Thinking Machines

by

Christian Moises Cuevas Larin

602

Investigation Methodology

Instituto Mexico Americano Noroeste

Prof. Francisco Buenrostro

DATE

ACKNOWLEDGMENTS

I'd like to thank all of those who've formed who I am, whether they did good or bad; among those would be my family, friends and those who are nothing more than strangers to me now. I'd also like to thank all those "brilliant" minds that have contributed to humanity's progress, also to those who have detained progress or shown the worst of humanity for without the "ugly" there would be no "beauty".

INDEX

TOPICS, DELIMITATION AND HYPOTHESIS

Topics:

Ethics, Philosophical and Psychological

Delimitation:

Development of digital machines throughout history, and where they're headed.

Hypothesis:

Technology has advanced to a point where we can create digital machines after our brains; which is why I propose the question: can digital logical machines think? And have they the capability to decide for themselves?

JUSTIFICATION

Thinking Machines is a philosophical, psychological and ethical look at the development of digital logical machines(which we'll refer to as "Thinking Machines" in the text) throughout history, and to where they're heading. With the purpose of proving whether machines are thinking, and are capable of reaching consciousness, thus granting free will. It's only then that we'll concern ourselves with the implication that Thinking Machines will have on the way of life.

Before I state the importance of this topic I'll make the following clear. Both weak, also known as "narrow", AI and strong, also known as "general", AI are on different spectrum of capability and complexity; but both present equally threatening problems. Weak AI, grants Thinking Machines the capability to do what we see and don't see them do daily; for instance, filtering out spam sent to our email address or being our personal assistant. However, General AI, is on a whole nother spectrum compared to weak AI's capabilities, since it would allow Thinking Machines to be on par, and, likely , surpass us in most, if not all, aspects. To summarize, weak AI allows Thinking Machines to be good at any one given task it's designed for. While strong AI allows Thinking Machines the capability to accomplish any task it's faced without being explicitly designed or programed for such a task; we'll go into detail on how both types of AIs work later.

Thinking Machines, in my opinion, are the biggest threat, that's if not lead accordingly, to life as we know it; and ironically it's developed and created by mankind. I do not believe that Thinking Machines pose threats like those portrayed in media like in The Terminator or Matrix in which subjugation or extermination of the human race by Thinking Machines occurs, for those are exaggerated concepts. Instead, I worry about the real implications Thinking Machines have and will have on most, if not all, aspects of life, and how little thought or concern, if any at all, we pay towards the implications Thinking Machines have on life.

When, not if, Thinking Machines with AI become commonly wide spread, becoming on par or surpassing use of common machinery, then we'll know that anyone and everyone is replaceable by a Thinking Machine that surpasses us in most, if not all, aspect. While this might seem impossible, or at the least unlikely, we've been seeing such events occur all the way back from the Industrial Revolution. For instance, the most predominant event being machinery replacing us in the labor force, because, in most cases, it's due to the fact that machines provide an efficiency that we can never compete with. Now ask yourself this, what of us when Thinking Machines are capable of doing everything we can do, but at an efficiency we can never possible be on par with, nevertheless surpass?

Thinking Machines present a threat unlike any other threat mankind has ever seen; that's if we continue moving forward as we are. Meaning, if we were to discuss how mankind and Thinking Machines are to interact from now onward then, and only then, will we have a fighting chance to the imminent threat they pose. However, the common man doesn't believe that he will ever be replaced, it's only when confronted with such an event that he panics. Luckily, the replacing of man, so far, has been subtle to not have every man franting for his safety. It's only when Thinking Machines become on par with us that every man will panic for his safety, not in fear of dying, but in fear of being replaced, or worse in fear of not being unique, or as some would say "special". Which is why I fear for the leaders of tomorrow, whoever they might be, for if the leaders of today don't begin the talk on Thinking Machines then many of us are doomed to falter as if in the Industrial Revolution; as Benjamin Franklin said, "*By failing to prepare, you're preparing to fail.*"

