Man In India, 97 (20) : 305-316

THE CRITICAL ANALYSIS OF FUNCTIONALISMAS THE MODERN VERSION OF MATERILAISM

Nasereddin Hassanzadeh Tabrizi*

Abstract: Functionalism appears to be one of the most modern version of materialism which has dominated the field in cognitive science and the philosophy of mind since 1970s. This sophisticated model of materialism is considered as the theory concerned with the software of thought rather than the hardware. Functionalism as such can typically be associated with physicalism because accordingly, our mental states are implemented in neural stuff but not in a spiritual stuff as the Cartesian model of dualism claims. Such a view of mind which has dominated the philosophy of mind in the last two decades of the twentieth century is known as the computational theory of mind. This paper tries to provide a critical analysis of the notion of artificial intelligence, the computational theory of mind, Allen Turing's Turing Test (2000) and John Searle's Chinese Room Argument. In this respect, the possibility of constructing a completely artificial thinking being made out of wiring, silicon chips, and other chemical things rather than axons, neurons, dendrites, and synopses which also possess mind, consciousness and identity, will be taken into consideration.

Keywords*:* Functionalism; artificial intelligence; physicalism; materialism.

INTRODUCTION

Functionalism is a modern version of materialism which has dominated the field in cognitive science and in the philosophy of mind since 1970s. It is considered to be a theory of mind concerned with the software of thought rather than the hardware. According toChurchland (2006),the core idea of functionalism is the thesis that mental states are defined in terms of their abstract causal roles within the wider information-processing system. A given mental state is characterized in terms of its abstract causal relations to environmental input, to their internal states, and to output. Being in pain, on this account, is a state characterized by its causal relations to behavior such as wincing and crying out, by its causal relations to external input such as the skin being burned, by its causal relations to other internal states such as the desire to make the pain go away, beliefs about the source of the pain and about what will bring relief, and so forth.

So, although functionalism is not a materialist theory per se, it can be considered as a very sophisticated model of materialism and many advocates of functionalism consider themselves materialists as a matter of fact. As Churchland (2006) declares, Functionalists have typically sided with physicalism by claiming that our mental states are implemented in neural stuff, not as the dualists would have it, in spiritual

* PhD, Western Philosophy, Faculty Member of Philosophy Department, Malayer University, Malayer, Iran. *E-mail: daneshjou138@gmail.com*

stuff. At one level of description we can talk about the causal and logical relations among perceptions, beliefs, desires, and behavior, and at the structural level we can talk about spiking frequencies of neurons, patterns of excitations, etc. It is because neurons are orchestrated as they are that the system has the functional organization it does, and thus the physical substratum sub serves the functional super stratum.

The functionalists made the revolutionary change in the philosophy of mind by stating that it is not the question of what is the stuff out of which the mind is made, but the question of what functions the mind actually performs. So, rather than analyzing the nature of the mental substance (material or non-material), the functionalists became interested in the way human mind functions. According toChurchland (2006), plainly, it is not because functionalists suppose that mental states have no material realization. Rather, it is because they envision that types of mental states could have too many distinct material realizations for a reductive mold to fit. As functionalists see it, for a reductive strategy to succeed, a type of mental state must be identical to a type of physical state, but, they argue, the identities are not forthcoming. The reason is that one and the same cognitive organization might be realized or embodied in various ways in various stuffs, which entail that there,cannot be one-to-one relations between functional types and structural types. A cognitive organization is like the computational organization of a computer executing a program.Computational processes are logical, or at least semantically coherent, and they operate on symbols as a function of the symbol's meaning, not as a function of its physical etiology in the machine, and the same program can be run on different machines.

The very reason that the functionalists were not much concerned with neuroscience was that mental states, on their account, are not limited to only one kind of material realization, namely, the neurons. Rather, mental states are like computer program and they can run on different machines. Based on Churchland (2006), in a general way one can imagine that on another planet there might have evolved creatures that, though very different from us in physical structures, might have a cognitive organization much like our own. Suppose, for example, they were silicon- based instead of carbon-based as we are. For these animals, having a goal will be functionally like our having a goal, but such a state will not be identical to having neurons responding thus and so, though to be sure the goal state will be embodied in their physical structure. Or suppose that in time we figure out how to manufacture a robot that has the same functional organization as human; it has goals, beliefs and pains, and it solves problems, sees, and moves about. Its information-processing innards are not neurons but microchips, and its cognitive organizations cannot therefore be identical to a particular neuronal organization, since neural stuff it has not got. Instead, its cognitive economy will be instantiated in electronic stuff.

