Doozandeh, P. (2015). *Introspection, emotion, and computational theory of mind* [unpublished manuscript]. Accessible in p-doozandeh.com

Introspection, Emotion, and Computational Theory of Mind

Pooyan Doozandeh

Author note:

This manuscript was completed in April 2015, and uploaded online in February 2021.

No funding agency or university supported the creation of this manuscript. Correspondence should be directed to Pooyan Doozandeh (pooyan.doozandeh@gmail.com).

Abstract

Is the Computational Theory of Mind (CTM) based on reasonable assumptions? By having an abstract look on CTM, this article discusses this question and ultimately finds the answer negative. I investigate several important assumptions of CTM; among those are introspection, formalization, and the emotionless intelligence. I first probe introspection and conclude that, although it is the sole origin of investigations about the mind, it can never be a reliable source of knowledge. Then, I shortly restate the famous argument of the impossibility of intelligence without human emotion. I continue by showing how CTM assumes introspection as valid, formalization of the theories of mind as possible, and the possibility of intelligence without emotion. By rendering CTM as a groundless assertion, I believe that artificial intelligence and cognitive science should not be ignored on the same basis; I instead offer a different interpretation of the goals of AI and cognitive science which have hitherto been the two progeny of CTM.

***Keywords*:** CTM, Introspection, Emotion, Frame Problem

Introduction

Artificial intelligence and cognitive science have many shared premises and come from the same background and history. These fields continue to operate without the slightest care of whether the assumptions behind them are justified or not. With regard to their original goals, however, an abstract look is required that is not only necessary to derive the fundamental assumptions but also assesses the philosophical implications of those goals. These assumptions must, therefore, be brought to surface so that we can clear the clouds and fully examine the possibility and limits of their goals.

In the present article my major discussion revolves around the Computational Theory of Mind. I think the shared premises of both artificial intelligence and cognitive science can be summarized in CTM and if we reduce our discussions to it, we can scrutinize both AI and cognitive science. This article begins by claiming that introspection is the sole origin of every investigation about the mind. By providing reasons and evidence from science and philosophy, I will try to show that introspection, naturally uncertain and volatile, cannot be justifiably considered as true. Next, I will briefly discuss the primacy of emotion in every aspect of human cognition and mind.

The primary section of this article investigates CTM from a higher perspective. By depicting an abstract picture of the CTM and showing its relations and similarities with psychology, I shortly discuss why psychology, unlike AI and cognitive science, can experimentally justify its theories of mind. We will see how the CTM makes several assumptions; among those are the validity of introspection, possibility of formalization, and the insignificance of desires in cognition. With responding to a criticism against the argument from emotion, I discuss these assumptions, showing how CTM is founded on uncertain principles.

Later, I will shortly debate the frame problem, arguing that it is a technical problem and can only be investigated in the realm of AI and cannot therefore be used for or against the CTM. I also briefly discuss an argument that uses the frame problem as an evidence for the CTM.

After bringing the basic assumptions of the CTM under question, I will show that, if we adopt the perspective that I depict, what possible justification and interpretation will arise for

the AI and cognitive science and on what grounds researchers in those fields should base their activities.

The Nature of Investigations in Mind

Literature in the philosophy of the Computational Theory of Mind has rarely investigated its assumptions, especially from a higher perspective which, I believe, is crucial. To begin with, it seems clear that mind and mental faculties cannot be the subject of conventional experimental enquiry. If we accept Kant's (1787/1929) epistemology, for any knowledge to be gained scientifically, there must exist two fundamental components: *material of knowledge (or object)* that takes the *form of knowledge (or subject)*[1]. This object-subject distinction is a desideratum for any conventional science.

With regard to mind, however, there cannot be any material of knowledge, rendering the experimental knowledge of mind impossible. In other words, in investigations concerning the mechanisms of the mind, human subject aims to know itself. There is no outer world or object and as a result, the relationship becomes subject-subject. The mind is directed towards itself and investigates its own mechanisms. It is widely known as *introspection* and is the sole origin of investigations about the mind. Descartes' cogito ergo sum, Hume's association of ideas, Freud's psychoanalysis, the methodology of today's psychology, and philosophy of mind in general (except behaviorism that rejects the existence of mind altogether) are all solely or primarily based on introspection.

