Erdem Cagatay – 76215383 26.11.2018

PHIL250

Final Paper

Professor Christopher Mole

**Quest of Imitating Human Mind on Artificial Media**

Artificial Intelligence (AI) has gained quite an importance for the decades since the very well-known work *Computing Machinery and Intelligence* [1] of famous scientist Alan Turing is published in the journal *Mind* in 1950*.* That was the accelerator of the scientific research on the area that is led by philosophical questions. Since then; many questions have been asked and with the help of those questions, AI has become more and more sophisticated by addition of novel branches such as deep learning, predictive analysis, language processing, image recognition, speech manipulation, neural networks, and so on. However, beyond the various curiosities that led to so many specification areas; there has always been some more widely-scoped interests against the potential applications of artificial intelligence, one of which constitutes the basis of discussion for this paper. Can we transfer/transfuse human mind onto an artificial medium? If so, how can we achieve it?

Over the years, the curiosity has driven many people into different approaches both in favor of positive and negative answers to the question raised above. Some people believe that this concept has no mean of realization while the other group of people embraces the possibility that the human mind can be very well simulated on some sort of artificial medium. What this paper will present on this issue is that we can manage to end up having such a concept, even though there are some definite steps to be taken and most of those steps won’t be taken directly by us. To be clearer; we -human beings- will not be the ones who will have direct responsibility of this act, but we will lead the way towards the final phase of fully simulating a human’s mind on an artificial medium for the reasons that will be presented further.

The paper will propose some certain milestones to be handled to achieve this goal. The paper will have its focus on the details of those milestones and how to achieve them as well as it will attempt to present some responses to possible objections raised against its arguments.

The approach held against this issue will be somehow related to the ideas on evolutionary learning on computers, which took place in the paper *Life and Evolution in Computers* [2]. It will be presented that in the future, machines will be expected to be the main actors of examining, interpreting, and reflecting the characteristics of the human mind onto an artificial intelligence by using an evolutionary way of advancing. What remains as the task of humans is to develop the main roots for that evolution. After humankind having improved some certain mechanisms and technology, machines will end up with the ability to put humane way of thinking to a machine.

The very first technology to develop must be a mechanism for machines to have them observe human behavior closely and relate them to certain mental states, emotions, reactions to specific events, etc. This methodology will be the key to examining the human mind and will constitute one of the main legs of advancement. At the current situation, there are some certain inventions that pose as the proof of possibility for such an advancement. In their astonishing paper, Pantic and Bartlett have demonstrated the results of their work on the programs which observe human facial expressions and relate them to certain moods [3]. In their work, they explain many newly developed techniques to read human facial expressions such as face-muscle action detection, appearance-based facial affect recognition, facial characteristic point detection, and so on. As their research indicates, some mechanisms for examination of human behavior can be and is being developed. As the technology advances further, it is expected that also minor parameters on the human body can be measured and can be related to certain conditions by programs. There has also been some introductory research on measuring some minor values of the human body such as blood rates, through the help of Artificial Intelligence [4]. It must be kept in mind that; this technology is not supposed to be developed as a helper for the reading human behavior, but instead, it is to let machines have a sort of database for learning human behavior.

After this crucial step, another milestone is to have a way for machines to demonstrate humane reactions to happenings. In other words, machines must be able to mimic human behaviors and emotions against certain situations. This may sound like a very imaginary concept, but there have been some advancements on this regard as well. In his paper *Toward Machines with Emotional Intelligence,* Picard has given striking examples of programs that show emotional responses as humans do [5]. Maybe the most surprising one of them is the fitness-helper program “Laura”, which demonstrates emotional responses to users’ interactions. Of course; at this level, this program is not way more advanced than a chatbot in terms of building emotional relationships, but still, it affected the performance of users in a significant way. Therefore, we can say that in the future, there will be programs/robots that will be able to communicate us in a reactional way that appeals to our emotions. This is also another must-have feature because so far proposed two mechanisms will constitute the basis for the continuous development of simulating the human mind. After machines will learn how to interpret the information they acquire with the help of firstly proposed technology, they will use those interpretations to show humane reactions by the means of emotional/behavioral intelligence skills offered here.

As it is seen, so far proposed technologies serve as the ‘input’ and ‘output’ of a prospective human mind simulation. Therefore, there is still a gap of interpretation between these two. This means that machines need to learn how to process the data that is learned from observing human behavior and after only that process can they exhibit the outcomes of their learning. At this point, the very first way of achieving this might seem as coding the programs how to process the information learned. Nevertheless, there are some big obstacles and sort of paradoxes before doing that. First, as humans, we do not fully know how the human mind works in detail. We don’t even have a precise definition for the mind. Thus, it doesn’t seem likely that we will be able to learn exactly how human mind works and will put that mechanism onto a machine. This situation can be considered as analogous to trying to interpret outer dimensions while we still live in 4-dimensional reality. We can deduce some points by experimental search or theoretical work on the topic, yet we cannot have any possibility of discovering them completely. Therefore, we will need an objective, self-existent agent with this functionality. This can be reached through the means of self-learning programs that will create the facilities of interpreting human mind. Self-learning machines may sound like a very fictitious concept, however; today, some research on this area has been practiced and their results are really promising.