To do this, the functionalists concentrated on the inputs, the causal relations between mental states and the outputs. The key questions the functionalists were concerned, were the questions like how is an idea or a thought born in the mind? What is the relation of this thought to the other thoughts? And what kind of behavior does this idea or thought lead to?

According to functionalism, what makes the belief that, for example, it is going to rain the sort of thing it is, is that a thought is brought about to my brain by the perception of a large amount of cloud in the sky which is received by my sense of sight which in turn leads into the desire to stay dry, which leads to the physical behavior of looking for my raincoat if I intend to go out. In a similar way, what causes a sensation of pain in my body is that it is damaged by an external cause which brings about some other mental states like pain and anxiety which leads in turning into physical behavior such as wincing, moaning and crying. John Searle the most famous critic of the theory of functionalism describes this view of the theory as follows:

To say that Jones believes that it is raining is to say that he has a certain event, state, or process going on in him that is caused by certain sorts of external stimuli – for example, he perceives that it is raining; and this phenomenon, in conjunction with certain other factors, such as his desire to stay dry, will cause a certain sort of behavior on his part, the behavior of carrying an umbrella. Mental states in short, are defined as states that have certain sorts of functions, and the notion of function is explained in terms of causal relations to external stimuli, to other mental states, to external behavior.

So, it is not only the question of the accessibility to the complicated system of the hardware of the mind or that there is only a little knowledge of the neurobiological mapping of the brain that the theory of functionalism becomes concerned with the software of thought rather than the hardware. Rather, the argument of the functionalists is that neuroscience cannot tell us how the mind functions. Churchland (2006)expresses thatresearch on neurons is not going to reveal the nature of the functional organization, but only something about the embodiment of the functional organization – and just one sort of instantiation at that. Neuroscience, it has been argued, is focused on the engineering details rather than on the functional scheme, and to this extent it is removed from the level of description that is appropriate to answering questions concerning learning, intelligence, problem solving, memory, and so forth. Knowledge of the structural minutiae is important for repairs, of course, and to this extent, neuroscience has obvious medical significance, but structural theory will not enlighten functional hypotheses and functional models.

To put it crudely, neuroscience, on this account, should not be the key-concern of the cognitive science to analyze the nature of mind. No attempt is also needed to reduce human psychology to neuroscience. Instead of being concerned so much with

the physical make up of brain the attempt of the functionalists is now to describe how the physical make up operates. John Searle explains that the functionalists naturally wanted to know what the nature of the inner mental brain states that enabled them to cause behavior was. How did the mental states differ from other sorts of brain states? One answer was to say that this is not really a suitable question for philosophy at all; it should be left to psychologists. We can treat the brain as just a "black box," which produces behavior in response to stimuli and we need not as philosophers worry about the mechanism inside the black box. This view was sometimes called "black-box functionalism." Black-box functionalism is intellectually, unsatisfying in that it does not answer our natural and intellectual curiosity. We want to know, really, how does the system work?

Thus, the mental activities, according to this approach, is not viewed or analyzed by the way what the essence of mind is, but in the way it functions. On this view there remains no mystery as to how minds and bodies are causally related, because, according to this theory, mind is to the brain as the program is to the hardware; our brains are the hardware on which our mental software runs. For a functionalist, brain is just a computer and the mind is the program. All our beliefs, thoughts, desires and feelings can be explained in terms of the program of a computer.

Artificial Intelligence (Computational Theory of Mind)

This view of mind which has dominated the philosophy of mind within the last two decades of twentieth century is known as the computational theory of mind. The computational theory of mind or the computer functionalism comes in two flavors. A brief look at these different positions is in order.