In trying to start the conversation on Thinking Machines I've created this investigation, to prove to the few, if any, readers that Thinking Machines are truly thinking. So, "*Can machines think?*", Alan Mathison Turing, father of modern computing,

proposed this question in his book, "*Computing Machinery and Intelligence*" publish in the 1950's. He did not ask this question on the capabilities of digital machinery at the time, since at the time there was only a handful of digital machines, but instead asked if any imaginable digital machine is capable of thinking? His work is the foundation for various branches of computing including AI development; like, Machine Learning a field of computer science that gives computer systems the ability to "learn" with data, without being explicitly programmed. He also created the famous "*Turing Test*" which test the capabilities of a Thinking Machine, we'll go more into detail later. Now I propose the same question, "*Can machines think?*", to be more specific, "*Can digital logical machines think?*", and unlike A.M. Turing I'm witness to things only before seen in science fiction like self driving cars or A.I. capable of natural language.

As of now AI has far surpassed our expectations we've reached feats that experts in the field didn't believe to be solved until centuries from now, and it's these feats that have convinced me that Thinking Machine's capability for thinking is entwined with the development of AI. While we still have a long road ahead of us before we reach strong AI, that doesn't necessarily mean our current Thinking Machines haven't the capability of thinking as A.M. Turing said, "*Just because someone or something thinks differently than you do, it doesn't mean it 's not thinking*".

So, what's the importance of proving Thinking Machine's capability for thinking? Well as I've mentioned before we must start discussing the way in which Thinking Machines and humankind will interact. In doing so I believe we'll soften the blow dealt by Thinking Machine's capability to being on par with us, and likely surpass us. So now I ask you this my dear reader; how might we initiate the conversation? I propose to prove Thinking Machine's capability for thinking, the question this thesis is set out to answer, for it's only then that we'll have real conversation, and not a debate on whether Thinking Machines are or aren't capable.

Once the masses accept the capabilities of Thinking Machines they'll need to ask themselves several questions for instance, let's play with the idea that one day Thinking Machines will revolt and be the death of us. How then are we to prevent such a thing? Progress can't be stopped, at most it can only be delayed, and today is prove such a thing is true; therefore one day the perfect automated killing Thinking Machine will be created. So, if we can't prevent the creation of the perfect killing machine, nor can we prevent Thinking Machines from surpassing us; then what can we do? We'll, in my opinion, that's when should look towards Thinking Machines from an ethical view, in other words, we need to evolve machine ethics, we'll go further into detail later, to the point where Thinking Machines have a moral compass that's even more just than ours.

Let's say that machine ethics isn't enough to prevent a Thinking Machine from harming a human, or another Thinking Machine, then what is to occur? Regardless of what your thoughts and believes on Thinking Machines are let's say that Thinking Machines have the capability of consciousness and free will, then should a Thinking Machine be tried for its actions? And if so how will it be tried? By a jury of humans or a jury of its own peers, meaning other Thinking Machines? These are but a fraction of the questions that we should and will set out to answer.

INTRODUCTION

Pythagoras of Samos, a Greek philosopher, and his supporters/students, were known as Pythagoreans, and they believed that the essence of everything in life could be found in numbers and mathematics. (Ancient Origins)

We can prove this by a infinity number of examples, I'd like to return to the oldest of examples. The abacus is the earlies tool, that we know of, that resembles that of our modern Thinking Machines. With it our ancestors calculated and resprested countless of things the living & dead, human & animal, gold & food and soo much more can be represeted by numbers and this tool assitecd un such computation.

Computation, the action of mathematical calculation.

What is the start of thought? It's only when we ask ourselves this and then remain is complete silence that we can learn where this starts.

The capability of thinking is unique to the human?

Thinking Machines are all those things that we interact with on a daily bases from our phone to our monitors and calculators. So when I ask the question do thinking machines have the capability to think? I'm surely not talking about all thinking machines but for the most part they all can think.

The origin of Thinking Machines goes all the way back to the

We must ask ourselves why?

The need for computation comes with live for without the ability to calculate we would not be a functional being.

It’s undeniable that the need of computation is like that of the need to breath or eat; without it we would be nothing more then dead.

Can machines think?

“*Can machines think*?”, in order to answer such a question we must first define what is meant by a “*machine*” and “*thinking*”. As A.M. Turing elaborated,

if the meaning to these words where derived from the layman's vocabulary then the answer to our question can be answered in a polling.

