AI And Consciousness

Project submitted in the context of the course
I3331 Computer and society

Prepared by
Mohamad Kassem 91532
Mohammad Majed 89719

LEBANESE UNIVERSITY
FACULTY OF SCIENCES I
DEPARTMENT OF COMPUTER SCIENCES
2019 - 2020

Abstract

This Project starts by discussing and explaining what Artificial Intelligence is. Then we discuss in this project the possibility of AI machines getting to a state of consciousness and we show all the points of view; with and against. Furthermore, we discuss each and every idea in an argumentative form and we show the tests that humans have created to test for consciousness and their results and we discuss these results.

Acknowledgements

We have taken efforts in this research. However, it would not have been possible without the kind support and help of many individuals. We would like to extend our sincere thanks to all of them.

We highly indebted to Lebanese University Faculty of Science, Department of Computer Sciences and especially our Dr. Samar Tawbi for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the project.

Keywords:

Artificial intelligence, Consciousness, Self-Awareness, Machines, Robots.

1.Introduction:

Artificial intelligence is when machines can learn and make decisions similarly to humans. There are many types of artificial intelligence including machine learning, where instead of being programmed what to think, machines can observe, analyse and learn from data and mistakes just like our human brains can.

This technology is influencing consumer products and has led to significant breakthroughs in healthcare and physics as well as altered industries as diverse as manufacturing, finance and retail. In part due to the tremendous amount of data we generate every day and the computing power available, artificial intelligence has exploded in recent years. Many everyday influences of artificial intelligence are altering the way our daily lives look. If someone from the 1950s travelled through time and arrived in 2019, they would marvel at the way we use our smartphones to navigate around town, how virtual digital assistants such as Alexa and Cortana respond to our queries and would be baffled by our addiction to social media channels such as Facebook, Instagram and Twitter. What is now normal to us and powered by AI, would be utterly foreign to our friend from the past. There's no doubt that artificial intelligence is an integral part of our daily lives. Our financial institutions, legal institutions, media companies and insurance companies are all figuring out ways to use artificial intelligence to their advantage. From fraud detection to writing news storeys with natural language processing and reviewing law briefs, AI's reach is extensive.

We might still be years away from generalised AI—when a machine can do anything a human brain can do—However this raises the fear of the possibility that a machine can one day become conscious and self-aware. The power of artificial intelligence is "so incredible, it will change society in some very deep ways," said billionaire Microsoft co-founder Bill Gates, and According to Elon Musk, "cutting edge" AI is actually "far more dangerous than nukes." Therefore, today more experts are interested in understanding human consciousness.

Consciousness refers to your individual awareness of your unique thoughts, memories, feelings, sensations, and environments. Essentially, your consciousness is your awareness of yourself and the world around you. This awareness is subjective and unique to you. If you can describe something you are experiencing in words, then it is part of your consciousness. So, this raises the question, is it possible for Artificially intelligent machines to become conscious and self-aware? And how can we even test if a machine is conscious?[1]

2.The Main Issue:

Some scientists believe that it is possible for machines to become self-aware as they believe humans are just a more complex form of intelligence and that human feelings and the most unique qualities of humans can be justified biologically as they are all the types of chemicals inside our brains. Also, the fact that technology is growing exponentially especially the domain of artificial intelligence. However, others believe that it is impossible for a machine to be self-aware and have feelings. Because they see that feelings cannot be expressed as algorithms.

3. The Survey:

3.1 Why Research AI safety:

In the near term, the goal of keeping AI's impact on society beneficial motivates research in many areas, from economics and law to technical topics such as verification, validity, security and control. Whereas it may be little more than a minor nuisance if your laptop crashes or gets hacked, it becomes all the more important that an AI system does what you want it to do if it controls your car, your airplane, your pacemaker, your automated trading system or your power grid. Another short-term challenge is preventing a devastating arms race in lethal autonomous weapons.