1. The Weak Artificial Intelligence
2. The Strong Artificial Intelligence

Weak Artificial Intelligence

There is the modest version of the Artificial Intelligence which limits itself in the claim that there is a kind of similarity between computer and human brain structure and that the understanding of the mechanism of the computer performance helps us to understand how human brain also functions. This view is considered to be modest because it acknowledges that having devices which perform such mental functions as visual perception, chess-playing or directing other machines does not by itself mean that they also possess minds. According to Heil (1998), the point of the computing machine analogy is not to suggest that we are mechanical robots, rigidly programmed to behave as we do. The point, rather, is that minds bear a relation to their material embodiments analogous to the relation computer programs bear to devices on which they run.

John Searle, the strongest critic of the notion of the Artificial Intelligence or the Computational Theory of Mind also makes the distinction between Weak Artificial Intelligence and the Strong Artificial Intelligence:

It is useful to distinguish what I call "strong" AI from "weak" or "cautious" AI (Artificial Intelligence). According to weak AI, the principal value of the computer in the study of the mind is that it gives us a very powerful tool. For example, it enables us to formulate and test hypothesis in a more rigorous and precise fashion. But, according to strong AI, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind, in the sense that computers given the right program can be literally said to understand and have other cognitive states.

This notion of the computational theory of mind can be traced back to seventeenth century when Thomas Hobbes (1588-1679) articulated the revolutionary idea that 'Reasoning is but reckoning.' (Leviathan, 1651) Hobbes very explicitly considered the whole process of man's thought as a mechanical process. In his History of Western Philosophy, Bertrand Russell (1955) summarizes Hobbes' mechanical view as following.

He considers geometry the one genuine science so far created. Reasoning is of the nature of reckoning, and should start from definitions. But it is necessary to avoid self-contradictory notions in definitions, which is not usually done in philosophy. 'Incorporeal substance', for instance, is nonsense. When it is objected that God is an incorporeal substance, Hobbes has two answers; first, God is not an object of philosophy; second, many philosopher have thought corporeal.

Man's thought and mental activities, on this view, are the product of some unknown incorporeal or spiritual substance whose essence is cogito or thinking. The whole process of mental activity, on this view is seen as the process of adding and subtracting which is built and shaped on the process of bodies in motion.

Strong Artificial Intelligence

With the advent of computer science in twentieth century such a purely materialistic view of human mind which considered all mental processes such as willing, believing, sensation, and feeling as mechanical processes, led to the most influential and appealing theory of mind, named the computer functionalism or the strong artificial intelligence. Mind, according to this theory, is seen as a software program or a digital computer and all mental activities become just a matter of reckoning or computing. Mind only receives data from the external world through the senses; it immediately starts sorting out the received data and manipulates them according to certain instructions. As Searle describes, mental states, on this view, are not defined by any intrinsic feature, rather they are defined by their causal relations, and these causal relations constitute their function. Beliefs, for example, are caused by perceptions and together with desires they cause actions. Such causal relations are

all that there is to having a belief.... Mental states are like carburetors, thermostats, and clocks. They are defined not by their physical structure and not by any Cartesian mental essence; rather, they are defined by their causal relations. A belief is just any entity that, standing in certain relationships to input stimuli and other mental states, will cause external behavior.

Turing Test or Turing Machine

The most famous computing pioneer, Alan Turing (1912-1954), for the first time raised the appealing question, "Can a machine think?" To answer this question, Alan Turing proposed a test which is known as the "Turing Test (2000)." In this test an interrogator puts some questions to (*a*) a man and (*b*) a computer. The interrogator is sitting apart from the two and the communication takes place indirectly through teletype machine. Now after the communication is over, if the received answers from both sides are so similar that the interrogator is not able to determine which answers belong to the man and which answer belong to the computer, then, the conclusion can be drawn that the computer also has a mind. Searle describes the Turing Test as follows:

We need a test that will tell us when a machine is genuinely behaving intelligently, and when it is not. This test was also invented by Alan Turing, and is called the Turing Test. There are different versions of it, but the basic idea is that we can side-step all the great debates about the other minds problem, about whether or not there really is any thinking going on in the machine, whether the machine is really intelligent, by simply asking ourselves, can the machine perform in such a way that an expert cannot distinguish its performance from a human performance? If the machine responds to questions put to it in Chinese as well as a native Chinese speaker, then we would have to say that the machine understands Chinese.According to computer functionalism, a computer and a human brain function in a very similar way. They both possess input and output devices and they receive data through input devices and manipulate and process the received information according to certain instruction and finally decide what to do with the manipulated information. According to Searle, just as the same computer program may be implemented to different sorts of hardware and thus is multiply realizable; so the same mental state, such as the belief that it is going to rain, might be implemented in different sorts of hardware, and thus also be multiply realizable.