Unfortunately, the nature of introspection is highly problematic. The obvious pseudo-logical contradiction lies in this question: is mind justified in trying to know itself? In other words, is it possible for a knower to know about its knowing faculty? As another objection, introspective theories of how the mind works cannot easily be brought into experimental investigations and there is no unanimous method of testing them. As a result, there is the danger of irrefutable solipsistic subjectivity; one may argue that his mind works in a certain way and he can unquestionably generalize it to every other human, accusing disagreeing party of ignorance and myopia about their own minds. There does not seem to be a simple way to

[1] This part of Kant's philosophy, which is one of his most important parts and a historic result of his investigations, is the one which is accepted by most of his successors. Even his subsequent critics generally agreed with Kant on this issue. For the detail of this part of Kant's philosophy, see Kant (1787/1929).

define a standard of methodology for investigations about the mind. Many disagreements in the history of the philosophy of mind are rooted in precisely the same issue.

Although many believe that their own mind is the most secure source of knowledge, there are many arguments against the validity of introspection. Dennett (1991)[2] shows how far we can be wrong about the visual capacities of our minds. He suggests an experiment in which you are asked to determine a card's identity from a deck of cards. You should focus your vision on one point and then choose one card from the deck. Without looking at the card, keep it at an arm's length and slowly bring it from a side to the center of fovea, while trying to recognize the identity of the card. The result shows that you should bring the card completely to the center of fovea in order to be able to recognize it, in spite of the fact that most people believe, prior to the experiment, that they are able to recognize it even away from their center of fovea.

Dennett tried to exemplify how we can be wrong in overestimating our visual capacity and he argued the same kind of mistakes exist in assessing our mental capacities. Schwitzgebel (2008) brings further evidence from both experimental and phenomenological realm to undermine the validity of introspection, calling it '… simply banal cases of self-fulfillment'. He argues that skeptics like Hume (1738/2004)[3], rightly proved that the truthfulness of our understanding of the world is highly in doubt. He further thinks that such doubt and confusion is not merely restricted to the outer world, but can even be more strongly applied to the inner world. In short, he believes that, although introspection is an integral part of every human's life, because of its volatile and uncertain nature, it cannot be used as a reliable tool to get at the truth. The impossibility of certain knowledge of the mind is also supported by many philosophers.[4]

---

[2] See also Dennett (2001).

[3] Hume was the first to seriously challenge people's confidence about their knowledge from the external world. He argued that most of what we call knowledge, especially those that come from science, is gained through inductive reasoning. He proved that people's grasp can only reach the two ends of inductive knowledge from the external world: cause and effect. We can never experience, and therefore investigate, the relationship between the cause and effect. This idea was later supported by Kant, who argued that we can only experience phenomena and our understanding cannot reach the noumena.

[4] For example, as an ontological argument, Heidegger suggested that most of what we know as human knowledge belongs to the realm of substances that have properties. In contrast of any substance, the nature of human mental phenomena cannot be defined in terms of an entity with properties. For further views from Heidegger on this issue, see Heidegger (1926/1962).

The Primacy of Emotion

By looking at nature, the first obvious fact about every living organism is its struggle for survival and reproduction. Not many assertions are supported unanimously by most philosophers and scientists like the primacy of desires in every form of life. An abstract look on our life also reveals that every human activity is and can be defined in terms of its relations to survival and reproduction. To mention a few instances of philosophers whose ideas about the primacy of desire are generally accepted, we can name Hume who, in his appreciation of the power and importance of desires, famously writes: 'Reason is and ought only to be the slave of passions; the only work it can claim to do is in serving and obeying them'. Schopenhauer (1818/1909) dedicated his magnum opus to show how every human endeavor is in relation to desires. The first serious argument in the realm of science, in defining human desires similar to those of animals, was brought by Darwin (1859).

