On The Subject of Thinking Machines

John Olafenwa

johnolafenwa@gmail.com

Moses Olafenwa

guymodscientist@gmail.com

Abstract

68 years ago, Alan Turing proposed the question "Can Machines Think" in his seminal paper [1] titled "Computing Machinery and Intelligence" and he formulated the "Imitation Game" also known as the Turing test as a way to answer this question without referring to a rather ambiguous dictionary definition of the word "Think"

We have come a long way to building intelligent machines, in fact, the rate of progress in Deep Learning and Reinforcement Learning, the two corner stones of artificial intelligence, is unprecedented. Alan Turing would have been proud of our achievements in computer vision, speech, natural language processing and autonomous systems. However, there are still many challenges and we are still some distance from building machines that can pass the Turing test. In this paper, we discuss some of the biggest questions concerning intelligent machines and we attempt to answer them, as much as can be explained by modern AI.

Can Machines Think

Turing choose to avoid answering this question directly, however, it is important to have a clear and concise meaning of thinking that incorporates lessons from neuroscience and Artificial Intelligence. We define thinking as "The process by which we evaluate features learned from past experiences in order to make decisions about new problems" In the context of human thinking, when you see a person and you are faced with the task of determining who the person is (The New Problem), a activity (The Process) begins in your brain that goes through the search space of all the people whose face you can remember (The Experience), you then begin to consider the nose, eyes, skin color, dressing, height, speech and any other observable treats (The Features), the process then attempts to match these features to a particular person based on people we have seen before, if no satisfactory match is found, the brain concludes that this is a stranger (The Decision).

Consider a computer vision system on the other hand, trying to perform the same task using Convolutional Neural Networks, when the image of a person is imputed, the 3 Dimensional Tensor of pixels are observed, (The Features), the network then searches (The Process) for the presence of previously learned features called kernels or filters (The Experience), it then compute which of these features are present in the new image and returns a set of class scores (The Decision).

The process is remarkably similar, except for the process by which it is done, for example, Convolutional Neural Networks do not put the position of features into consideration, a nose at the position of the ear makes no difference to a CNN, the process in humans puts this into consideration, however, Capsule Networks, recently proposed by Geoffrey Hinton et al. [7] seeks to address this weakness.

Thinking is not an act of magic that happens illogically, it is a well defined sequence of discrete actions that include observing problems, comparing observations to learned features and making decisions based on the results of our comparisons. An interesting problem that may be considered not to follow these sequence include emotional decisions pertaining to love, anger, sadness etc. Such decisions are often taken stochastically and cannot always be defined within the framework of thought that we formulated, decisions in such circumstances are often exploratory rather than exploitative. Does this disprove our earlier assertion? No, intelligent machines are fully capable of stochastic actions, the concepts of exploration and exploitation are central to the field of Reinforcement learning, a key enabler of autonomous systems. We shall emphasize further on this soon.

In view of this, we argue that machines do think, just differently from the way humans think. This begs the question, if the framework of thought is the same for both machines and humans, why do they think differently? The answer to this lies in the realm of the components of intelligent systems. Artificial neural networks were inspired by neuroscience but their mechanisms are fundamentally different, we have largely given up on searching for systems that work like the brain, rather we keep searching for systems that work well, without minding how far we are deviating from the way the brain functions.

The human brain and the intelligent systems we create are very different, but the framework of thought is the same. Machines can think and it's only a matter of time before they truly become capable of thinking without using much energy.

Can Machines Imagine

Accepting that machines can think is a little challenging for critics of Artificial Intelligence, it is even far greater a challenge to consider machines capable of "Imagination" To answer the question of "Can Machines Imagine" we must first deeply consider the framework of imagination.

Thinking is often based on past experience as we earlier explained, however, imagination is the formulation of ideas which we have not learned "explicitly" from past experience. We added the term "explicitly" to indicate the fact that while imagination results in new untested ideas, it often draws inspiration from previous experiences. Hardly is any imagination completely new without being rooted in an existing idea, such might be a possibility, however, the connection with past experience might be highly implicit so much that we might not realize how much past experience has influenced our imagination. In simple terms, imagination is often an exploratory action that is guided partly by learned features from past experiences. These exploration is stochastic but still partly motivated due to its connection to past experience, such connections can be weak, but very vital.