A very serious question that must be posed here is that if a device performs an action similar to that of man's behavior, can it be concluded that the device also possesses mind? Or, can it be said that the computational theory of mind plausibly leads to the notion that human being is ultimately nothing over and above a mechanical robot programmed to act in certain way? Why can we not say that the multiple varieties of computer activities such as very professional chess playing, the using of language and controlling of other machines are clear evidence of computer's possession of mind?

Searle's Chinese Room Argument

The response of the critics of the strong Artificial Intelligence to these questions might be that performing various intellectual activities must not lead us to the conclusion that the intelligent devices also must possess minds. The essential difference between human mind and the intelligent devices is that man's intellectual behavior is not pure intellectual, but it involves consciousness, emotions, intentionality, wishes, etc. In other words, human thought is not a matter of implementing a program or a pure syntactic manipulation of information, it is, rather, conscious, thoughtful, meaningful or semantic understanding of data as well.

John Searle, in his controversial thought experiment, known as the Chinese room argument, argues that having a 'mind' is not a matter of performing various high-level intellectual activities or giving correct answer to questions and passing the standard Turing Test, but it is a matter of semantics manipulation of data or a conscious understanding of what is being done. Searle's Chinese room thought experiment is as follows:an argument explicitly directed against Strong Artificial Intelligence was put forth by the present author. The strategy of the argument is to appeal to one's first person experiences in testing any theory of the mind. If Strong AI were true, then anybody should be able to acquire cognitive capacity just by implementing the computer program simulating that cognitive capacity. Let us try this with Chinese. I do not, as a matter of fact, understand any Chinese at all. I cannot even tell Chinese writing from Japanese writing. But, we imagine that I am locked in a room with boxes full of Chinese symbols, and I have a rule book, in effect, a computer program, that enables me to answer questions put to me in Chinese. I receive symbols that, unknown to me, are questions; I look up in the rule book what I am supposed to do; I pick up symbols from the boxes, manipulate them according to the rules in the program, and hand out the required symbols, which are interpreted as answers. We can suppose that I can pass the Turing Test for understanding Chinese, but all the same, I do not understand a word of Chinese. And if I do not understand Chinese on the basis of implementing the right computer program, then neither does any other computer just on the basis of implementing the program, because no computer has anything that I do not have.

In his 'Chinese room' thought experiment John Searle does not intend to disprove that a machine can think, on the contrary, he is of the opinion that human mind is a machine and it obviously can be involved in the process of computation. What Searle very strictly denies is that thinking is just a matter of sorting out of information or it is the manipulation of syntactic symbols. The man in Searle's Chinese room has the capability of giving correct answer to all questions as well as any Chinese native speaker which satisfies the requirements of the Turing Test for being intelligent, but whatever this proves, it does not prove that the man really understood Chinese language; he just manipulated

some formal symbols according to certain rules or instructions. Searle expresses thatI see no reason in principle why we could not give a machine the capacity to understand English or Chinese, since in an important sense our bodies with our brains are precisely such machines. But I do see a very strong argument for saying that we could not give such a thing to a machine where the operation of the machine is defined solely in terms of computational processes over formally defined elements; that is, where the operation of the machine is defined as an instantiation of a computer program.

Searle rejects the traditional contrast between man and machine and considers human being as a peculiar biological machine which has cognition and the unique qualitative states of mind such as understanding, learning and intentionality. On Searle's view, these qualitative features cannot be shared by any other machine and they cannot be realized in terms of computational processes. Searlearguesit is not because I am the instantiation of a computer program that I am able to understand English and have other forms of intentionality (I am, I suppose, the instantiation of any number of computer programs), but as far as we know it is because I am a certain sort of organism with a certain biological (*i.e*. chemical and physical) structure, and this structure, under certain conditions, is causally capable of producing perception, action, understanding, learning, and other intentional phenomena. And part of the point of the present argument is that only something that had those causal powers could have that intentionality.