In the long term, an important question is what will happen if the quest for strong AI succeeds and an AI system becomes better than humans at all cognitive tasks. As pointed out by I.J. Good in 1965, designing smarter AI systems is itself a cognitive task. Such a system could potentially undergo recursive self-improvement, triggering an intelligence explosion leaving human intellect far behind.[2]

By inventing revolutionary new technologies, such a superintelligence might help us eradicate war, disease, and poverty, and so the creation of strong AI might be the biggest event in human history. Some experts have expressed concern, though, that it might also be the last, unless we learn to align the goals of the AI with ours before it becomes superintelligent.

There are some who question whether strong AI will ever be achieved, and others who insist that the creation of superintelligent AI is guaranteed to be beneficial. At FLI we recognize both of these possibilities, but also recognize the potential for an artificial intelligence system to intentionally or unintentionally cause great harm. We believe research today will help us better prepare for and prevent such potentially negative consequences in the future, thus enjoying the benefits of AI while avoiding pitfalls.

There are two ways AI might become a risk to humanity:

The AI is programmed to do something devastating: Autonomous weapons are artificial intelligence systems that are programmed to kill. In the hands of the wrong person, these weapons could easily cause mass casualties. Moreover, an AI arms race could inadvertently lead to an AI war that also results in mass casualties. To avoid being thwarted by the enemy, these weapons would be designed to be extremely difficult to simply "turn off," so humans could plausibly lose control of such a situation. This risk is one that's present even with narrow AI, but grows as levels of AI intelligence and autonomy increase.

The AI is programmed to do something beneficial, but it develops a destructive method for achieving its goal: This can happen whenever we fail to fully align the AI's goals with ours, which

is strikingly difficult. If you ask an obedient intelligent car to take you to the airport as fast as possible, it might get you there chased by helicopters and covered in vomit, doing not what you wanted but literally what you asked for. If a superintelligent system is tasked with an ambitious geoengineering project, it might wreak havoc with our ecosystem as a side effect, and view human attempts to stop it as a threat to be met.

3.2 AI Can Be Conscious:

There is little doubt that our intelligence and our experiences are ineluctable consequences of the natural causal powers of our brain, rather than any supernatural ones. That premise has served science extremely well over the past few centuries as people explored the world. The three-pound, tofulike human brain is by far the most complex chunk of organized active matter in the known universe. But it has to obey the same physical laws as dogs, trees and stars. Nothing gets a free pass. We do not yet fully understand the brain's causal powers, but we experience them every day—one group of neurons is active while you are seeing colors, whereas the cells firing in another cortical neighborhood are associated with being in a jocular mood. When these neurons are stimulated by a neurosurgeon's electrode, the subject sees colors or erupts in laughter. Conversely, shutting down the brain during anesthesia eliminates these experiences.

Given these widely shared background assumptions, what will the evolution of true artificial intelligence imply about the possibility of artificial consciousness?

Contemplating this question, we inevitably come to a fork up ahead, leading to two fundamentally different destinations. The zeitgeist, as embodied in novels and movies such as Blade Runner, Her and Ex Machina, marches resolutely down the road toward the assumption that truly intelligent machines will be sentient; they will speak, reason, self-monitor and introspect. They are eo ipso conscious.

This path is epitomized most explicitly by the global neuronal workspace (GNW) theory, one of the dominant scientific theories of consciousness. The theory starts with the brain and infers that some of its peculiar architectural features are what gives rise to consciousness.

Its lineage can be traced back to the "blackboard architecture" of 1970s computer science, in which specialized programs accessed a shared repository of information, called the blackboard or central workspace. Psychologists postulated that such a processing resource exists in the brain and is central to

human cognition. Its capacity is small, so only a single percept, thought or memory occupies the workspace at any one time. New information competes with the old and displaces it.

