UNIVERSITY OF Southampton

SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

FOUNDATIONS OF ARTIFICIAL INTELLIGENCE

Blocksworld Puzzle

Authors:
Subash Poudyal
28395253

Moderators:
Dr. Richard Watson
Enrico

Contents

UNIVERSITY OF
Southampton

1 Turing Test

The gist of the Turing test is to determine whether or not a machine (digital computer) is capable of thinking like a **human**. Turing presents it in a scenario he names "An Imitation Game"; human and a computer are required to have a conversation with a human solely judged by the means of text, if the judge was unable to differentiate, the computer has passed the Turing test showing that it is "intelligent".

In my personal opinion, in this current era, Turing test wouldn't be *the* test to test intelligence. Turing relies on the concept of behaviourism as it would only work with the belief that intelligence is based on behaviour and ability to answer simple questions proving to show "mental activity". Does one need to physically show behaviour to be intelligent? And are humans the only intelligent animals? Is AI just something that understands language? In addition, I noticed that Turing, although mentions it for the future of tests, never mentions how long the test is to be taken for.

Opinions about the test seem to vary as some still believe it is a significant test whereas some claim that it is a bottleneck or even hindrance to progress of Artificial Intelligence; raising the question: are we testing to imitate intelligence or attaining intelligence?

2 Searle's Chinese Room

Similar to the Turing test, Searle's Chinese Room is an experiment to test for "Artificial Intelligence". Searle divides the topic of AI into two: **strong AI** and **weak AI**. Strong AI being the "real" AI as it'd show to have intelligence rather than just an algorithm. Searle uses Chinese room as an analogy; imagining himself (who doesn't "understand a word of Chinese") in a room where he uses some form of manipulation algorithm to answer questions asked by the native Chinese speakers from outside the room. "I have inputs and outputs that are indistinguishable from those of the native Chinese speaker, and I can have any formal program you like, but I still understand nothing". Searle uses this argument throughout to argue what intelligence is and how artificial intelligence should be able to do.

In his paper, Searle replies to a number of critiques and their arguments some of which claim that the analogy of Chinese room focuses on the wrong idea. My personal favourite argument was 'The Other Minds Reply' which was very similar to what I was thinking while reading the paper; how we ***know*** other people understand Chinese is by their behaviour and if the computer can pass this, wouldn't this be considered as good as a native Chinese speaker? to this, Searle replies by saying that it's not ***how*** one knows others have cognitive states but rather ***what*** it is and how it can't just be a computational process without any cognition.

Chinese room seems to have initiated a lot of debate amongst the AI experts as there seems to be no conclusion or an agreement. Overall, it definitely helped play an important role in defining what intelligence, or even better, what an artificial intelligent agent is? (Decades later, we still debate on this topic based on our expectations and completions).

3 The Stanford Cart and the CMU Rover

In this paper, two autonomous vehicle are discussed: Cart and the CMU Rover. Cart was developed and functional whereas CMU Rover seemed to be just a prototype, "nearly operational"; an idea. And it seems, an inspiration to future development of autonomous vehicles.

Cart is described to be autonomous and reliable for short runs but slow, taking 10 to 15 minutes to move 1 meter and repeating the process of "thinking" and re-planning the route after the meter run, taking about 5 hours to complete a 20 meters long course. Cart relied solely on the on-board camera for it's vision.

Although Cart was functional and quite reliable, it's speed wasn't impressive. To improve this, the CMU Rover was designed. not only did the CMU have the camera on a tilt/pan/slide mount, it also featured short

UNIVERSITY OF
Southampton

range infrared and long range sonar proximity detectors with a dozen on-board processors for high speed decision making and communication.

Decades later, somewhat similar to CMU, autonomous vehicle manufacturer like Tesla use array of sensors and cameras and on-board processing system to navigate through obstacles effectively. Tesla uses 9 cameras, a radar with range of up to 160 meters and ultrasonic sensors with range of 8 meters [1]. This comes to show the advancement in technology since the development of Cart and CMU but also begs the question, will we ever reach the full autonomy when it comes to robots and vehicles? Is what we have enough or will the future require heavy modification to the technology we now call "revolutionary"?

4   Checkers Game

Samuel uses Checkers as a game to solve using "machine learning" as opposed to Go or Chess because of the complexity in rules (lack of) while still being considered as a "logical" game. For this, Samuel uses minimax to develop gameplay intelligence instead of brute force as it'd take "10^{21} centuries" to consider all the moves.

Minimax is a decision rules based algorithm for minimising the possible loss and maximising the potential gain; mostly used in a two player game / scenario where both parties know everything about the possible moves of the adversary.

Compared to current era, Samuel's approach was somwewhat like the AlphaGo Zero; using tree search to select the next move in a Go game. AphaGo uses Monte Carlo Tree Search (MCTS) as opposed to minimax. However because of computational advancements, rather than looking at 3 to 15 moves ahead (as Samuel's minimax did), AlphaGo, after 1,600 searches picks the move with the highest chance in winning the game [1][2].

