

Richard Taye Oyelakin

Obafemi Awolowo University, Nigeria, Ile-Ife,
e-mail: richyman2009@yahoo.com

QUESTIONING THE IMPLIED AUTONOMY OF MACHINE STATE IN PUTNAM'S COMPUTATIONAL HYPOTHESIS OF THE MIND

What appears as the main issue of debate between Putnam's computational hypothesis and Searle's Chinese room experiment is whether or not machine state is sufficient to account for the nature of human mental state. Putnam argues that the nature of machine states is synonymous to the nature of the mental states. For him, it follows that an understanding of the nature of Machine states is adequate to understanding the nature of the mental states. Searle's challenge against Putnam's computational hypothesis is anchored upon the popular Chinese Room Experiment. The experiment shows that it is possible to satisfy Putnam's requirements for having a particular mental state without actually having the mental states in question, although Boden to the contrary. However, the debate is built upon an implied autonomy of machine state. That is, it is assumed that machine state has an independent existence from the mental state. Correspondingly, it is argued that for machine state to be used as an analogy in understanding the nature of mental state, it must be autonomous to mental state. The question which is being engaged in this paper is whether machine state is actually autonomous. For instance, how much can we understand by using the nature of the (mechanical) state of a wrist watch to study the nature of the mental state of the watch maker? The paper maintains that this autonomy of machine states from mental state is questionable. This is because, (1) There is yet no self-created/programmed computer machine, (2) Machine state is created or designed by human mental states and, (3) Only the nature of an autonomous entity could sufficiently be used to study the nature of another autonomous entity. The paper further argues that if an extreme position of the computational functionalism is maintained, then it raises more challenging questions and leads to more complicated problems. The paper, therefore, concludes that the view that intends to use the nature of machine state to study the nature of mental state is circular and the view that equates machine state with mental state is trivial.

Key words: Chinese room, Computational Hypothesis, autonomy, Symbols and Codes, Computer machine, Machine table.

Ричард Тайе Оелакин

Обафемии Аволово Университеті, Нигерия, Иле-Ифе қ.,
e-mail: richyman2009@yahoo.com

Путнамның есептеу ақыл гипотезасындағы күмәнді машиналардың автономиясы

Путнамның есептеу гипотезасы мен Сирлдің қытай бөлмесінде жүргізген тәжірибесі арасындағы пікірталастағы орталық мәселе адамның психикалық күйінің табиғатын түсіндіру үшін машиналық күй жеткілікті ме деген мәселе. Путнам машиналық күйлердің табиғаты психикалық күйлердің табиғатымен синонимдес деп тұжырымдайды. Оның пікірінше, осыдан Машина күйлерінің табиғатын түсіну психикалық күйлердің табиғатын түсінуге адекватты деген қорытынды шығады. Сирлдің Путнамның есептеу гипотезасына қарсы шығуы танымал қытайлық бөлме экспериментіне негізделген. Эксперимент көрсеткендей, Путнамның нақты психикалық күйге ие болу туралы талаптарын нақты сол психикалық күйлерсіз қанағаттандыру мүмкін, дегенмен Боден керісінше талап етеді. Алайда, талқылау машина күйінің автономиясына негізделген. Яғни, машина күйі психикалық жағдайдан тәуелсіз өмір сүреді деп болжануда. Тиісінше, психикалық күйдің табиғатын түсінуде аналогия ретінде машиналық күйді қолдану үшін, ол психикалық күйге қатысты автономды болуы керек деген пікір айтылады. Бұл мақалада машинаның күйі шын мәнінде автономды ма деген сұрақ қарастырылады. Мысалы, қол сағатының (механикалық) күйінің табиғатын сағат жасаушының психикалық күйінің табиғатын зерттеу арқылы не түсінуге болады? Құжатта машиналық күйлердің психикалық жағдайдан тәуелсіздігі күмән тудыратындығы айтылған. Себебі (1) өздігінен құрылған бағдарламаланған компьютерлік машина әлі жоқ, (2) машинаның күйін адамның психикалық күйлері жасайды немесе салады және (3) басқа автономды құрылымның табиғатын зерттеу үшін автономды құрылымның табиғатын

ғана жеткілікті дәрежеде қанағаттандырады. Мақалада әрі қарай есептеу функционализмінің шектен тыс позициясын қабылдау неғұрлым күрделі мәселелер тудырып, күрделі мәселелерге әкеліп соқтырады деп тұжырымдайды. Сонымен, автор психикалық күйдің табиғатын зерттеу үшін машиналық күйдің табиғатын пайдалануды көздейтін көзқарас айналма болып табылады, ал машинаның күйін психикалық жағдаймен теңестіру көзқарасы тривиальды деп тұжырымдайды.