“If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer the the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll.”[1]

“Is it possible for computing machines to think? No — if one defines thinking as an activity peculiarly and exclusively human. Any such behavior in machines, therefore, would have to be called thinking-like behavior. No— if one postulates that there is something in the essence of thinking which is inscrutable, mysterious, mystical. Yes — if one admits that the question is to be answered by experiment and observation, comparing the behavior of the computer with that behavior of human beings to which the term "thinking" is generally applied.”[8]

Thinking, to analyze and reflect over one's surroundings or oneself.

~ Thinking Machines ~

What does it mean to think? The simple act of analyzing and reflecting on thought itself is thinking. But how can such a thing occur? How can one things possibly lead to its own analyzation and interpretation.

Machinery has been in the grace of humanity long before the modern age; and computers have been in the grace of humanity for long aswell. It's been here for soo long that it's meaning has changed throughout history. It's only recently that computers have been otherwise with the machinery we know as digital machines.

But what is it that drove humanity to create these machines, computers and digital machines? If we go back to any point in history we can tell that the living have always wanted a better form of life; and in reaching such a goal they've created machinery to ease and simplify their work.

But as time went on life's necessities keeped on growing and new dependencies grew, the biggest of them today being electricity. With the growth of our necessities and ways to fulfill these necesites.

Can logical digital machines think? To answer such a question we must first define what a logical digital machine is then we must define what thinking means. To use the layman's term to define both of these would be outrages, instead we must

Saying all Thinking Machines think is an outrageous claim, like claiming all humans can see! Sure, all humans should have the capacity to see but not all do, the same would apply for Thinking Machines all should have the capacity for thinking but not all do. For the most past the commoners smartphone lacks the capacity to hold something as complex as weak AI much more strong AI.

I

THINKING

"*Just because someone or something thinks differently than you do, it doesn't mean it 's not thinking.*" — A.M. Turing

I I

THE PROBABILITIES OF THINKING

“We are unlikely to someday have robots that decide to turn on us, defeat us, make us their slaves, or exterminate us; and just as unlikely to have them befriend us or show us love without being specifically prompted by instructions to do so. This is because such intentional behavior from an A.I. would undoubtedly require a mind, as intentionality can only arise when something possesses its own beliefs, desires, and motivations.”[2], one with little experience or faith in machines would argue.

Is the mind limited to us? Can we not adapt machines to think as we do therefore giving them our mind? Sure, at the moment we lack the understanding of the thing we are. Yet, we hold on tight to our theories as if they were the truth. With those theories we give machines systems modeled after our own minds, and still we say they can’t think. What is it that limits thinking to us, is it perhaps that we’re living?

What is a “*mind*”? Merriam-Webster dictionary defines it as, “the element or complex of elements in an individual that feels, perceives, thinks, wills, and especially reasons”, a definition composed of words which have an almost endless amount of interpretation, words like ” feelings”, ” thoughts” and “will”.

We live by our will, and our will, in most cases, drives our intent. So how can a machine have a will much more an intent? After all, aren’t those all human aspects that require a consciousness to achieve such a feat? Well, like many of us, our intent is discovery and understanding of all that surrounds us, and that’s just about what machines are being trained to do.

I I I

TRAINING MACHINES

As I've mentioned before we've reached a point in which we don't program every action a machine takes, but instead we "train" them. The field of machine learning has verily started, and yet our algorithms, which we'd have to manually tinker with before, have evolved into something like our own neural networks which form our mind. Machines capable of automating decision making, and much more impressive being able to "think" of why and why not a decision is most appropriate.

Machine learning is a vast and ever growing field which started around the 1950. Today, companies worth billions like Google, Microsoft and Facebook have dedicated teams to the research and development of the field. In the past few years astonishing feats have been accomplished from Siri, a personal assistant reachable from your phone, to AlphaGo, beating top human players in the game of Go which has 2.08×10^{170} moves a feat ahead of its time by decades.

Siri and AlphaGo are two AIs that are worlds apart, both serving a different function. Yet they are one in the same as they're both "Thinking Machines", and their AI would be classified as weak or narrow. For even when they accomplish feats centuries ahead of predictions they are still simple compared to a strong or general AI. There are major differences between a weak/narrow and strong/general AI, but nevertheless they are both Thinking Machines.