I argue that the representation of desires in our life is through human emotions.[5] The relationship between our daily activities and desires is provided by experiencing emotion. Whether we feel it consciously or being directed by it unconsciously, emotion permeates in every aspect of our lives. It is, therefore, an inseparable part of our cognitive abilities and although our knowledge of mind is limited and uncertain, as many philosophers and scientists agree, I think we cannot imagine the activity of the mind without the constant influence of emotion. Think about any cognitive or non-cognitive activity. What is the purpose or the goal of that activity? If we ignore short-term goals, the only abstract reason which remains is the relationship between the activity and human desires. It cannot be out of two sides: either it is related to survival or to reproduction. We often tend to stop bringing reason for our activities by answering with sentences like: '… because it feels good'. What in the end 'feels good' is precisely our true motive, the motive in our personal situation which leads, in our side, to the fulfillment of desires. Whether the activity is actually leading to a better situation for our

[5] The definition of my usage of emotion in this article is not similar to those that are used in the literature of cognitive theories, feeling theories, or hybrid theories. My usage is close to qualitative feels or emotional qualia. It is an umbrella term and in my view, emotion is a holistic entity that, unlike cognitive theories, cannot be separated into different representations like grief, sadness, joy, etc. Emotion, in my terms, is an inseparable part of our consciousness and we have some indefinable feelings toward everything, whether external or internal. The quality of our experience of emotion with regard to an entity and in a specific situation depends on our biological structure which is both innate and acquired. The function of the emotion, in this definition, is to ensure our survival and reproduction and thus can be viewed as the representation of human basic desires.

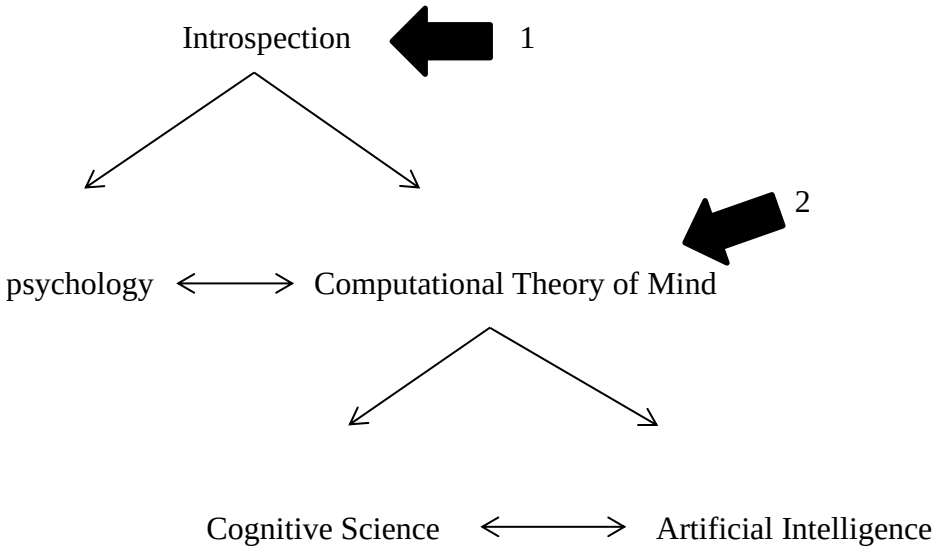
survival and reproduction is a question that we are not easily able to answer. It only ends in the survival of the fittest which is out of our reach. Such 'good feeling' as a motive can also be an indicator of how desires assert their influences in our lives through emotion, through 'feeling good'.

Damasio (1994), in his popular and well-known book brings experimental evidence from neuroscience to show the importance of emotion in our cognitive abilities. His patient, Elliot, was damaged in an area of his brain that was responsible for the experience of emotions and, as a result, lost his abilities in many cognitive tasks, like goal-oriented activities. Damasio concludes that emotion plays a key role in many aspects of our cognitive abilities. A very important conclusion of Damasio, that many writers and researchers who allude to his book seem to ignore, is his idea that the human organism must be viewed as a holistic living organism with interconnected and interwoven parts that cannot be imagined to work without their relation to other parts.[6] From that point, I support my argument that mind and its cognitive capacities cannot be imagined without emotion. The precise way that desires use other mental faculties, like reason, to ensure its survival and reproduction cannot be clearly delineated. The mind, with all of its so-called 'components', should therefore be defined in the context of brain and body; so distinctions like that of emotion and reason, memory and learning, mind and brain, and mind and body are not justified, at least if one is to view them scientifically and with regard to influences from other interconnected components.

An Abstract Look on the Computational Theory of Mind

So far I have discussed two issues, namely the invalidity of introspection and the primacy of emotion in mind. Now I will continue by mentioning how the arguments and evidences in those areas have implications on the assumptions of the CTM. The first part of the present section revolves around introspection. Before any further discussion, I will present an abstract picture of the CTM from an outsider's perspective that will help us derive the fundamental assumptions behind the CTM.