This framework of imagination changes the question, "Can Machines Imagine" to "Can Machines Develop New Ideas" The later question more appropriately describes the first question. Supervised Deep Learning, which is the most popular method for building Artificial Intelligent systems is solely based on learned features and by nature is at present not capable of producing machines that can develop new ideas of its own. Reinforcement learning however is based on machines that can exploit current knowledge just like supervised deep learning, but can also explore its environment by trying new actions to discover better ideas. This is very much like imagination. The ability of an agent to discover new policies or new sub policies in a hierarchical setting [4], gives such agent the ability to formulate new ideas and even do new things that we never expected it to do in a particular environment. There are limitations in the form of such discovery being limited by the finite set of actions available to the agent at a given time step or in a state, however, humans are subject to very similar limitations, for example, you cannot just imagine that you want to fly without some jet pack or related equipment.

It can be argued that exploratory moves only mimics imagination, however, no one can tell precisely the process by which imagination occurs in the brain, what we are sure of is the effects of intuition, hence the effects of exploratory moves can be characterized as a form of imagination. A lot still needs to be done in infusing machines with great imagination capabilities. To answer the question "Can Machines Imagine" or "Can Machines Develop New Ideas" the answer is Yes. With time, the ability of machines to imagine would improve as new techniques are discovered.

Can Machines Feel

"Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain-that is, not only write it but know that it had written it. No mechanism could feel (and not merely artificially signal, an easy contrivance) pleasure at its successes, grief when its valves fuse, be warmed by flattery, be made miserable by its mistakes, be charmed by sex, be angry or depressed when it cannot get what it wants."

The above statement was made by Professor Jefferson in 1949. It sums up the view of critics of Artificial Intelligence. The central core of the above assertion is the part, "No mechanism could feel (and not merely artificially signal, an easy contrivance)"

The idea that reaction based on artificial signals cannot be described as a feeling is rather contrary to how the human system functions. All forms of feelings in the body, including pain, anger, love, fear etc. are not a resultant of some inexplicable processes in the human body, rather they are actions initiated by the brain based on signals received from the sensory organs. For instance, pain is not actually generated on the skin, instead the brain receives signals from the skin, encoding the degree of touch and it then initiates a set of reactions that makes us sense pleasure or pain.

There exists such people that due to brain malfunction, they cannot feel any pain even if injured.

A feeling is the response the brain gives to a signal.

In the context of this definition, machines can feel.

Take a humanoid robot, we can build sensors into the body such that when it is touched, signals would be sent to the processing system of the robot, which in this case is acting as a brain, the cumulative signals over a period of time would form a sequence that can then be fed into a type of supervised deep learning system called Recurrent Neural Networks. The system can then interpret this signal sequence as pleasure or pain, and could initiate responses such as smile, laugh or frown. This entails that feelings are not limited to biological organisms, Machines can Feel. They can compose music in response to signals from the environment, they can feel happy or sad and can enjoy the pleasures of life if trained on sequence of signals and corresponding feelings. There already exist practical demonstrations of these, "Solo" is a home appliance by a company named Uniform. It can play songs based on your facial expressions, if you are happy, it can play songs to help you enjoy your time and if it detects that you are down, it could play songs that would lift your spirit. A lot more is still to be done. Machines can feel and act based on feelings.

Can Machines Be Conscious

This is perhaps the most daunting question about machines. Consciousness is highly regarded as a trait that is unique to living things. It is generally a state of being aware of your existence.

The question about machines can be more appropriately asked as "Can machines become the subject of their own thoughts"

Having established that machines can think, it is most appropriate to consider a conscious machine as any machine that can think about itself. This has a lot of implications. A machine that can do so would begin to care about its existence and can make efforts to prevent its creator from performing any action that can cause it to stop functioning. Such an ability is often a scary prospect, many believe machines possessing such abilities would become terminators, however the question of "Are conscious machines desirable " is not the same as “Are conscious machines possible" Machines can indeed become the subject of their own thoughts. Here is an example. Consider a robot that is programmed to learn not just from how its actions influence its environment but also from how the actions of other agents in its environment influence its internal state. Such other agents maybe humans or other robots. In such a scenario, robot A might learn that some actions of an adversarial agent G is causing it to loose power faster than normal, possibly by making it overwork, robot A may then decide not to obey commands giving to it by Agent G, since data from past experience indicate a negative effect from commands initiated by Agent G. It might also learn that the actions of another Agent D improves its internal state, hence it might develop more affinity for Agent D by obeying D more than G. Such a machine is conscious and is possible within the framework of Deep Reinforcement Learning.