Түйін сөздер: қытай бөлмесі, есептеу гипотезасы, автономия, шартты белгілер мен кодтар, компьютерлік машина, машиналық үстел.

Ричард Тайе Оелакин

Университет Обафемі Аволово, Нигерия, г. Иле-Ифе,
e-mail: richyman2009@yahoo.com

Сомнительная автономия машины в вычислительной гипотезе разума Патнэма

Главный вопрос в спорах между вычислительной гипотезой Патнэма и экспериментом Сирла в китайской комнате заключается в том, достаточно ли машинного состояния для объяснения природы психического состояния человека. Патнэм утверждает, что природа машинных состояний синонимична природе ментальных состояний. По его мнению, из этого следует, что понимание природы состояний Машины адекватно пониманию природы ментальных состояний. Вызов Сирла против вычислительной гипотезы Патнэма основан на популярном эксперименте с китайской комнатой. Эксперимент показывает, что можно удовлетворить требования Патнэма о наличии определенного психического состояния, фактически не имея этих психических состояний, хотя Боден утверждает обратное. Однако дискуссия строится на подразумеваемой автономии состояния машины. То есть предполагается, что машинное состояние существует независимо от ментального состояния. Соответственно, утверждается, что для использования машинного состояния в качестве аналогии в понимании природы психического состояния оно должно быть автономным по отношению к психическому состоянию. В этой статье рассматривается вопрос, действительно ли состояние машины автономно. Например, что мы можем понять, используя природу (механического) состояния наручных часов, чтобы изучить природу психического состояния часовщика? В документе утверждается, что эта независимость состояний машины от психического состояния сомнительна. Это потому, что (1) еще не существует самозданной запрограммированной компьютерной машины, (2) состояние машины создается или конструируется человеческими ментальными состояниями, и (3) только природа автономной сущности может быть в достаточной степени использована для изучения природы другого автономного образования. В статье далее утверждается, что если придерживаться крайней позиции вычислительного функционализма, то возникают более сложные вопросы, ведущие к более сложным проблемам. Таким образом, автором делается вывод, что точка зрения, согласно которой необходимо использовать природу состояния машины для изучения природы психического состояния, является круговой, а точка зрения, согласно которой состояние машины приравнивается к психическому состоянию, тривиальна.

Ключевые слова: китайская комната, вычислительная гипотеза, автономия, символы и коды, компьютерная машина, машинный стол.

Introduction

Putnam's computational hypothesis of the mind specifies that the nature of machine states could be used to study the nature of the machine states. This is exemplified in the analogy of Turing machine. The hypothesis claims that at the fundamental level of description, an appropriately programmed machine is a mind. But, Searle challenged this position by deploying his popular Chinese room experiment. This experiment demonstrates that Putnam's hypothesis might be necessary about the nature of mental states but not sufficient. The reason is that it is possible to satisfy the requirement of the hypothesis without having the mental states in question. This is

because while computation is syntactical cognition is semantical.

However, Boden, challenged Searle's submission against Putnam's hypothesis. For Boden, nothing differentiates the digital symbols and codes from linguistic symbols and codes. For her, both comprises synonymous characteristics. In that wise, what is called meaning is arbitrarily conferred on codes and symbols depending on circumstances and or conventional needs. This leads to the claim, contra Searle, that the nature of machine state is sufficient to study the nature of the mental states. Debate between Putnam and Searle is based upon the question whether the nature of machine state is sufficient to study the nature of mental states. It is

argued in the paper that machine state could not be used to study the mental state except the former is autonomous to the latter. However, in the debate between two scholars is an implied autonomy of machine states. The paper questions this assumed autonomy of the machine state and then raises an issue on whether machine state is actually sufficient in the study of the nature of mental states. The paper argues that this assumption is faulty in the sense that machine state is a product of mental states. It then, becomes superfluous using the nature of machine states to study the nature of mental states. This is because, largely speaking, it is based on the presumption of what is exactly at issue.

