Thinking Machines

Art by Google's DeepDream

By CMCL

INDEX

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 convincing the reader that machines are thinking, and are capable of reaching consciousness, thus granting free will. It's a reality that has yet to unfold. It's only when we think of it from this point of view that we concern ourselves with the implication that Thinking Machines with strong/general artificial intelligence, AI, will have on the way of life once commercially available.

In a nutshell, such a event would mean anyone and everyone could be replaced by a Thinking Machine that surpases us in every aspect. While this might seem impossible, or at the very least very unlikely, we've already been seeing this since long ago, starting all the way back in the Industrial Revolution. For instance, we keep getting replaced by machinery in the work field, and it's due to the fact that machines provide a efficiency that we can never compete with. Now ask yourself this, what will we do when Thinking Machines are capable of doing everything we can do, but at a efficiency we can never possible be on par with, nevertheless surpass?

I believe that if we were to analyze Thinking Machines from today onwards in our education system and government, on all levels, I don't believe such a event would be necessarily bad. For we'll be able to determine how Thinking Machines will interact with us and vice versa, but before we do such a thing we need to ask ourselves the following.

"*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 based on the capabilities of digital machinery at

the time, since at the time there was only a handful of digital machines, but instead asked the question if any imaginable digital machine is possible of thinking? His work is the foundation for various branches of computing including AI development with his famous "*Turing Test*", 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 in 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.

It's without a doubt in my mind that machines capability for thinking is entwined with the development of AI, and just because we haven't reached strong/general AI doesn't necessarily mean that our current machines aren't capable of thinking as A.M. Turing said, "*Just because someone or something thinks differently than you do, it doesn't mean that it's not thinking*". So what's important about machine's capability for thinking?

Well today, our way of life is affected by machinery, and AI at a smaller scale, in most cases it eases and simplifies the process by which we accomplish things while still requiring human intervention in some way or another; and in other cases it does exactly what we wish to accomplish without the need for human intervention besides maintenance and turning it on and off. Just a quick glance, and one can see that machinery affects many aspects of our life like our culture, economy, transport and other aspects of our life that are both insignificant and crucial.

Machinery has such a big affect on our way of life that one would asks, how did machinery become so embedded into our life? I'd say that the embedding of machinery into our life started with the Industrial Revolution, in which the way of life was transformed so drastically in such a short time that many could not adapt and suffered immense consequences. It was only after we surpassed multiple social, cultural,

economic and other problems the hard way that we analyzed ourselves as a whole and noticed something was wrong.

The wrong, which originated way back in the Industrial Revolution, is still present today permitting multiple problems, that were faced back then as well, to occur to this day the most threatening problem, in my opinion, is the way machinery is replacing our role at work.

So what importance does machinery and the problems caused by it's embedding into our life all the way back from the Industrial Revolution have to do with the philosophy, psychology and ethics of Thinking Machines? As George Santayana said, "*Those who do not learn history are doomed to repeat it.*", and as I've mentioned before AI, at the moment, already affects multiple aspects of our life on a small scale; but every year it's applications are growing at such a rate that AI will be on par with machinery in a matter of years.

It's without a doubt in my mind that we're once again going to achieve a revolution this time being a "AI Revolution", and if we don't prepare for this revolution I assume the worse for mankind as Benjamin Franklin said, "*By failing to prepare, you're preparing to fail.*"

I do not claim to know the future, but one, such as myself, can analyze the past and see where Thinking Machines are heading. At our current phase I believe we'll achieving the AI Revolution within the next century. I predict such a thing will occur based off Moore's law which states, "*The number of transistors in a dense integrated circuit doubles approximately every two years*"; however, this law has been heavily modified since its emergence in 1965. Due to physical and manufacturing limitations the span of time it takes for growth to occur has increased from two to two and a half years, and I expect the span of time to continue to increasing. One day we'll indefinitely hit a

wall due to physical limitations, such a thing would mean that Thinking Machines, most likely all machines, would move from silicon CPUs, Central Processing Unit, to CPUs made with higher grade material, we'll go further into detail later, such a thing will open doors for Thinking Machines that now seem impossible to open.

So what will happen when a AI Revolution occurs? We'll first we must ask ourselves, what does a AI Revolution even intend? In a nutshell, a AI Revolution would mean we've surpassed weak AI and have achieved strong/general AI, we'll go further into detail of what weak and strong AI is later. Such a feat would mean anyone and everyone's role at any given place is at stake of being replaced by a machine, even the geniuses of the world like Stephen Hawking, Bill Gates, Jeff Bezos, Elon Musk and soo many more agree that such a event is unavoidable.

So if such a event is unavoidable then what's the point of researching Thinking Machines? We'll at the time of writing this we have capability of analyzing why Thinking Machines will one day surpass us in every aspect, and in doing so we can adjust them and ourselves accordingly. 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? Well if history has shown us anything it's that you can't stop progress, therefore one day the perfect killing 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 a ethical view, in other words, we need to evolve machin ethics, we'll go further into detail later, to the point where Thinking Machines have a moral compass that's even more just than that ours.

Let's say that machin 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 position on Thinking Machines is 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.

I
THINKING

"*Just because someone or something thinks differently than you do, it doesn't mean that 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 very thing we are. Yet, we hold on tight to our theories as if they were the truth. With those very 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 a 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 a 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 that of 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 that 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: 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