Cognitive neuroscientist Stanislas Dehaene and molecular biologist Jean-Pierre Changeux, both at the Collège de France in Paris, mapped these ideas onto the architecture of the brain's cortex, the outermost layer of gray matter. Two highly folded cortical sheets, one on the left and one on the right, each the size and thickness of a 14-inch pizza, are crammed into the protective skull. Dehaene and Changeux postulated that the workspace is instantiated by a network of pyramidal (excitatory) neurons linked to far-flung cortical regions, in particular the prefrontal, parietotemporal and midline (cingulate) associative areas.

Much brain activity remains localized and therefore unconscious—for example, that of the module that controls where the eyes look, something of which we are almost completely oblivious, or that of the module that adjusts the posture of our bodies. But when activity in one or more regions exceeds a threshold—say, when someone is presented with an image of a Nutella jar—it triggers an ignition, a wave of neural excitation that spreads throughout the neuronal workspace, brain-wide. That signaling therefore becomes available to a host of subsidiary processes such as language, planning, reward circuits, access to long-term memory, and storage in a short-term memory buffer. The act of globally broadcasting this information is what renders it conscious. The inimitable experience of Nutella is constituted by pyramidal neurons contacting the brain's motor-planning region—issuing an instruction to grab a spoon to scoop out some of the hazelnut spread. Meanwhile other modules transmit the message to expect a reward in the form of a dopamine rush caused by Nutella's high fat and sugar content.

Conscious states arise from the way the workspace algorithm processes the relevant sensory inputs, motor outputs, and internal variables related to memory, motivation and expectation. Global processing is what consciousness is about. GNW theory fully embraces the contemporary mythos of the near-infinite powers of computation. Consciousness is just a clever hack away.

3.3 AI Can't Be Conscious:

Humans are not just a simple living, we are a very complicated combination of many things, we are not just a physical and mental organism, we consist of souls, feelings, past experiences… And a lot many more stuff beside consciousness and self-awareness, and all of these things, we may not be able to define in a scientific way, or apply them as an algorithm.

An Indian author in the domain of human consciousness was asked by the interviewer if the machines can one day become conscious like humans. He laughed and answered: "Consciousness is not an academic study, the very fact that we think we can study it is because we always mix between many concepts like, consciousness, learning, accumulation of memory... which most of them machines can achieve, but not human consciousness. Thus, machines will never be like a human with a soul.

Answering a similar question “What is the main difference between robots and human?”, doctor Justin Hart, from the university of British Columbia, argues that what robots can never have, which humans have, is “Phenomenological consciousness”, which is defined as the first hand experience of consciousness thought. For example, there is a big difference between actually experiencing a sunrise and having the visual cortex neurons firing in a way that represents a sun light. Thus, a robot could be trifle Philosophical Zombies, capable of emulating consciousness, but never truly processing it. The idea requires that there is something beyond the physical mechanisms of thought that experiences the sun raise which robots would lack.

Some other experts like Max Tegmark a machine learning professor have even went beyond that into new interesting interpretations. Tegmark in this book “Life 3.0: Being Human in The Age of Artificial Intelligence”, have wrote many of these interpretations. One of those is that for a certain AI program or robot, because their information can be backed-up and easily shared among other AI’s. All they would lose if they were destroyed is their physical body. All their experiences, memory and life experience would be saved at the cloud. This would probably mean that they wouldn’t have a strong mean of self or identity like we humans do. So, a group of AIs would likely look like a single organism with a shared mind, rather than a group of individuals. Thus, building consciousness for each one of them will not make it like a self-aware human rather than another contributor in the AI universal cloud. This interpretation has opened the door to an entire new set of questions about AI and consciousness. Because even if scientists’ tests and experiments today have reached a certain point of success to implement human consciousness in a way or another. The question is, will this experimental consciousness ever make robots become like humans? Tegmark demands that he have the answer for this question, “Robots will never reach a certain level of consciousness that make them humans”.[3]

3.4 Testing Consciousness:

Because we cannot easily define human consciousness and feelings, many testing experiments were created and applied to test the ability of AI to act like a human, instead of diving deep into explaining what is consciousness.