[6] In the penultimate paragraph of the main part of Descartes' Error, Damasio adds: '…the comprehensive understanding of the human mind requires an organismic perspective; that not only must the mind move from a nonphysical cogitum to the realm of biological tissue, but it must also be related to a whole organism possessed of integrated body proper and brain and fully interactive with a physical and social environment.'



As we can see, I think the origin of both psychology and CTM is introspection. The first ideas for researchers in both fields come in their own mind of how different mental faculties work. Two bold arrows indicate assumptions that we will investigate their justifiability. The assumption of the first arrow is that introspection, as the only method to start giving theories of how the mind works, is unquestionably justified. I have tried to show, in the first part of this article, that introspection cannot be a secure source of knowledge of mind by bringing scientific evidence and philosophical reasoning, and therefore cannot be justified as an assumption. The psychology, as a result and in the next step, will go on to support its theories of the mechanism of a certain mental faculty by scientific investigations, of which the most important is, in my view, testing on human subjects. If the theory survived the tests, psychologists would justifiably depend on the inductive reasoning of their scientific test, although we can never be certain about, and can never fully know, the real mechanisms behind those theories. In other words, psychology only sees the two ends of its activity, namely introspective theories and their performance on humans. No psychologist can ever know the precise details of what is happening beyond his/her own mind, since human knowledge of the mind is limited.[7] So, the truth of the assumption of the first arrow, which assumes the introspection as valid, cannot be determined.

[7] In different terms, psychology works between the ends of two black boxes: mind and human subjects. It forms its theories of mind, as the first black box and it tests those theories on human subjects, as the second black box.

In the other side, what researchers in the side of CTM do is, at first, to form a theory of how the mind works and then formalize the theory in terms of computational rules. In forming a theory of the mechanism of a certain mental faculty, researchers in CTM have two options: to form a theory on their own which cannot be experimentally tested and is, therefore, entirely unjustified,[8] or to borrow theories from psychology that were tested on human subjects. In either of those cases, we should not forget that we could not determine the truthfulness of our theories in the first place; neither could we investigate the details of the mechanisms behind those theories.

Now, what CTM takes for granted is, in one hand introspective theories are valid and true, and in the other, we can further investigate the details of what is happening behind. One might be surprised by how many unwarranted assumptions are behind CTM, but there is still another important assumption which is indicated in the picture above by the second bold arrow. In psychology, the story ends with human subjects, but in CTM, researchers enter into the realm of computers to implement the theories of mind into machines. The second bold assumption is that the theories that were formed in the previous step can now be clearly stated and then be translated into the formal rules of computation. Even if we are convinced of the truthfulness of the theories of mind, how can we be certain that those theories can be stated in terms of mathematical and logical steps? To put it differently, in psychology the first end was the theory and the second end, human subject. In CTM, however, the second end does not finish simply and requires detailed knowledge of mechanisms behind the theory that is only a prerequisite for, and does not guarantee, its implementation in terms of computer's

---

As long as the theories hold in human subjects, we have no choice but to accept them, regardless of our ignorance of mechanisms inside the two black boxes. In short, the psychologist does not need to know the details of how theories function. This perspective reminds us of Hume's problem of causality that defined human knowledge of the external world in a similar sense: we cannot know whether two objects or events that are constantly conjoined are really connected or not. We only see the two ends: cause and effect. By their constant conjunction, we have the tendency to form a rule in our mind and as long as the rule holds, we are confident in the truth of the rule, even for future occasions. We tend to accept causality regardless of our ignorance concerning the nature of connection between the cause and effect.

[8] Some cognitive scientists claim that they can form theories of mind, by an interdisciplinary approach, and they do not need to test those theories on human subjects; they argue that they can use computers for experimental testing of their theories. Two difficulties emerge from this methodology. The first is that the conceiver of the theory is subject to solipsistic introspection (as we mentioned earlier in this article) and the second is that since the implementation of theories require detailed knowledge of human mind and mental mechanisms and because our understanding of mind is limited, when we try to formalize them, we inevitably change the nature of theories and as a result, we are no longer testing those theories. We are just testing our computerized translation of those theories.

programming languages.[9] Hence, I argue that even if I am unable to formally prove the impossibility of formalizing our own theories of mind in terms of computational rules, there is no reason or evidence that can prove or even support the possibility of formalization. That, in turn, renders the second assumption unwarranted.