Putnam's Hypothesis and Searle's Chinese Room Experiment

What Putnam's hypothesis (Putnam, 1975: 429-440), argues is that the nature of mental state is determined by its causal relations to stimulus input, behavioural output, and corresponding mental states, as specified by the Table of instruction. It is argued that this process is computable by any Turing Machine. This is characterized by the analogy of Turing machine (Turing, 1950: 433-460, Boden, 1990: 40-46). The claim, therefore, is that whatever is constitutive of the mental states is nothing over and above and it is equivalent to the description of the nature of the machine states. It then follows that there is no significant different between the nature of the mental states and machine states. The hypothesis largely pushes the position that machine state could be used to study mental state. In fact, Putnam's hypothesis suggests a synonymy between the two. This then means that for Putnam, syntax and semantics are, strictly speaking, indistinguishable.

The Chinese Room Experiment (herein CRE) is a direct attack on the claim that thought can be represented as a set of computable symbolic functions. Searle describes an hypothetical person (Searle-in-the -room) who only speaks English language. He is in a room with only Chinese symbols in baskets and a rule book written in English for moving the symbols around. The Searle -in -the room is then ordered by some Chinese-out- of- the- room to follow the instruction in the rule book in order to send certain symbols out of the room when supplied with certain Chinese symbols. Furthermore, the Chinese speakers communicated with the Searle- in -the- room via the Chinese symbols. The experiment supposes again that Searle-in -the-room was able to send correct Chinese symbols out, as answers, by

following the instructions written in English. It did appear that Searle-in -the-room understand Chinese language. For Searle, the exercise on the juggling of the Chinese symbols in the room consists of pure syntactic process. Therefore, according to the experiment, it would be absurd to claim that the English speaker (Searle-in-the -room) understands Chinese language simply based on these syntactic processes.

Searle raises two main arguments from the experiment. The first is that it is possible to satisfy Putnam's computational hypothesis and not having the mental state in question. The Searle- in- the-room only has the syntax of the Chinese symbols and not the semantics, although he was able to produce correct answers to the questions. The knowledge of semantics of the language differentiates Searle-in- the- room from native Chinese speakers. While Searle- in- the- room has the syntax of the symbols, he does not possess their semantics.

Searle argues that;

The limitation was corrected by computer functionalism to the extent that it at least specified a mechanism: the computer program that mediated the causal relations between the external input stimuli and the external output behavior. But the difficulty with that theory is that the program is defined purely formally or syntactically, and consequently does not, qua program, carry the intrinsic mental or semantic contents that human mental states actually have, (Searle, 2008: 60).

Searle seems to be making a prima-facie distinction between machine state and mental state here. This is arguing that the nature of the machine states is syntactic. It only consists in specifying the structural arrangements of the digital codes and symbols used in the computation based on some certain recursive rules or instructions. For him, this structural process obviously lacks the intrinsic meaning or semantic content of the codes and symbols involved. Correspondingly, Searle argues that this is what differentiates a human being from a computer machine. Whereas a machine state consists of syntactic process, mental states, in addition and more importantly, consists of the semantic content of the codes and symbols. The question of how either the machine or the mind converts information to digital or mental symbols and codes is raised by Anderson (2013)

For instance, in the activity of number addition; it is supposed that the machine does not have the meaning or thought of or *about* the numbers. It is incapable of independently conceptualizing numbers in various ways or raising the perennial question

about the possibility or otherwise of the ontology of numbers beyond the way it is programmed. It only adds in accordance to the appropriate table of instruction. But, not only are the human beings able to add these numbers, questions about the meaning and ontology of these numbers are parts of such mental phenomenon, again so it is supposed. Besides, it is also opined by some biological naturalists that there is *something it is like* to know that $2+2=4$. This refers to the subjective or phenomena experience. It is argued that this is limited to organism with electro-chemical organic properties which are only found in animals. The main point, contrary to Putnam's assertion, is that syntax is not semantics, (Searle, 2008: 68). It is argued that computational hypothesis is purely syntactical while mental states is both syntactical and semantic in nature. Machine states, therefore, is insufficient to account for the nature of mental states.

