic:
  - Gödel's Incompleteness Theorem - Any consistent formal system is incomplete. In simple words, there are statements that are true but cannot be proven. Example : This statement is False.
  - A similar result is shown by Turing about computability - there are problems that are not computable.
  - However, a more important practical aspect is the ability to solve it in reasonable time. If a solution takes the natural age of the universe to compute, it's not a *good*solution! – Tractability and NP-Completeness.

Foundations of AI

- Mathematics - Probability:
  - Probability is a way to model uncertainty.
  - This is at the heart of inductive logic - reasoning from specific instances to general rules, basically undelies all machine learning algorithms.

Foundations of AI

- Economics - Understanding Human Decisions:
  - Economics is a rich source on how humans make decisions, and so is a rich source of data for AI.
  - The concept of utility and how various humans interact is a subject of study – Game Theory. Example is the famous *Prisoner's Dilemma* and Nash Equilibrium.
  - Book - Thinking Fast and Slow by Daniel Kahneman - discusses a lot of cases where humans make irrational decisions.
  - Basically, humans seem to make decisions which are "good enough" rather than the "best" – *Satisficing*.

Foundations of AI

- Neuroscience - How does the brain work?
  - Brains are fascinating - Story of Phineas Gage!
  - Deep learning is inspired by the structure of the brain - neurons and synapses.

Foundations of AI

- Psychology - How do humans think?
  - behavioral psychology - Given a stimulus how do living organisms respond? (Pavlov's dog)
  - Cognitive psychology - How do humans think? (Memory, Reasoning, etc.)
  - Two books which give pretty good insights are - *Sapiens* by Yuval Noah Harari and *A Brief History of Intelligence* by Max Bennett.

Foundations of AI

• Control Theory - How do we control systems? More interestingly, How do we control ourselves?

- As yourself - How can you make an LLM control itself?
- The answer lies in understanding control theory and cybernetics - Basically well defined/designed feedback loops.
- It wasn't until RLHF finetuning that LLMs made an impact!

Foundations of AI

- Computer Science - Algorithms, Data Structures, etc.
  - None of the ideas are relevant if you cannot implement them!
  - Which means designing good architectures - CPU, GPU, TPU, etc. (or maybe something else!)
  - Good algorithms and data structures are the key to efficient computation to process and model the ideas. (Backpropagation, etc.)
  - Good Programming Languages to implement the ideas - Python, C++, etc.

History of AI

- 1943-55
  - Lack of computers during this time implied people relied heavily on theoretical work.
  - Questions they tried to solve - Model of a neuron and How much such a model can represent mathematical functions?
  - McCulloch and Pitts proposed a model of a neuron - a simple binary switch, and Minsky and Papert showed that a single layer of such neurons cannot represent all functions.
  - Turing proposed the Turing Test in 1950, and discussed the possibility of machines thinking.

History of AI

- 1956-70
  - The Dartmouth Conference in 1956 is considered the birth of AI.
  - Although not many breakthroughs happened during the conference, the field was established.
  - It's computer science, but has a lot of interdisciplinary components - philosophy, psychology, etc. (as discussed before!)

History of AI

- 1952-69
  - The first wave of AI.
  - Several ideas/tools were proposed - Logic Theorist, General Problem Solver, etc.
  - No matter which problem was picked, the researchers could make the computer solve it! - Ranging from geometry to analogy problems (typical in IQ tests)

History of AI

- 1966-73 - The problem with overexpectation and hype!
  - Story - When trying to translate russuian to english, the sentence "The spirit is willing but the flesh is weak" was translated to "The vodka is good but the meat is rotten"!
  - The hardware did not meet the pace of ideas - the computers were slow and memory was expensive. And intelligence could as well be a NP-Complete problem!
  - Neural Networks which were popular then did not have enough representational capacity.

History of AI

- Knowledge-based systems 1969-1979
  - One issue was - people were trying to start from scratch! - But humans don't start from scratch, they have a lot of knowledge from books!
  - DENDRAL incorporated knowledge about chemistry to solve problems in organic chemistry and identify molecule from mass spectrometry data.

History of AI

- Economic Value of AI (1980-present)
  - At the end of the day, irrespective of the curiosity of researchers, the technology has to create value.
  - An automated expert system helped people configure computers at Digital Equipment Corporation in 1986 - saved $40 million.

History of AI

- The scientific approach to AI (1987-present)
  - Recall Birds vs Airplanes - Why should it be that artifical intelligence be like human intelligence.
  - Hidden Markov Models, Bayesian Networks
  - Basically Emergence of Probability and Machine learning and statistics.
  - Still the main source of economic value!

History of AI

- Hardware, data and Deep Learning
  - Parallel Processing and other technologies helped propel the growth of neural networks. An interesting article - *The Hardware Lottery* by Sara Hooker
  - Moreover, since humans started living their life online, data was very easy to get and record.

History of AI

- The broad lessons from history:
  - Economics trumps always. Anything which is not feasible and does not reduce costs might not last. See `https://www.goldmansachs.com/intelligence/pages/gs-research/gen-ai-too-much-spend-too-little-benefit/report.pdf`
  - Humans have an amazing ability to adapt and accept - ChapGPT was a phenomenon in 2021, but after 3 years there are open source models Llama3.1 which can do the same thing.
  - Efficiency is important - Bayesian methods can theoretically solve every problem out there, but it's highly inefficient.
  - What you measure is what you get - Evaluation metrics and designing them are probably the most important value addition humans still make.

In this course?

- This course is designed to model and understand the rational agent perspective.
- We dive into Modelling Rational/Intelligent Agents. Four approaches to achieve this
  - Searching (especially A^* search) for various optimal solutions, for example game-playing etc.
  - Logical Agents - Use knowledge for inference
  - Bayesian Networks - Modelling randomness
  - Reinforcement Learning - Learning across time and changing environments.