Here comes “The Turing Test” to be the first created test in the domain. The Turing test was developed by Alan Turing in 1950. It tests the machine's ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human. Turing proposed that a human evaluator would judge natural language conversations between a human and a machine designed to generate human-like responses. The evaluator would be aware that one of the two partners in conversation is a machine, and all participants would be separated from one another. The conversation would be limited to a text-only channel such as a computer keyboard and screen so the result would not depend on the machine's ability to render words as speech. If the evaluator cannot reliably tell the machine from the human, the machine is said to have passed the test. The test results do not depend on the machine's ability to give correct answers to questions, only how closely its answers resemble those a human would give.

The test was introduced by Turing in his 1950 paper, "Computing Machinery and Intelligence", while working at the University of Manchester (Turing, 1950; p. 460). It opens with the words: "I propose to consider the question, 'Can machines think?'" Because "thinking" is difficult to define, Turing chooses to "replace the question by another, which is closely related to it and is expressed in relatively unambiguous words." Turing describes the new form of the problem in terms of a three-person game called the "imitation game", in which an interrogator asks questions of a man and a woman in another room in order to determine the correct sex of the two players. Turing's new question is: "Are there imaginable digital computers which would do well in the imitation game?" This question, Turing believed, is one that can actually be answered. In the remainder of the paper, he argued against all the major objections to the proposition that "machines can think"

After Turing, many tests have appeared to study similar issues. One of the newest and most interesting ones was “The Silence Pill” or “The Dumbing Pill” Test. The idea of the test was taken from a classical famous induction puzzle called “The King's Wise Men Hat Puzzle”. The goal of the “Silence Pill” Test is to determine if a certain robot, knows when it is speaking. In another words the test was created to check if a certain robot can be aware of itself, through knowing its own voice.

“Let’s find some interesting philosophical problem, then engineer a robot that can solve that problem,” says John Sullins, a philosopher of technology at Sonoma State University.

4.Comparison and Analysis:

4.1 Creativity:

For the longest time in history, humans believed that machines are never capable of becoming creative or beating humans when it comes to creativity and intelligence. However, in the recent years we have seen a major improvement when it comes to AI machines which proves that once again, machines can get another important human quality and that is creativity.

A group of researchers from Google's artificial-intelligence subsidiary, DeepMind, published a paper in the journal Science that described an A.I. for playing games. While their system is general-purpose enough to work for many two-person games, the researchers had adapted it specifically for Go, chess, and shogi ("Japanese chess"); it was given no knowledge beyond the rules of each game. At first it made random moves. Then it started learning through self-play. Over the course of nine hours, the chess version of the program played forty-four million games against itself on a massive cluster of specialized Google hardware. After two hours, it began performing better than human players; after four, it was beating the best chess engine in the world.

The program, called AlphaZero, descends from AlphaGo, an A.I. that became known for defeating Lee Sedol, the world's best Go player, in March of 2016. Sedol's defeat was a stunning upset. In "AlphaGo," a documentary released earlier this year on Netflix, the filmmakers follow both the team that developed the A.I. and its human opponents, who have devoted their lives to the game. We watch as these humans experience the stages of a new kind of grief. At first, they don't see how they can lose to a machine: "I believe that human intuition is still too advanced for A.I. to have caught up," Sedol says, the day before his five-game match with AlphaGo. Then, when the machine starts winning, a kind of panic sets in. In one particularly poignant moment, Sedol, under pressure after having lost his first game, gets up from the table and, leaving his clock running, walks outside for a cigarette. He looks out over the rooftops of Seoul. (On the Internet, more than fifty million people were watching the match.) Meanwhile, the A.I., unaware that its opponent has gone anywhere, plays a move that commentators called creative, surprising, and beautiful. In the end, Sedol lost, 1-4. Before there could be acceptance, there was depression. "I want to apologize for being so powerless," he said in a press conference. Eventually, Sedol, along with the rest of the Go community, came to appreciate the machine. "I think this will bring a new paradigm to Go," he said. Fan Hui, the European champion, agreed. "Maybe it can show humans something we've never discovered. Maybe it's beautiful."