Suppose you have an idea of how you learn a foreign language and the mental faculties involved in the process of your learning. You can talk about your ideas to a friend and he/she will either agree, by ascertaining that the same mechanisms happen in his/her mind, or disagree, by claiming that he/she has a different idea of how him/herself learns a foreign language. The first question is: can you state the details of how you learn a language in terms of algorithmic rules than can be brought into the realm of logic or mathematics? Even if you do, the second question is: how do you reconcile the differences between your own ideas with those of your friend's? This example clarifies two challenges facing the CTM; the validity of introspection, which lies in the differences between the theories of different people, and the possibility of formalizing the theories, which is questioned through the challenge of formalization.

In the second part of the present section I will discuss the primacy of emotions in our mental and cognitive abilities, which was mentioned in the previous section,[10] with regard to its implication on the assumptions of CTM. It is not a new objection against CTM and many writers have claimed that emotions play an important role in our cognition and since machines cannot have our emotional qualia, machines will fall short of human intelligence. I think it does not need further elaboration and is per se a convincing argument which, as I mentioned, is not only supported by philosophy, but by experimental enquiries of scientists like Damasio.

---

[9] Dreyfus (1972/1979/1986) argues that much of what constitutes human intelligence are skills that are rooted in our unconscious instincts; these unconscious skills cannot be formalized. Turing (1950), in defense of formalization, foresaw criticisms and wrote: 'we cannot so easily convince ourselves of the absence of complete laws of behavior ... The only way we know of for finding such laws is scientific observation, and we certainly know of no circumstances under which we could say, 'We have searched enough. There are no such laws.''

The obvious part of these discussions is that, at least, there has not been any proof that can convince us that we are able to formalize whatever humans can think, understand, or learn. Gödel (1931) proved that there can always be statements that a formal system cannot prove, while humans can see the truth of these statements. For further discussions, see Gödel (1931), Lucas (1961), and Penrose (1989, 1994).

[10] For further reading about the importance of emotion in different aspects of cognition and mind in general, see de Sousa (1987), Nash (1989), Lang et al. (1997), Ramachandran and Blakeslee (1998), Schupp et al. (2007), Rubin (2010), Attar and Müller (2012), Bate and Cook (2012).

However, there have recently been some arguments against that claim and I will continue by selecting one of them and investigate its logic.

Megill (2014) argues that although emotional qualia are an integral part of human cognition, they are not essential to the functioning of cognitive abilities and therefore, we can have those abilities in machines that lack emotions. The key question in Megill's article is: 'Are emotional qualia a *necessary* or *essential* component of the performance of human cognitive abilities?' As a response to this question he proposes four possible positions that one can take. He rejects one of the four which says 'cognition cannot possibly proceed without emotional qualia', based 'on empirical grounds'. The remaining three claim that cognitive abilities can be possessed without emotional qualia. He concludes that, as a result, cognitive abilities that we perform can still be performed by a machine.

With regard to his article, a very important issue is raised and its clarification affects not only his point, but many other articles in this discourse. What exactly does he mean by 'human cognitive abilities'? If we look for the answer to this question in his paper, there are two cognitive abilities that are mentioned: facial recognition and selective attention (these two abilities are also presented as 'empirical grounds' to refute the position that he rules out). I think this issue is worthy of further elaboration. In the domain of computers, tasks like facial recognition are performed by a series of algorithmic processes that, in the ground level, are performed by a fast processor and memory. The major job of the processor is to add two numbers and store its result; whatever we perform by a computer is the result of a very fast processor, performing millions of additions in a minute. So, we can claim that, in the realm of computers, every task can be reduced to a series of additions and possible transitions between memory and the processor.[11] Now, instead of facial recognition, we can think of the operation of addition; is adding two numbers a 'human cognitive ability'? In a different perspective, addition, like facial recognition, is an ability that both computers and humans can perform. Before we made computers, was addition a particular cognitive ability of humans? Megill's answer to this question is positive and, looking from that perspective, he would go on to say

[11] Despite many claims, I believe in order to understand the fundamental functioning of a computer, the only important part that one must know is that computers are composed of logical gates that are performing addition in a very fast pace. In higher perspectives (as discussed in the realm of the theories of computation), there are units, instructions, memory, and software that are reducible to logical gates that perform addition and memory that store them.

that if anything else is capable of performing addition, it possesses, at least some of, human cognitive abilities.