It seems that the origin of the perplexity or the confusion of the advocates of the Strong Artificial Intelligence is in the fact that they do not have a clear-cut definition of the term *understanding* when they attribute it to a machine. The term understanding like any other term is totally context-dependent and any careless use of the term can be misleading and may cause philosophical misunderstanding.

The term 'understanding' in the context of Strong Artificial Intelligence has been completely misunderstood by the advocates of this theory. And this reminds us of Wittgenstein's famous remark that philosophical issues arise whenever language goes on holiday. The advocates of the Strong Artificial Intelligence have missed a very important point; they have not defined the term 'understanding' definitely. This term is often used metaphorically; we attribute it to inanimate and artificial things but all the same we do not use it in all contexts in the same way. Searle points out the abuse of the term 'understanding' in the following way:

We often attribute 'understanding' and other cognitive predicates by metaphor and analogy to cars, adding machines, and other artifacts, but nothing is proved by such attributions. We say, "The door knows when to open because of its photoelectric cell," "The adding machine knows how (understands how, is able) to do addition and subtraction but not division," and "The thermostat perceives changes in the temperature." The reason we make these attributions is quite interesting, and it

has to do with the fact that in artifacts we extend our intentionality; our tools are extensions of our purposes, and so we find it natural to make metaphorical attributions of intentionality to them; but I take it no philosophical ice is cut by such examples. The through which an automatic door "understands instructions" from its photoelectric cell is not at all the sense by which we understand English.

On Searle's view it is absolutely impossible to attribute the cognitive predicates such as knowing, understanding and learning to program computers. These qualities can be attributed to program computers only in a metaphorical sense, the way we may attribute these two cars, thermostats and adding calculators. But such cognition or understanding is qualitatively different from the human cognition or understanding.

If the sense in which Schank's programmed computer understands stories is supposed to be the metaphorical sense in which the door understands, and not the sense in which I understand English, the issue would not be worth discussing. But Newell and Simon (1976) write that the kind of cognition they claim for computers is exactly the same as for human being. I like the straightforwardness of this claim, and it is the sort of claim I will be considering. I will argue that in the literal sense the programmed computer understands what the car and the adding machine understand, namely, exactly nothing. The computer understanding is not just (like my understanding of German) particular or incomplete; it is zero.

Searle's core of argument in 'Chinese room' thought experiment is that a programmed computer has no capability of understanding in the way that a human being understands. The man in the Chinese room thought experiment that represents the computer model of dealing with questions correlates two kinds of symbols according to certain permissible system of rules without any understanding of the meaning of the symbols. This experiment attempts to show that understanding is not a matter of the exchange of the patterns of symbols; it is, rather, the thoughtful and the semantic involvement of the individual.

Can A Machine Think?

But the question which still remains to be answered here is, 'Is it possible to construct a machine which goes over and above syntactic level of implementing the symbols and reaching the level of semantic understanding of the human being?' Or, 'Can we make a thinking machine?' Searle does not seem to be disagreeing with the possibility of constructing a thinking being made out of wiring, silicon chips and other chemical things rather than axons, neurons, dendrites and synapses, a completely artificial thinking being which possesses mind, consciousness and intentionality. Accordingly, to the question, 'Could a machine think? 'Searle's answer is, obviously, yes. As Searle explains, we are precisely such machines. Yes, but could an artifact, a man-made machine, think? Assuming it is possible to produce artificially a machine with a nervous system, neurons, with axons, and dendrites,

and all the rest of it, sufficiently like ours, again the answer to the question seems to be, obviously, yes. If you can exactly duplicate the causes, you can duplicate the effects. And indeed it might be possible to produce consciousness, intentionality, and all the rest of it using some other sorts of chemical principles than those that human beings use. It is, as I said, an empirical question.

On Searle's view, consciousness in the sense of having thought, understanding and intentionality is just the peculiar quality of human mind and such quality cannot be obtained or created merely by sorting out of data or through implementing information. Searle's argues "But could something think, understand, and so on solely in virtue of being a computer with the right sort of program? Could instantiating a program, the right program of course, by itself be a sufficient condition of understanding?"