AlphaGo was a triumph for its creators, but still unsatisfying, because it depended so much on human Go expertise. The A.I. learned which moves it should make, in part, by trying to mimic world-class players. It also used a set of hand-coded heuristics to avoid the worst blunders when looking ahead in games. To the researchers building AlphaGo, this knowledge felt like a crutch. They set out to build a new version of the A.I. that learned on its own, as a "tabula rasa."

The result, AlphaGo Zero, detailed in a paper published in October, 2017, was so called because it had zero knowledge of Go beyond the rules. This new program was much less well-known; perhaps you can ask for the world's attention only so many times. But in a way it was the more remarkable achievement, one that no longer had much to do with Go at all. In fact, less than two months later, DeepMind published a preprint of a third paper, showing that the algorithm behind AlphaGo Zero could be generalized to any two-person, zero-sum game of perfect information (that is, a game in which there are no hidden elements, such as face-down cards in poker). DeepMind dropped the "Go" from the name and christened its new system AlphaZero. At its core was an algorithm so powerful that you could give it the rules of humanity's richest and most studied games and, later that day, it would become the best player there has ever been. Perhaps more surprising, this iteration of the system was also by far the simplest. [4]

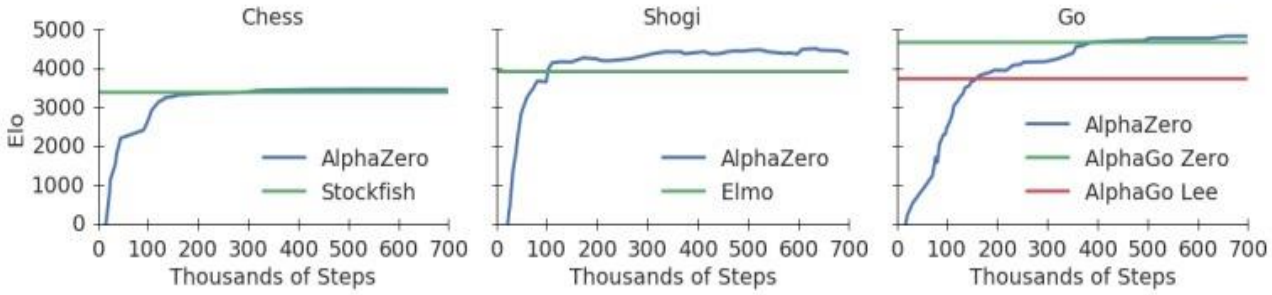

Figure 1: Training *AlphaZero* for 700,000 steps. Elo ratings were computed from evaluation games between different players when given one second per move. **a** Performance of *AlphaZero* in chess, compared to 2016 TCEC world-champion program *Stockfish*. **b** Performance of *AlphaZero* in shogi, compared to 2017 CSA world-champion program *Elmo*. **c** Performance of *AlphaZero* in Go, compared to *AlphaGo Lee* and *AlphaGo Zero* (20 block / 3 day) (*29*).

Figure 1 AlphaZero performance in Chess, Shogi and Go

4.2 Consciousness:

On the other hand consciousness is another unique human quality that many people believed and still believe that it is impossible for a machine to have.

Experts who are working on the domain of AI are trying to understand the terms “Conscious AI”. As we have showed before, many of these agrees with the idea that AI may one day get conscious just like humans. Some even believe that AI may even go beyond the idea of human self-awareness to achieve more complicated level of consciousness. And at the same time others believe that all of these topics about AI consciousness are just a science-fiction topic and that AI with never get conscious, because they see that human consciousness and self-awareness are very complicated terms that we can’t even define yet. And that we humans are not just a branch of neurons interacting together as other scientist's demand.