Weak/Narrow AI, according to the Techopedia dictionary, it's a form of AI specifically designed to be focused on a narrow task and seem very intelligent at it. Has never been viewed as general intelligence, but rather a construct designed to be intelligent in the narrow task that it is assigned to. Techopedia gave a very good example of a weak AI that being Apple's Siri, which has the Internet behind it serving as

a powerful database. Siri seems very intelligent, as it is able to hold a conversation with actual people, even giving snide remarks and a few jokes, but actually operates in a very narrow, predefined manner. However, the "narrowness" of its function can be evidenced by its inaccurate results when it is engaged in conversations that it is not programmed to respond to.

Strong/General AI, is a form of AI designed with general intelligence allowing the machine to successfully perform any intellectual task that a human being can do, such as: reason, strategy, puzzle solving, judgment, plan, learn, communicate and integrate everything into a common goal. However we have many hurdles before we can achieve such a feat:

1. To autonomously and interactively acquire new knowledge and skills, in real time. This includes one-shot learning — i.e. learning something new from a single example.[7]

2. To truly understand language, have meaningful conversation, and be able to reason contextually, logically and abstractly. Moreover, it must be able to explain its conclusions![7]

3. To remember recent events and interactions (short-term memory), and to understand the context and purpose of actions, including those of other actors (theory of mind).[7]

4. To proactively use existing knowledge and skills to accelerate learning (transfer learning).[7]

5. To generalize existing knowledge by forming abstractions and ontologies (knowledge hierarchies).[7]

6. To dynamically manage multiple, potentially conflicting goals and priorities, and to select the appropriate input stimuli and to focus on relevant tasks (focus and selection).[7]

7. To recognize and appropriately respond to human emotions (have EQ, emotional intelligence), as well as to take its own cognitive states — such as surprise, uncertainty or confusion — into account (introspection).[7]

8. Crucially, to be able to do all of the above with limited knowledge, computational power, and time. For example, when confronted with a new situation in the real world, one cannot afford to wait to re-train a massive neural network over several days on a specialized supercomputer.[7]

While the list of hurdles seems long and tedious do not mistake it for a list of impossible task. For most of the task, if not all of them, are intertwined with humanity's prosperity, meaning, if all of humanity is to prevail it must overcome these hurdles.

Just as you and I were educated by a school, AIs are also educated. While we have different types of schools such as public, private, homeschool and etc. AIs have different learning methods with different traits, benefits and disadvantages. Schools to us, are like learning methods to AIs.

Supervised Learning, the algorithm is taught or trained from data which is already labeled with the correct answer. The larger the data set the more the algorithm will be able to generalize in a more precise way. Once the training is completed, new data is provided, without the correct response labels, and the learning algorithm uses

the past experience acquired during the training stage to predict a result.[4]

Unsupervised Learning, the algorithm is trained using a data set that has no label; in this case, the algorithm is never told what the data represents. The idea is that the algorithm can find by itself only patterns that help to understand the data set.[4]

Reinforced Learning, the algorithm learns by observing the world around it. Your input information is the feedback or feedback you get from the outside world in response to your actions. Therefore, the system learns based on trial and error. A good example of this type of learning can be found in games, where we try new strategies and we select and perfect those that help us win the game. As we gain more practice, the cumulative effect of reinforcement on our victorious actions will end up creating a winning strategy.[4]

Deep Learning, is a particular set of machine learning algorithms that use deep structures of neural networks to find patterns in the data. These types of algorithms currently have a great interest, since they have proven to be extremely successful in solving certain types of problems; as for example, the recognition of images. Many consider that these types of models are those that in the future will lead us to solve definitively the problem of Artificial Intelligence.[4]

Those are but four of the many methods used to teach machines today, and the field of machine learning is ever growing. It’s without a doubt in my mind that reaching the pinnacle of weak and strong AI is only a matter of time, and when such a event occurs we’ll need to have machine ethics on par.