Searle would say this is weak AI. And having come a long way in terms of computational power and algorithmic advancement, it begs the question, as the computers get more powerful, will the algorithms equally evolve? And are we still creating intelligent agents based on logical games?

5   Common Sense and the Mind of HAL

The hunt for a perfect AI has been going on for decades. And how we define a perfect AI depends on our definition of 'intelligence'. In the somewhat a review, by David G. Stork, he claims that AI should have common sense. Sure, it may be able to 'understand' language but does it really understand it the way we do? In other words, does it have intuition and the general common sense along with contextual information while communicating?

With advancement in technology, self driving cars are becoming more popular. As Stork argues, the car has no intuition, it doesn't know if a kid by the side of the road is more likely to jump in than an old man. This is something that humans have learned from experience and some may even say, the initial instinct and intuition is genetic but how does a machine learn these? Sure, it can learn from a knowledge base but we can't have database of everything; we humans have contextual information and prior knowledge to connect the dots. While AI like IBM Watson may play and win games such as Jeopardy, they don't really understand the language itself.[1][2]

To test this, Stork started a company that can understand language. With different models (top down and bottom up), the goal is to solve the problem of NLP. I wonder how much training dataset that will require for the computer to be understanding to a human level?

6   Examining the Society of Mind

This paper written by Push Singh examining a famous book by Marvin Minsky: Society of Mind. I personally found it quite strange to build a report on this paper.

Singh first explains the fundamental topics from Minsky's book such as *agents* and *frames* for collaborative agents. Then explores some familiar topics such as the CYC project in order to build a "common sense" or knowledge database and the Soar cognitive architecture; similarities and difference between them.

The goal is clear: understand and build cognitive machine. But how does one go about that? By building "agents" and "frames" could be one way. Maybe the "agents" Minsky describes are what "neurons" in an artificial neural net we use today?

Minsky's approach to cognition and the "mind" is interesting when compared with distributed cognition or "DCog". While Minsky seems to see Mind as it's own entity and try to understand it's working as a single big picture made up of smaller workings, DCog emphasizes the ways that cognition is off-loaded into the environment through social and technological means. It is a framework for studying cognition rather than a type of cognition.

7   Nouvelle AI

What is intelligence and what exactly are we trying to create when we say 'Artificial Intelligence'? Brooks in his paper, 'Intelligence Without Reason'[1] tries to summarise this. Rather than a top down approach where we think of AI as a computational problem, Brooks implies it's more useful if we see it as a bottom up approach where we first understand the neuroscience aspect of intelligence with the help of biological examples from nature.

Brooks argues that what we call 'AI' is based on old techniques (mainly search) which was then limited due to the computational architecture rather than anything else. And now that computers are more powerful, our techniques haven't gotten comparatively better. Brooks refers to these top down approach as Sense Model Plan Act (SMPA) model such as that mentioned in CART paper[2] and holds it responsible for the lack of speed and importance in intelligent agents.

In the hope of building decentralised biological like agents, Brooks proposes behaviour-based agents with specific characteristics: Situatedness, Embodiment, Intelligence and Emergence. Comparing it to what Turing or Searle would call 'AI', Brooks has quite a contrasting idea, one that may lead us to Artificial General Intelligence that companies like OpenAI[3] and Neuralink[4] are trying to achieve.

8   The Architecture of Mind

David E. Rumelhart, in his paper "The Architecture of Mind" [1] attempts to describe the workings of the brain. In the area of AI research, computers and computer architecture has always influenced our imaginative design of our brain. Although many researchers in his time still believed in the classical system, David (an example of 'Connectionists') believes in the neurally inspired models where there are a set of processing units instead of a central processor, and the knowledge is being stored in the connections themselves. These systems are good at parallel processing, like our brain. David quotes, "Our goal in short is to replace the computer metaphor with the brain metaphor".

Like R.A Brooks [2], David is focused on understanding the inner workings of the brain as well as the computer architecture required to achieve such goal. R.A.Brooks' research on such a model is based on a holistic approach. Their assumption is that "human level intelligence is too complex and too little understood to be correctly decomposed into the right sub-pieces at the moment, and that even if we knew the sub-pieces we still wouldn't know the right interfaces between them"[3]. Their proposal is that intelligent systems should be built incrementally, having complete systems at each step, just like as in the evolution of the biological intelligent systems.

Although research and pioneers like Minsky didn't agree with the idea of perceptrons as the way forward in development of AI, David along with Hinton and Williams (1986) believed they had developed a generalization

UNIVERSITY OF
Southampton

of the perceptron learning procedure called *generalized delta rule.*

It is fascinating to see how the current deep learning strategies and methods have come into play. From just an idea about simple perceptrons to current technologies such as RNNs and CNNs used by OpenAI[4] and Deep Mind[5] makes you question; what is the future of AI?

References

[1] "About tesla: Tesla." [Online]. Available: https://www.tesla.com/about