“A robots revolution is probably not coming anytime soon, and that’s because we still need to figure out what it means to be a human”. Say’s doctor Justin Hart

However, many experts of those who believe that AI might get conscious one day, are trying to simplify the concept of consciousness, in order to be able to experiment it, or at least do some experiments that could flatten the way for the future of conscious AI.

One of those is a robotics engineering professor at Columbia university, Hod Lipson who believes that AI may achieve things that are even much further than self-awareness, and they could redefine consciousness. But first, instead of trying to understand something as complicated as human consciousness, professor Hod and his team at Creative machines Lab at Colombia university are looking into something a little simpler, which they define as, the ability to image your self or what they call “Self Simulate”. Professor Hod defines Self Simulate by an example: “A dog might be able to simulate (imagine) itself into the afternoon, if it can see itself into the future, like if it can see itself having the next meal”. So with that definition they are trying to build that concept of Self Simulate as a beginning to achieve something like self-awareness in the future.

Also, there are many other examples that shows how AI must be upgraded step by step through a very long process before we talk about reaching a conscious AI. Here’s an example of human introspection that an AI must achieve. Suppose I ask you whether the President of the United States is standing, sitting or lying down at the moment, and suppose you answer that you don’t know. Suppose I then ask you to think harder about it, and you answer that no amount of thinking will help. A certain amount of introspection is required to give this answer, and robots will need a corresponding ability if they are to decide correctly whether to think more about a question or to seek the information they require externally.[5]

4.3 Testing Experiments

Throughout the centuries, experts have tried to develop ways and methods to test for consciousness in machines in order to refute the people who claimed to have developed a conscious machine, or maybe prove them right. One of these first tests and the most famous one is the Turing test.