The second point is that computational hypothesis only attempts to simulate cognitive and mental capacities. But according to Searle, "simulation is not duplication." (Searle, 2008: 68). Simulating a particular phenomenon is like imitating the phenomenon. Machine states hypothesis is simulating mental states in the sense of artificially programming a system to demonstrate human cognitive capacities. A calculator is artificially programmed to demonstrate computer machine's arithmetic capabilities. This is done to show that human arithmetical cognitive capabilities are computationally or mechanically demonstrable. However, the description of a calculating process in a calculator is not synonymous to the human cognitive ability in calculation.

It is clear from Searle's argument that, in this case, imitation cannot be the duplicate of the original. There are however, some things which might be successfully simulated or imitated. Indeed, it is possible to simulate digestion, rain storms, arithmetic abilities, and so on. Anything which is capable of precise definition may be successfully simulated. But, for Searle;

it is just as ridiculous to think that a system that had a simulation of consciousness and other mental processes thereby had the mental processes as it would be to think that the simulation of digestion on a computer could thereby actually digest beer and pizza, (Searle, 2008: 68).

The point is that it is implausible to think that the simulation of a phenomenon or an event is the real phenomenon or the event. For instance, it is faulty to think that the accident which occurred in a movie is actually real. The claim is that machine state is just

a simulation of the mental state and that it cannot be *the* mental state. It is even apparent that a simulation inherently distinguishes original from the imitation. This is because, "to simulate" means there is something original to simulate and must be different from its imitation. It is impossible to simulate an inexistent phenomenon. For Searle, the only means to arrive at the mental states is to duplicate it and not to simulate.

You would have to duplicate, and not merely simulate, the actual causal powers of human and animal brains. There is no reason in principle to suppose that we would have to have organic materials to do this, but whatever material we use we have to duplicate the causal powers of actual brains, (Searle, 2008: 62).

This means that for a system to duplicate mental states, it must be such that it possesses the right sort of properties with which to duplicate the causal powers of the brain. The point which comes out of this analysis is that for the nature of mental states to be accurately and adequately accounted for, the human organic system has to be duplicated. The computational hypothesis fails, as an account of the mental states, because it is just a simulation and not a duplication of mental state. It should be noted the issues here is mainly about the plausibility of using the nature of the machine state to study the nature of the human mental state. This is not a question of the possibility or otherwise of human mental states being programmed computational machines. Whether or not this is the case is another critical and metaphysical issue to be separately addressed.

Searle's Distinction between Syntax and Semantics

What is apparent in the CRE is that there appears to be distinction between syntax and semantics. This opinion is vividly shared by Ned Block. (Block, 1993: 819-831) Searle identifies this distinction as a major challenge against computational hypothesis. One of Searle's, (Searle, 2008: 70) main arguments against Putnam's hypothesis is that computation is about mere syntactical description. Syntactical description is about the structural arrangements of the symbols, codes, or sentences as specified in the machine table. It is not about the semantic of the mental content. Block's summation of the issue might be instructive;

At the most basic computational level, computers are symbol-crunchers and for this reason the computer model of the mind is often described as the symbol manipulation view of the mind, (Block, 1993: 828).

Semantics deals with the question of meaning of terms, concepts and sentences, and how this is determined. However, a pressing semantic question is; how is meaning determined? But, this is a question about meaning which is only raised in semantics but not by any digital automaton whose main concern is only dealing with syntax. For computational hypothesis to account for the mental states, it must be able to account for the semantics of mental terms, concepts, and sentences. For Searle; "The program by itself is insufficient to constitute mental states because of the distinction between syntax and semantics", (Searle, 2008: 70). It is notable, according to Searle that, just as Searle-in-the-room, all that is done in the computation is mere structuring, arrangement and re-arrangement of digital codes and symbols. These activities are constrained within the principles of syntax. In support of this point, Searle argued that syntactical knowledge does not guarantee semantical knowledge. According to him;

The program by itself is insufficient to *constitute* mental states because of the distinction between syntax and semantics. And it is insufficient by itself to *cause* mental states because the program is defined independently of the physics of its implementation. Any causal power the machine might have to cause consciousness and intentionality would have to be a consequence of the physical nature of the machine. But the program *qua program* hasn't got any physical nature. It consists of a set of formal, syntactical processes that can be implemented in the physics of various kinds of machinery, (Searle, 2008: 70).