Along with many other works on the topic, Rashwan and Al Sallab’s research presents an example of how machines can use deep networks and adaptive learning to improve their certain skills [6]. In the paper; it is given that with a minimal level of supervising, machines will get better at grouping visual data after some process of ‘learning’. A similar advancement can be had with the behavior interpreting programs as well with using the evolutionary perspective in [2]: Some basic guidelines may be given, and the program is then set to evolve. At this point, it is also critical to have a measurement of success on the regard. Do we expect to fully copy a person’s mind, or do we expect to end up with a newborn human’s mind? According to which parameter is accepted, the evolution of the program(s) can find their paths as it is presented in [2].

Those were the critical steps to be taken to achieve transferring a human mind into an artificial medium. If they are realized and implemented, theoretically there will be no obstacle before a program that can observe, interpret, and practice the human mind. Having presented the required path to be taken, it is very likely for people to have some questions or objections to the arguments claimed.

One might object to this approach by saying that a machine can never fully experience having a human mind because its physical entity is not as humans have. This might seem reasonable at first, but it is better keeping in mind that our senses and physical responses against outer impulses are nothing but our brains’ signals to/from our nerves. Nerve signalization can also be simulated after achieving the first and the third step that allows a machine to trace a human’s thinking process. There has already been some research on this area before, and results show that it is possible to create a virtual ‘sensitivity’ for machines. In the article [7], it is presented that with the help of electronically mimicked distributed network of receptors, neurons, and synapses, researchers were able to create an artificial sensory system for robots. By using this technology and enhancing it with the methodology presented as the 3rd step above, it is theoretically possible for robots to experience human sensation. Thus, this objection can be falsified.

Another objection to arguments presented in this paper might be that interpreting and mimicking human behavior does not make it the exact same way as the human mind does. In other words, a machine is still not capable of becoming human even though it perceives the world and reacts to happenings just as the human mind does. This objection may be considered close to the system objection in the paper *Minds, Brains, and Programs* by Searle [8]: Machine functions the same way as humans do but still, it does what it is programmed to do. This objection is related to how objector views the human nature. In my opinion, humans can be considered as bionic systems of which programs are coded into their DNAs. We are how we are because the coded data in our DNAs helped us through the natural elimination process. We believe that we do conscious choices, but they are nothing but optimization of pleasure according to the resources we have. Even moral actions that we take serve as the fulfillment of our self-judgmental pleasure. Leaving daily small conscious choices aside, we are ruled by the very basic instincts that are inherited from our ancestors. We make money to have an acceptable life because we hate being alone and our instincts order us to socialize and have offspring. We want the world to be saved because we want the next generations to keep on surviving as it is coded into our DNA. Simply, one should think of life as the programs running on a carbon-based unit. Therefore, mechanization of this bionic system can be achieved if proposed technological advancements in the paper are reached.

To conclude, the human mind can be mapped and reflected onto a mechanical medium, but this won’t be achieved by humans themselves. Instead, it is going to be machines which have the technology of examining human behavior and practicing that behavior by its means. For machines to achieve this, they also need to have an evolutionary learning mechanism that will make them get better and better without human control. So far scientific research has shown us that these sorts of developments are possible and will likely take place in the future. On the other hand, there might be some objections to this thought which claim that the mechanization of the human mind is impossible because of humankind’s unique nature. This is disputable: to me, humans are nothing but carbon-based, DNA-programmed bionic systems. Therefore, their senses and mechanisms can be replicated and adapted to work on machines.

**References**

[1] Turing, A. M. *Computing Machinery and Intelligence*. Blackwell for the Mind Association, 1950.

[2] Mitchell, Melanie. *Life and evolution in computers.* Santa Fe Institute Working Paper 2000-01-001, 2000.

[3] Pantic, Maja, and Bartlett, Marian Stewart. *Machine Analysis of Facial Expressions*. INTECH Open Access Publisher, 2007.

[4] Ali, Shawkat, et al. *Non-Invasive Blood Glucose Measurement Performance Analysis through UWB Imaging*. 2016 3rd International Conference on Electronic Design (ICED), 2016.

[5] Picard, Rosalind W. *Toward Machines with Emotional Intelligence*. Vol. 23, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001.

[6] Al Sallab, Ahmed, and Rashwan, Mohsen. *Self Learning Machines Using Deep Networks*. 2011 International Conference of Soft Computing and Pattern Recognition (SoCPaR), 2011.

[7] Kim, Yeongin, et al. *A Bioinspired Flexible Organic Artificial Afferent Nerve*. Vol. 998, Science, 2018.

[8] Searle, John R. *Minds, Brains, and Programs*. Cambridge University Press, 1980.