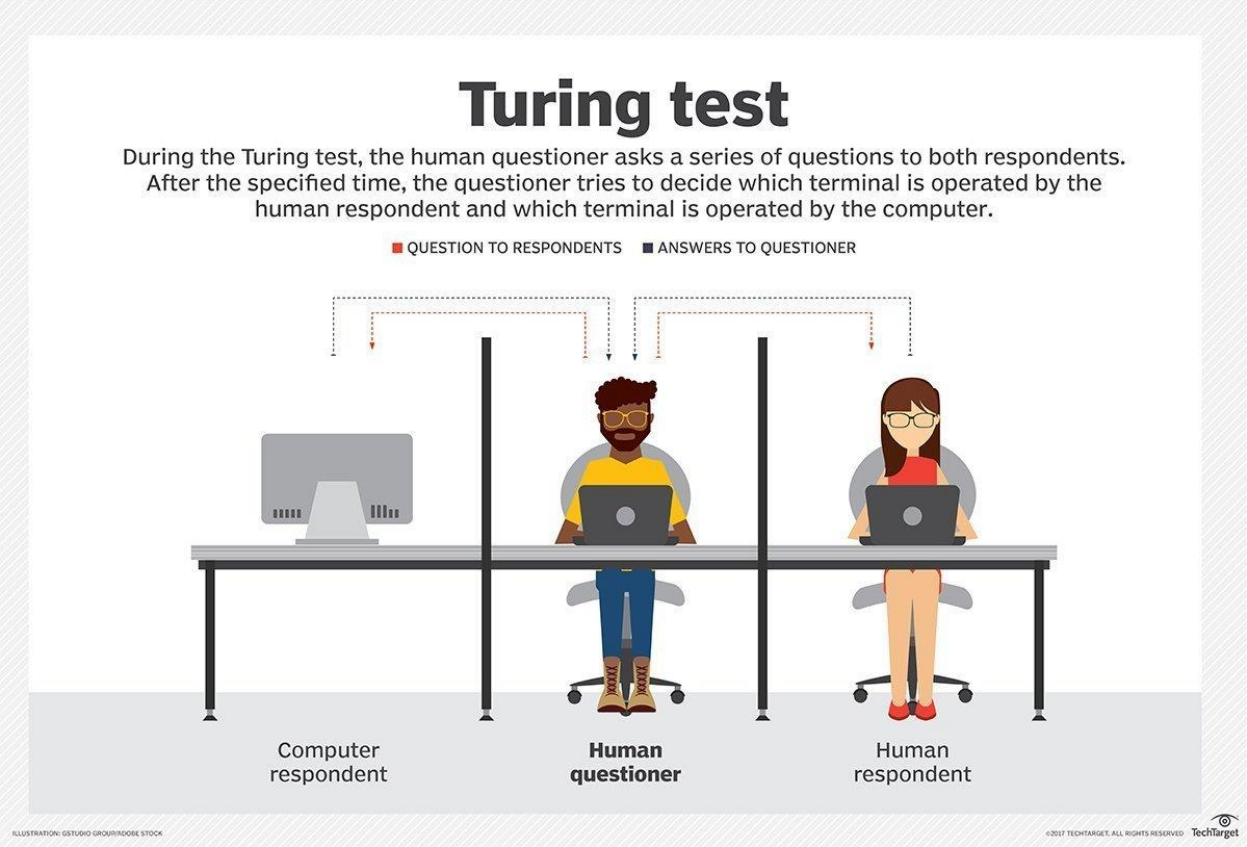

Figure 2 Turing Test

Turing Test is a method of inquiry in AI for determining whether or not a computer is capable of thinking like a human being. The test is named after Alan Turing, the founder of that Test. As well as, we have defined the Turing test in details before in section 3.4 (Testing Consciousness).

And “Figure 3” here shows how the Turing test can be applied as an experiment to test the ability of an AI to act like a human.

When Turing have first invented his test in 1950, it was hard for any computer to pass his test back then. Until 7 June 2014 where Turing test was passed for the first time by "Eugene Goostman" a chatbot that have won the competition at the Royal Society event organized by Keivn Warwick of the University of Reading. Goostman won after 33% of the judges were convinced that the bot was a human. 30 judges took part in the event. Each judge partook in a textual conversation with each of the five bots, at the same time, they also conversed with a human. In all, a total of 300 conversations were conducted, this made Goostman the first machine to pass a Turing test. In order to fake the judges, Goostman claimed that he is a 13 years old boy from Odessa, Ukraine who doesn't speak English that well, makes for a semi-convincing chatbot. His answers are at times enthusiastic and unintelligible like those from any normal 13-year-old would be; add in a shaky grasp of English, and there you go. [6]

However, there has been plenty of criticism about the value of the Turing test, as well as over the implementation of the University of Reading contest. "Chatbots like Eugene Goostman get away with changing the subject, not giving direct answers, and joking around," says Leora Morgenstern, chair of the executive committee at Commonsense Reasoning, an organization focused on enabling artificial intelligence systems to solve problems the way humans do. Turing assumes that an interrogator can determine if a machine is "thinking" by comparing its behavior with human behavior. Every element of this assumption has been questioned: the reliability of the interrogator's judgement and the value of comparing only behavior.

"What does fooling a human judge really have to do with intelligence?". Says Evan Ackerman.

Also, it requires that the machine have to be able to execute *all* human behaviors, regardless of whether they are intelligent. It even tests for behaviors that we may not consider intelligent at all, such as the susceptibility to insults, the temptation to lie or, simply, a high frequency of typing mistakes. If a machine cannot imitate these unintelligent behaviors in detail it fails the test. The example of Goostman and the other AI's that have passed the Turing test after, suggests that a machine passing the test may be able to simulate human conversational behavior by following a simple (but large) list of mechanical rules, without thinking or having a mind at all.

These Chatterbot programs have repeatedly fooled unsuspecting people into believing that they are communicating with human beings. In these cases, the "interrogator" is not even aware of the possibility that they are interacting with a computer. To successfully appear human, there is no need for the machine to have any intelligence whatsoever and only a superficial resemblance to human behavior is required. This leads Morgenstern to believes that there should be more effective ways of determining whether a computer program has the ability to demonstrate consciousness, or at least the ability to think and interact like humans. And he suggested that a better test would be: "seeing whether AI is able to use a large body of knowledge to correctly reason about an intentionally ambiguous statement". Such statements

are called “Winograd schemas”, named for Stanford computer science professor Terry Winograd, who first proposed their use.