Searle's point might be understood in two distinct but correlated senses. First, computational functionalism is inadequate as an account of the nature of mental states because it is abstractly formulated independent of the physical structure of the implementing system. For him, for computational hypothesis to sufficiently account for the nature of mental states, the account must be in conjunction with the account of the nature of the implementing physical structure. In the case of Putnam's machine structure, the account of the nature of machine states does not include the account of the implementing physical structure. In fact, this problem is made even complex by Milikan when he says "So long as people assimilate studies of consciousness to studies of phenomenal experience, they are side stepping the real issues", (Milikan, 2014: 13)

Second, the syntactical knowledge is distinct from semantical knowledge and one does not presuppose the other. Computational hypothesis is

syntactical in nature and by this it consists of formal processes which can only be implemented in the physical structure of different kinds of machines. Mental states, on the other hand, possess semantical properties. The reason is that an adequate account of the nature of mental states includes the account of the physical structure of the implementing system. This enables the system to generate mental states. That is why whereas different and many physical substrates could implement a particular logical or computational function, in the strict sense, only one electro-chemical organic system could realise a particular mental state. This is the point Putnam made by deploying *computational plasticity*, (Putnam, 1988: xiv). Therefore, computational hypothesis is not able to account for mental states. It is too restrictive.

For instance, let us compare these two statements of belief; (1) "Ade loves his parents", and (2) "Blgars grears his gerondo." In matters of structure, there may be no controversy that the two statements are syntactically the same. Both of them consist of subject and predicate. Both of them satisfy required grammatical rules for a standard sentence. But the identifiable problem is about the meaning of the second sentence. Whereas the first sentence makes a clear conventional sense in terms of meaning, the second, in this sense, does not. Why the first sentence makes a sense is that there are corresponding environmental evidences which the statement refers to. The point made in this example is that syntactical equivalence does not guarantee semantical equivalence. For Searle, it is impossible to generate semantical content from mere abstract computational process. This is because "There isn't any such thing as understanding in addition to symbol manipulation, there is just the symbol manipulation." (Searle, 2008: 69). This is the reason for the conclusion that computational process lacks this semantical content. Whereas once there is a good semantic mastery of symbols and codes, syntactic questions get automatically settled. For instance, it is impossible to know the meaning of the terms "elephant" and "rat" and then maintain a syntactic arrangement; "The rat swallows the elephant." Except it is being used idiomatically, the syntax betrays the semantics.

In syntactical structure, environmental facts do not have any influence and therefore, meaning is not involved. For instance, the logical validity of (P → Q) does not necessarily involve the meaning of either P or Q. In other words, whatever P or Q means has no influence in the validity of the judgment of the rule. This is underlined by Newell and Simon in

“Logic, and by implication all of mathematics, was a game played with meaningless tokens according to certain purely syntactic rules.” (Newell and Simon, 1990: 112). Whereas the understanding of the mental concepts carries with it the meaning of the terms used. Therefore, machine states which relies on mere syntactical structure are not sufficient as an account of mental states.

Searle claimed that computational hypothesis muddled up the difference between syntax and semantics. The hypothesis presumes that syntax is sufficient for semantics, that is, “The symbol manipulation is all there is to understanding.” (Searle, 2008: 70). But, what produces semantics is more than mere abstract code, and symbol manipulation. Here is the correlation; whereas the process of symbol manipulation could be abstractly described independent of the implementing physical system, this is not true of semantics. The nature of mental states combines both the syntactical and semantical underpinnings. Every meaning is attached to point of view. Every point of view is attached to a set of environmental facts. The point of view must belong to some agents. For instance, if I say “x is white”, I won’t be making any sense until some physical or empirical facts come into play in the description. In addition such as understanding requires a peculiar organic system. However, as Searle argues, computational hypothesis faces a difficulty because it abstracts syntax away from the physical nature of the implementing organism. It is argued that semantics, which deals with the meaning of sentences of propositional beliefs, is caused by neurobiological processes in the brain. Therefore, “Any causal power the machine might have to cause consciousness and intentionality would have to be a consequence of the physical nature of the machine.” (Searle, 2008: 70). But, machine states hypothesis denies this.