I, in contrast, do not think that the answer to the question is an easy one.[12] Humans are capable to perform many cognitive tasks. If we create a machine that can have the same end-result of those cognitive tasks (like facial recognition or selective attention), we are not justified to claim that those machines *have* cognitive abilities.[13] Besides, we should never forget the claim that I made in the first section of this article that, although we can have some ideas about the functioning of our mind, our understanding of how it works and its components is highly limited. We cannot know what is the precise functioning or even definition of human facial recognition or selective attention. So, I doubt if either humans or machines are doing the 'facial recognition'; since we do not have a unified definition of what facial recognition is. Moreover, 'cognitive abilities' are characteristic of humans. By that I mean they are human abilities that happen in human cognition and in human context; assigning the name of those abilities to machines, at least in serious discussions, is simplistic.

In general, Megill bases his argument on the assumption that if anything has the performance of human cognitive abilities, it has, at least some aspects of human intelligence; so emotion is not an essential part of those cognitive abilities. Upon that unproved assumption, he, like many others, establish his arguments.

For that reason, if I am to choose between his four possible positions, I would say 'cognition cannot possibly proceed without emotional qualia'; nor can it be imagined outside the realm of humans.

---

[12] If we exclude the ability of addition and counting in some animals, like crow, it is imaginable that two objects in nature move, for example by wind, to the nearby of another two objects and form the addition of four objects. Does nature have some of human cognitive abilities? Some would argue that the definition of adding two numbers is a mathematical definition and like all other mathematical definitions, it is the product of human thinking; that seems quite convincing but we can still imagine the operation of addition, like many other operations, to happen without any cognitive interference.

[13] The nature of my objection in this section is not new. John Searle's famous Chinese room argument has the same logic: imagine a person who does not know Chinese is in a room with an instruction book that contains answer for every Chinese message. One Chinese person, from outside the room, inserts a message in Chinese from the input. The person inside the room, which does not know Chinese, takes the message and finds an answer from the instruction book and sends it back to the Chine se person. The person outside the room thinks that there should be a person inside the room that understands Chinese that could answer the message. Does the person inside the room understand Chinese? The person inside the room can be replaced with a computer. This argument is also used against the Turing test. For further reading, see Searle (1980).

The Frame Problem

In this section I will shortly discuss the frame problem which has been an important issue in the context of the Computational Theory of Mind. The frame problem is the name of a problem which was first introduced by McCarthy and Hayes in their classic paper, 'Some philosophical problems from the standpoint of artificial intelligence'. It was encountered when AI researchers aimed to design a robot to perform a certain task, but the machine was hung up on the innumerable details of its environment and could not determine and focus on the relevant knowledge from its knowledge-base. It has been, and I believe is still, a technical problem that concerns researchers and engineers whose purpose is to build robots that can perform goal-oriented tasks. Some philosophers however, like Dennett[14], suggest that it is an important epistemological problem and was ignored by philosophy. Dennett believes that it was only by the introduction of AI that we began to investigate how we humans deal with goal-oriented tasks in our everyday life.

I argue that the frame problem is still a technical problem that affects only AI researchers. The people who claim that frame problem is an important philosophical or epistemological problem and those who try to investigate how nature solved the problem for us seem to ignore that they have already assumed the validity of the Computational Theory of Mind, the theory that, as I mentioned earlier, suffers from several unjustified assumptions. They aim to answer the question: how do humans derive relevant knowledge, from their vast knowledge-base, that is relevant to a specific task, in such efficient manner? This question stems from erroneous analogy between human knowledge and information stored in the memory of a computer. I think that in the realm of human life, there is no frame problem and that is precisely the reason that generations of philosophers have never discussed the issue. I argue that we cannot know how we deal with situations in our daily lives; how we ignore irrelevant details and focus on a few relevant ones. That is not a problem for us to investigate; since every investigation faces difficulties and contradictions. For example, some have claimed that Damasio's Somatic Marker hypothesis is the answer why humans do not suffer from the frame problem.[15] They argue that in our life, every detail, option, choice, and nearly everything

[14] Dennett writes: 'a new, deep epistemological problem – accessible in principle but unnoticed by generations of philosophers – brought to light by the novel methods of A.I., and still far from being solved'. See Dennett (1984).