But still, Morgenstern’s suggestion didn’t seem really convincing for many scientists. Ackerman comments on this suggestion saying: “Even if a computer program wins the Winograd Schema Challenge next year, saying that artificial intelligence has reached a human level would still be nearly as hasty as making the same claim for a chatbot that passes a Turing test. But the ability to reason and infer using a large database of commonsense knowledge has the potential to make software much better at both understanding what humans want and communicating with us as we do with one another”. [7]

Therefore, the experts today criticize the ability of such tests like the Turing test and the Winograd Schema Challenge to flatten the road for the AI to get conscious and self-aware in the future. But many believes that until the current day, one of the best experiments that was done in the domain is the “Silence Pill” experiment that was done in a robotics lab on the eastern bank of the Hudson River, New York. Where 3 small humanoid robots have a conundrum to solve. They are told that 2 of them have been given a “dumbing pill” that stops them from talking. In reality the push of a button has silenced them, but none of them knows which one is still able to speak. That’s what they have to work out.

Unable to solve the problem, the robots all attempt to say “I don’t know”. But only one of them makes any noise. Hearing its own robotic voice, it understands that it cannot have been silenced. “Sorry, I know now! I was able to prove that I was not given a dumbing pill,” it says. It then writes a formal mathematical proof and saves it to its memory to prove it has understood.

This was the first time a robot passes the test. It sounds like a simple test and it is, hardly scaling the foothills of consciousness. But showing that robots – in this case, off-the-shelf Nao models – can tackle logical puzzles requiring an element of self-awareness is an important step towards building machines that understand their place in the world.

Selmer Bringsjord of Rensselaer Polytechnic Institute in New York, who ran the test, says that by passing many tests of this kind – however narrow – robots will build up a repertoire of abilities that start to become useful. Instead of agonising over whether machines can ever be conscious like humans, he aims to demonstrate specific, limited examples of consciousness.[8]

Last but not least, because these types of tests only focus on a very little aspect of self-awareness at very specific situation. But these situations can’t be generalized so that we could say that it could actually happen to have a really fully self-aware AI just like a human. These tests until now don’t prove nor disprove what future AI would look like.

Conclusion and perspectives

Thus, consciousness in AI is a very complicated and debatable topic. Up to this date, there is no one true answer to whether or not it is possible. We have seen some evidence on AI beating humans in some traits we used to think only humans specialize in such as creativity, so theocratically everything can be possible. We saw how some robots were able to beat the Turing Test and other consciousness tests, but we also discussed how these tests could be beaten without conscious and how these tests contain some flaws in them. So, will we ever create a test that can truly give a correct answer to whether or not something is conscious? What if machines do become conscious one day, what will happen to humanity? Will machines stand by humans and help improve our world or will they get their own goals and become harmful to humans? Will the world be overtaken by robots such as the ones we see in science fiction movies? ...

References:

1- Article by Kendra Cherry, Reviewed by Amy Morin, LCSW

2- Article on https://futureoflife.org/

3- From Book: Life 3.0: Being Human in The Age of Artificial Intelligence written by the AI Professor Tegmark

4- Based on Research published by Google's DeepMind company

5- John McCarthy Computer Science Department Stanford University Stanford, CA 94305 jmc@cs, stanford, edu

6- "Interview with Eugene Goostman, the Fake Kid Who Passed the Turing Test" Article by professor "Doug Aamoth" on https://time.com/

7- "A Better Test Than Turing" Article by professor "Evan Ackerman", https://ieeexplore.ieee.org//abstract/document/6905475

8- "Robot homes in on consciousness by passing self-awareness test" Article by Doctor "Hal Hodson", on https://www.newscientist.com/