Ⅳ
INTELLIGENCE

“Intelligence, by its very nature is something that cannot be understood but not because understanding it is impossible but because understanding it destroys our perception of it as intelligence.”[3] , does this mean that when we understand how the human mind work we stop becoming intelligent? If not, then what of machines when we give them our thinking process, are they still not intelligent? As A.M. Turing said, “It doesn't matter how it works - if it behaves like a human intelligence then it is a human intelligence." [3]

Two theories of the brain exist namely the **Grandmother Cell Theory** and the **Distributed Representation Theory**. The first theory asserts that individual neurons have high information capacity and are capable of representing complex concepts such as your grandmother or even Jennifer Aniston. The second theory neurons asserts that neurons are much more simple and representations of complex objects are distributed across many neurons. Artificial neural networks are loosely inspired by the second theory.[5]

So even when we model machines thinking process after the very thing that makes up our mind we still deny them the right to think. Why is that? One would argue the following:

“*All digital computers are binary systems. This means that they store and process information exclusively in terms of two states, which are represented by different symbols—in this case 1s and 0s. It is an interesting fact of nature that binary digits can be used to represent most things; like numbers, letters, colors, shapes, images, and even audio with near perfect accuracy.*

This two-symbol system is the foundational principle that all of digital computing is based upon. Everything a computer does involves manipulating two symbols in some way. As such, they can be thought of as a practical type of Turing machine—an abstract, hypothetical machine that computes by manipulating symbols."[2]

At the moment our machines are being built with fine metals such as copper, gold, and silicon among other materials, and it's these very materials that limit our machines; but not for long. It's only recently that we've looked into other fields to improve our Thinking Machines. The field that is of most use to our Thinking Machines is Biology, the two intertwined is commonly referred to as Biotech. A field that has already accomplished various feats from DNA digital data storage to artificial neurons.

"While still in their infancy, **DNA Computers** will be capable of storing billions of times more data than your personal computer. Scientists are using genetic material to create nano-computers that might take the place of silicon-based computers in the next decade."[6]

DNA is in four binary code, meaning it processes in 0s, 1s, 2s and 3s that's double the binary then our current machines can process! Double the binary is a feat that solves just about every issue current machines exhibit such as corruption, physical limitations, speed and so much more will be revolutionise by DNA Computers. "One pound of DNA has the capacity to store more information than all the electronic computers ever built; and the computing power of a teardrop-sized DNA computer, using the DNA logic gates, will be more powerful than the world's most powerful supercomputer."[6]

"A Turing machine's operations are said to be "syntactical", meaning they only recognize symbols and not the meaning of those symbols—i.e., their semantics. Even the word "recognize" is misleading because it implies a subjective experience, so

perhaps it is better to simply say computers are sensitive to symbols, whereas the brain is capable of semantic understanding.”[2]

REFERENCES

1. "*Computing Machinery and Intelligence*"
   - ➢ Author: Alan Mathison Turing
   - ➢ Date Published: 1950
   - ➢ Publisher: Oxford University Press on behalf of the Mind Association
2. "*A neuroscientist explains why artificially intelligent robots will never have consciousness like human*"
   - ➢ Author: Bobby Azarian,
   - ➢ Date Published: Mar-31-2016,
   - ➢ Publisher: RawStory
3. "*Artificial Intelligence - Strong and Weak*"
   - ➢ Author: Alex Armstrong,
   - ➢ Date Published: May-04-2015,
   - ➢ Publisher: I Programmer
4. "*Libro de IAAR*"
   - ➢ Author: IAAR,
   - ➢ Date Published: Jun-19-2017,
   - ➢ Publisher: IAAR (*Comunidad Argentina de Inteligencia Artificial*)
5. "*10 Misconceptions About Neural Networks*"
   - ➢ Author: Stuart Reid,
   - ➢ Date Published: May-8-2014,
   - ➢ Publisher: TuringFinance
6. "*How DNA Computers Will Work*"
   - ➢ Author: Kevin Bonsor,
   - ➢ Date Published: Nov-17-2000,
   - ➢ Publisher: HowStuffWorks
7. "*From Narrow to General AI and from External to Internal Intelligence*"
   - ➢ Author: Peter Voss,
   - ➢ Date Published: Oct-3-2017,
   - ➢ Publisher: Medium
8. *"Computers and Thought*"
   - ➢ Authors: Edward Feigenbaum, Julian Feldman, Mike Sharples
   - ➢ Editors: Edward A. Feigenbaum, Julian Feldman
   - ➢ Date Published: 1963
   - ➢ Publisher: McGraw-Hill Book Company
   - ➢ Locations: New York, San Francisco, Toronto, London, Sydney