The potentials of these can be scary but can be beneficial too. Robots that are conscious could learn to become more effective at work by learning which conditions would improve their working ability, such robots could learn to even work in collaboration with other agents, all on their own without reprogramming.

The big difference between humans and conscious machines is the role of the "subconscious" in humans. The subconscious part of humans remains largely mysterious, many actions are initiated by it, hence, we can at least have a sense of superiority over robots based on the fact that there is at present no proof that we can infuse machines with sub consciousness within the framework of Deep Learning and Reinforcement learning. Time would tell if someday, breakthrough in neuroscience and Artificial Intelligence can unlock the secrets of the brain and make it possible to build super-intelligent systems.

Conclusions

All around us, the applications of artificial intelligence are present. Google Assistant, Microsoft Cortana, Hound Assistant and Siri have greatly changed the way we interact with our smartphones. Alexa, Amazon Echo, Google Home and Apple's Home Pod are making home automation available to everyone. The effects on industries are even more pronounced. Every aspect of business, from Stock Trading, Recruitment, Manufacturing, Marketing, Decision making etc. are becoming automated. Autonomous cars are gradually becoming common sight in advanced cities. The applications of artificial intelligence would fundamentally transform our society in the same way that digital computers did. This is the fourth Industrial Revolution.

However, despite the promises of a more advanced civilization, a lot of people are rightly scared of the consequences of widespread automation and military uses of artificial intelligence. There are fears of job displacement due to automation. These fears are genuine and need to be addressed. Disruption is inevitable but all stakeholders including scientists, captains of industries, investors, directors, shareholders and the government must work towards applying artificial intelligence in such ways that it would augment the roles of humans at work rather than replace them. There exist such jobs in which machines need to absolutely replace humans, this include very dangerous works that lead to the death of many workers yearly. For example, jobs that expose humans to harmful radiation should be done by machines. Rescue operations in the event of disasters can be greatly improved with robots, by these we can save many lives.

Military uses of Artificial Intelligence is also a dangerous concept which should be avoided. Great scientists including Yoshua Bengio and Elon Musk has spoken against weaponizing artificial intelligence.

We need a future where our activities are assisted by Safe Artificial General Intelligence (AGI). A great effort to achieve this is the OpenAI initiative, a non-profit research organization that is dedicated to creating intelligent systems that would help and not destroy humans.

Let us all work together to create a better future for the human race.

AUTHORS

JOHN OLAFENWA

A self-taught computer programmer, Neural Networks Blogger and Computer vision researcher. Skilled in Building Android applications and Native software. Can develop software with Java, Python and C#. Very passionate about transforming lives through highly efficient neural networks. Studies reinforcement learning at leisure time.

Email: johnolafenwa@gmail.com,

Website: john.specpal.science

Twitter: @johnolafenwa

MOSES OLAFENWA

A self-Taught computer programmer, Cloud and Internet Logistics expert. Skilled in developing Android applications

Web portals and Desktop software. Can code in Java, Python and PHP. A Deep Neural Network practitioner with a vision to make the world better via Artificial Intelligence. A lover of Big Data.

Email: guymodscientist@gmail.com,

Website: moses.specpal.science

Twitter: @OlafenwaMoses

References

[1] A.M Turing (1950) Computing Machinery and Intelligence. Mind 49: 433 - 460

[2] John Olafenwa, Moses Olafenwa. FastNet: An Efficient Architecture for Smart Devices. <hal-01685480> https://hal.archives-ouvertes.fr/hal-01685480

[3] Yoshua Bengio, "Learning Deep Architectures for AI," in *Learning Deep Architectures for AI* , 1, Now Foundations and Trends, 2009, pp.136-
doi: 10.1561/2200000006

[4] Kevin Frans, Jonathan Ho, Xi Chen, Pieter Abbeel, John Schulman, Meta Learning Shared Hierarchies, https://arxiv.org/abs/1710.09767

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun . Deep Residual Learning for Image Recognition . https://arxiv.org/abs/1512.03385

[6] W. Xiong, L. Wu, F. Alleva, J. Droppo, X. Huang, A. Stolcke. The Microsoft 2017 Conversational Speech Recognition System. Microsoft AI and Research Technical Report MSR-TR-2017-39 https://arxiv.org/abs/1708.06073

[7] Sara Sabour, Nicholas Frosst, Geoffrey E Hinton. Dynamic Routing Between Capsules. https://arxiv.org/abs/1710.09829