[15] See Megill and Cogburn (2005).

that we do not have a positive emotion towards, is ignored by the mind and as a result, we only focus on the elements that tend to appeal to our emotion. But people who claim this do not explain how exactly we *ignore* those details? In their view, our mind must list all of details of a situation and mark each detail as positive or negative, with respect to our emotions about them; they are not aware that doing so would further put us into the frame problem. The act of ignoring insignificant and irrelevant details of an environment is still *the act of ignoring* them. This solution, like similar solutions, assumes that we can know how we humans solve this problem; the problem which is the result of their faulty analogy. Furthermore, we should not forget that our knowledge of how our minds deal with the details of everyday life is limited and we are probably looking for the solution that does not exist, at least to our understanding.

Most post-AI thinkers, nonetheless, are divided into two groups: those who think that the frame problem is important evidence against CTM (by pointing to human's inability to make robots perform goal-oriented tasks) and those who, in contrast, think that it can even be used as evidence for CTM. I think both sides are equally on the wrong track, but for the sake of brevity, I will shortly discuss the latter group.

Megill and Cogburn (2005) claim that frame problem can be used as evidence for CTM. They start by appreciating the role of emotions in human cognition. Later, they point to Damasio's Elliot whose part of the brain that was responsible in experiencing emotions was damaged and they argue that he actually suffered from an approximation of the frame problem; since he could no longer perform goal-oriented tasks. They go ahead by claiming that if a human who cannot properly experience emotions, suffers from the frame problem, then an emotionless human is comparable with a computer; since they both suffer from the same symptoms. They thus conclude that '… the Computational Theory of Mind is on the right track, but is simply grossly incomplete'.

The goal of Megill and Cogburn's article is to provide evidence for CTM, but if one reads it, he/she will easily recognize that the writers have already assumed the validity of CTM and they write in its language. Elliot did not suffer from the frame problem. As I mentioned, no human suffers from the frame problem; since frame problem has a certain technical definition that cannot be applied to any human. In other words, the fact that Elliot suffered from the lack of emotional qualia and could not perform goal-oriented tasks does not mean that he suffered from the frame problem. Lack of emotional qualia is not a sufficient reason to equate an

individual human with a machine. Besides, even if we are convinced that there are some people who suffer from the frame problem, it means that having emotional qualia is crucial for intelligence or goal-oriented tasks. This leads us to the question: can machines *have* emotional qualia? The article does not discuss this inevitable question and if machines cannot have emotional qualia, then their argument can, inversely, be used against CTM!

Possible Interpretation and Conclusion

At the beginning of this article I discussed the introspection and the unjustifiability of its method. Then I mentioned the superiority of desires, which is supported not only in philosophy but by neuroscience, and their representation in our consciousness by experiencing emotion. Later, I argued that the Computational Theory of Mind is founded on two unwarranted assumptions, namely the truth of introspection and the possibility of formalization. The primacy of emotion in human cognition can also be used against CTM. I also tried to show that the frame problem cannot be used for or against CTM; since it is an issue which belongs to the realm of robotics, not humans.

Now, what are we to conclude from this point of view? It seems that CTM and its two progeny (AI and cognitive science) have the entire picture in the wrong way (by making analogies between computers and minds). Are all efforts in artificial intelligence and cognitive science futile?

My personal answer is ‘no’. The primary reason for the defense on the side of AI and cognitive science can be that both of those fields produce numerous methods and devices that are used widely in science and technology. From novel methods in image processing and recognition in AI to proposing possible ways for machines to learn by induction in cognitive science, we can think of many other instances that both fields can be used to promote technology. For my part, I think there is no fundamental distinction between the scope and nature of discoveries in a technology field, like electrical engineering, and AI. The major difference can be that AI and especially cognitive science inspire from our ideas of how our own minds function.

The goal of this article is to provide evidence and reasoning that machines and human mind are basically dissimilar and CTM is not justified in its analogy between thinking and calculation. Human abilities and the results that those acts can bring, at least to some degree,

have the potential to be performed by machines, although methods in achieving those end results are fundamentally different. From this perspective, however, one important issue is raised: *to what extent* can we simulate human practical performance in machines? In other words, there can be no easy proof concerning how much machines can do what humans are capable of doing. If we ignore human mind and its cognitive details, one would complain to this article and justly argue that the goal of activities that form the viewpoint of AI and cognitive science is and has always been to have *useful capabilities* of humans in machines, regardless of whether machines literally *have* mind, consciousness, or emotion. There does not seem to be a simple answer to this objection and in this respect, another issue rises concerning how we can recognize if a machine truly has a human characteristic. From my point of view, discussions in those issues do not amount to any useful result. What is evident, by what I have presented so far, is that human mind and its details are fundamentally different from machines and computers. As a result and for the sake of pragmatism, we would better at first make it clear what capabilities are useful that we need machines to have. Then it would become easier to build systems that simulate the end results of those capabilities, with no regard for the origin and detailed mechanisms of how humans do them.