Searle, however, further identifies that at some levels of description processes in the brain are syntactical. This is because, essentially, sentences and symbols are also variously and differently arranged in the brain. For Searle, “there are so to speak, “sentences in the head.”” (Searle, 1993: 836). What this means is that there is the structural arrangement of sentences in the brain. For instance, the sentence “I will go there tomorrow” has a peculiar syntactical arrangement which together with some other factors determines its meaning. “I go will tomorrow there” will not provide a good syntactical arrangement. Hence, it will distort its meaning. To the question; how are sentences come to be so arranged, could simply be traced to a conventional

exercise which is committed to be done in the head, again, influenced by relevant environmental factors. But, this structural arrangement is combined with the awareness of the meaning of the sentences. Scholars such as Ned Block realized that thought process is done through a combination of syntax and semantics. For him, “When it finds a match, it sends a signal to a third component, whose job it is to retrieve the syntactic and semantic information stored in the dictionary”, (Block, 1993: 819). For instance, “it rained, therefore, the ground is wet” combines the two notions viz; syntax and semantics. The difference between “(P → Q)” and “if it rains, then the ground is wet” is in the meaning of the terms involved in the second proposition combined with its syntax. (P → Q) may be variously interpreted, but “if it rains, then the ground is wet” maintains a particular and definite interpretation and meaning. Again, consider the following notations. Abstractly, if “F = G” and “G = H”, then “F = H”, where the notation = is used to represent identity, sameness or equivalence, without the need to understand what the symbols represent or their meanings. Of course, the validity of the inference still follows if the symbols are represented by any statement. That is an example of a syntactic arrangement devoid of its semantic content.

Furthermore, thought is possible because these sentences consist of words whose meaning is clear to the thinker or speaker. In a situation where the meaning of words is not clear, thinking is impossible. Whereas syntax only deals with the taxonomy principles (classification) of these sentences, this is not sufficient to account for the nature of mental states. Note that I would have communicated nothing, if instead of saying “if you read very well, then you will pass your next exam”, I only exclaim “if P then Q” which may preserve various interpretations, including my own. However, machine state hypothesis is limited to structuring and arranging codes, symbols and sentences in some specified way to produce an outcome. This cannot produce mental states. The nature of mental states includes having the knowledge of the meaning of sentences, codes and symbols. Therefore, this, in essence, shall involve being able to define terms and concepts in relation to environmental fact in order to define the meaning of a particular concept or term.

This discussion therefore centres on two germane questions; the first is the question concerning how the brain works in structuring these sentences. The second is how these sentences in the head get their meanings. The first question has to do with syntax

while the second question deals with semantics. For syntax to guarantee semantics, "it has to have a meaning or semantic content attached to the symbols." (Searle, 2008: 60). This is because, syntax or computation or symbol manipulation in itself is not constitutive of, nor sufficient for, thinking because it is defined entirely syntactically and thinking has to have something more than just symbols. But computational description is only constitutive of mere symbol manipulation which is devoid of semantic attachment. What goes on in the mental states are more than mere symbol manipulation. When I express a sentence such as "this idea come from my brain," there is something which I know that I express. Beyond that, there appears to be the feeling of awareness of my expression. Thus, there is something that machine states hypothesis lacks which makes it unable to account for the nature of mental states. These set of points leads to the conclusion that machine hypothesis fails to assume that the nature of the mental states is synonymous with the nature of machine states. As we proceed to Boden's reaction, it is important to observe a heavy dependence on human singular point of view and judgment in Searle's objection. But, there, obviously, is a question about the authenticity of human singular point of view in assessing nature. We may need to pass this over for now.

Boden's Argument against Searle – in -the – Room Experiment

Margaret Boden argues to show that there is nothing in the nature of mental states that machine states hypothesis does not and cannot account for. For her, Searle's hypothesis fails to support the point that syntax cannot account for semantics, (Boden, 1990). For instance, there is a measure of cognitive understanding of language by Searle-in-the-room. Meaning that the man in the room certainly understands the rules which were written in English. He was able to arrange the Chinese symbols by his understanding of the instruction written in English language. This means that right in that room, meaning is defined and something is understood. This further means that even though computer may not have a cognitive understanding of external symbols and stimulation, it has a cognitive understanding of its own program language through which it manipulates other symbols. This shall make a position that computers also possess an understanding of the semantics of its own language. The reason is that, the instruction table defines the meaning of concepts and or symbols used in

statement of belief. But, does and can computer possess independent awareness outside of the way it is programmed? Boden seems to be silent about this.