This is the right question to ask, though it is usually confused with one or more of the earlier questions, and the answer to it is no; "Why not?"

Because the formal symbol manipulations by themselves do not have any intentionality; they are quite meaningless, since the symbols do not symbolize anything. In the linguistic jargon, they have only syntax but no semantics. Such intentionality as computers appear to have is solely in the minds of those who program them and those who use them, those who send in the input and those who interpret the output.

So, to the question 'Can a computer think?' Searle's answer is, yes – But he emphasizes that it must be special machines, only human brain. This is, for Searle, an empirical question. It is conceivable that in the future there will be a new generations of very sophisticated computers made of silicon chips and wiring which is also capable of semantic understanding. Such a device, modeled on the human brain can think. So, in Searle's view the creation of consciousness and intentionality is a matter of creation of 'semantic understanding'. Create a brain like the brain of human being, and you have created a thinking being, because if you duplicate the cause, you can duplicate the effects.

It seems that a very important point that John Searle, like many others, misses here is that thinking is not the product of brain. It is doubtful that even if a brain exactly like human brain can be modeled in the future, it will function like a thinking organism. Thinking is not, as it is commonly assumed, created in the brain; such a complex process of mental activity must not occur in vacuum. The creation of thought must be the joint product of the brain and bodily activities. Eric Mathews describes thatbrains certainly seem to be necessary for thinking; a being which did not have a brain, and a brain of a particular kind, could not be said to think, as far as our scientific knowledge goes. But notice that it is the *being with the brain* that thinks, not the brain itself. Thinking is not a set of brain-processes, but a human activity.

Mathews also holds that something very important occurs whenever we are involved in the process of thinking and states that whenever we think about something, we are actually considering certain propositions which we express them in sentences of a particular language. One cannot be said to have a thought unless he understands what was meant by the words used in its expression. In that sense, thinking is using language.

On this view any thought or mental activity is as social as it is personal. Thought or any other metal activity does not occur in vacuum and brain cannot be the exclusive source of the production of thought. To clarify this point, Mathews refers to Wittgenstein's argument of the impossibility of using a private language. Mathews' argues the use of language gets its meaning from concepts shared by a whole linguistic community. The use of words with meaning depends on following certain rules which say when it is correct to use the term in question and when it is not. But in a private language, words couldn't be said to be governed by rules, because there would be no possibility of real correction of someone who made a mistake in following the rules.

One might ask how Wittgenstein's private language argument is related to thought. Mathew's point here is thatif thinking involves considering propositions, then having a thought cannot be simply a matter of something going on in private within the individual, whether that something is processes in the soul or brain processes. It seems reasonably well confirmed scientifically that brain-processes must take place in order for us to have a thought; but they are not sufficient. In addition, we have to understand the meanings of the words in which our thoughts are expressed. That understanding in turn cannot be identified with certain brain-processes, because it requires interaction with other human beings in the world outside our brains. Simply put, we need to learn what words mean, and the process of learning requires our use to be corrected by others. The meanings we learn are in that sense shared with others in our linguistic community.

This is the real reason why machines (and indeed brains on their own) cannot think. Machines can, as Searle rightly says, follow syntactic rules; they can order symbols in particular approved ways. But they are incapable of semantics, not because of the materials they are made of, but because they do not participate in a society, in shared activities pursued with a purpose.

References

Churchland, P.S. (2006). From Neurophilosophy in 'The Philosophy of Mind' MIT Press, New Delhi, p. 169-171

Heil, J. (1998). Philosophy of Mind: 'A contemporary Introduction', Rutledge, London, , p. 91

LEVIATHAN, S. I. H. S. (1651). Rebels With A Cause: Self-Preservation and Absolute.

Mathews, E. (2005). Mind: Key Concepts in Philosophy, London, pp. 89-91

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, *19*(3), 113-126.

Russell, B. 1955. History of Western Philosophy, Rutledge, London, p. 534

Searle, R.J. (2004). Mind: A Brief Introduction, Oxford, New York, pp. 44-145.

Searle, R.J. (2006). Brains, Minds and Programs: 'The Philosophy of Mind' MIT Press, New Delhi,

Turing, A.M. (2000). Computing Machinery and Intelligence: Minds, Brains and Computers, University of California, p. 153