References

Attar, C. H., & Mu¨ller, M. M. (2012). Selective attention to task-irrelevant emotional distractors is unaffected by the perceptual load associated with a foreground task. *PLoS ONE, 7(5),* e37186.

Bate, S., & Cook, S. J. (2012). Covert recognition relies on affective valence in developmental prosopagnosia: evidence from the skin conductance response. *Neuropsychology, 26(5),* 670–674.

Damasio, A. R. (1994). *Descartes' error*. New York: Putnam.

Darwin, C. (1859). *On the Origin of Species*. John Murray Publications.

Dennett, D. (1984). Cognitive Wheels: The Frame Problem of Artificial Intelligence, in Hookway, ed. *Mind, Machines and Evolution,* Cambridge: Thyssen Grove Volume, Also in Pylyshyn (1987).

Dennett, D. (1991). *Consciousness Explained*. Boston: Little, Brown, and Co.

Dennett, D. (2001). Surprise, Surprise. *Behavioral and Brain Sciences, 24: 982.*

de Sousa, R. (1987). *The rationality of emotion.* Cambridge, MA: MIT Press.

Dreyfus, H. (1972). *What computers can't do?* New York: Harper and Row.

Dreyfus, H. (1979). *What computers can't do: A critique of artificial reason*. 2nd Ed. New York: Harper & Row.

Dreyfus, H., & Dreyfus, S. (1986). *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. New York: Free Press.

Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. *Monatshefte für Mathematik und Physik, 38:* 173–98.

Heidegger, M. (1926/1962). *Being and Time*. Blackwell Publishers Ltd.

Hume, D. (1738/2004). *Treatise of Human Nature*. Oxford: Clarendon Press.

Kant, E. (1787/1929). *Critique of Pure Reason*. Translated by Norman Kemp Smith. Macmillan and co, Ltd.

Lang P. J., Bradley M. M., & Cuthbert B. N. (1997). Motivated attention: Affect, activation, and action, In R. F. Lang, Simons, Balaban, M. T. (Eds.). *Attention and orienting: Sensory and motivational processes* (pp. 97–135). Hillsdale: NJ: Erlbaum.

Lucas, J. (1961). Minds, Machines and Gödel. *Philosophy*, Vol. 36, Issue 137, pp. 112-127, Cambridge Univ Press.

Megill, J. (2014). Emotion, cognition and Artificial Intelligence. *Minds & Machines*, 24:189–199.

Megill, J., & Cogburn, J. (2005). Easy's getting harder all the time: The computational theory of mind and affective states. *Ratio*, 17(3), 306–316.

Nash, R. A. (1989). Cognitive theories of emotion. *Nous, 23,* 481–504.

Penrose, R. (1989). *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford University Press.

Penrose, R. (1994). *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford University Press.

Ramachandran, V. S., & Blakeslee, S. (1998). *Phantoms in the brain: Probing the mysteries of the human mind*. New York: HarperCollins.

Rubin, D. C. (2010). Emotion and autobiographical memory: Considerations from posttraumatic stress disorder. *Physics of Life Reviews*, 7, 132–133.

Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences*, 3 (3), 417–457

Schopenhauer, A. (1818/1909). *The World as Will and Idea*. Translated by R. B. Haldane, M.A. And J. Kemp, M.A. Kegan Paul, Trench, Trübner & Co.

Schupp, H. T., Stockburger, J., Codispoti, M., Jungho¨fer, M., Weike, A. I., & Hamm, A. O. (2007). Selective visual attention to emotion. *The Journal of Neuroscience, 27(5),* 1082–1089.

Schwitzgebel, E. (2008). The Unreliability of Naive Introspection. *Philosophical Review*, Vol. 117, No. 2.

Turing, A. (1950). Computing Machinery and Intelligence. *Mind, (236),* 433–460.