Boden argues that natural language behaves like un-interpreted symbols and codes. It is the already in-built capacity (program) which enables the possibility of what is called conventional definition and cognitive understanding. These codes and symbols do not have their independent meanings outside conventional definition. For instance, "two plus two equals four" has no independent meaning outside its conventionally stipulated definition, hence are abstract codes and symbols. This point is reinforced by Huttenlocher, (Huttenlocher, 1973:174). For him, simply because mental states are like automaton states in this regard, the illustrated method for defining automaton states is supposed to work for mental states as well. In other words, the mind and machine only manipulate coded and un-interpreted symbols. Whatever method works for one is expected to work for the other. This is because mental state and machine state run the same functional organization as specified by an appropriate psychological theory or machine table.

However, what becomes clear in the above is that Searle in the Chinese room experiment obviously assumes that meaning of concepts and terms used in belief fixation is determined. All that is been shown by Boden's argument is that Searle in the room experiment is inadequate to show that a computational hypothesis is unable to sufficiently account for the nature of the mind. It follows that computational hypothesis is still a plausible account of the mind. From Boden objections, Chinese room hypothesis may not be adequate to show that machine states are different from mental states unless it is able to show that language codes and symbols have their inherent semantics apart from the conventional labels.

Questioning the Implied Autonomy of Machine State

Running through this popular debate, however, there is an implied assumption that machine state is a sufficient phenomenon to be used to study the nature of mental state. To say the least, it is assumed that machine state is different from mental state. But does this difference imply autonomy? In this debate, the autonomy of machine states is inherently and apparently implied. This section intends to questions such as; Is machine state autonomous from mental states? In the strict sense, is machine state sufficient for the study of the nature of mental

state? To address these questions, we shall assume a premise based on the existing facts;

(1) There is yet no self-created/programmed machine, (Oyelakin, 2019: 140).

However, there have been recent technological advancement leading to self-improving computers. Note that self-improving, (Barrat, 2013: 7, 99, 176), is not the same as self-creating/programming as the case may be. If this premise is granted then it follows that digital/ computational machines are not self-created. If they are not self-created, then they must be a product or creation of some ingenuity. The existing facts attest to the point that the digital machines are products or creations of human cognitive capacity. Human cognitive capacity is a property of human mental states. It may then be safely affirmed that human mental states created or programmed machine states. That is; machine state is a creation of human mental state. Now we have another premise.

(2) Machine state is created or designed by human mental state.

Now, then next is to understand what it may mean to say that one entity is autonomous from the other. What does it mean to say that one object or entity is autonomous from the other? Merriam-Webster's Online Dictionary defines autonomy in terms of self-existence, self-containment, or independent from external control. Following from this, an object or entity is autonomous just in case it is an atomic entity capable of being self-existent, self-contained, or self-determination, existing independent/ without external control. The question is whether machine state and mental state enjoy such autonomy from each other. From the first and second premises, it is evident that while mental state is capable of being autonomous from machine state, the latter is not yet capable of being autonomous from the former. From the following analysis, I maintain that only through the nature of an autonomous object or entity could the nature of another autonomous object or entity be studied. Otherwise, what sense would there be to suppose that the state of Ford automobile engine is used to study the state of the mind of Henry Ford? Could there be the state of the engine without the state of the mind of Henry Ford?

For instance, the following are examples of autonomous entities; Man, and Shark, Whale, and Oak tree, Bat, and Bird, Whale, and Shark, Oxygen, and Hydrogen. The list appears endless. Each of these is autonomous from the other. Now, an attempt to use the nature of the mental states of Whale to study the nature of the mental states of Bat may make a good sense, howsoever that enquiry is

to proceed. Same for each of the autonomous pair listed. This is because, nothing in the nature of each is in any way connected to the other. Then,

(3) Only the nature of an autonomous entity could sufficiently be used to study the nature of another autonomous entity.

It must be clearly stated here that this is only the sense in which the nature of a particular entity may be used to study the nature of another autonomous entity. Once there is any form of connection between two entities which affects their autonomy, then it becomes superfluous to attempt to use the nature of one of them to study the other. For instance, it is superfluous to attempt to use the nature of the (mechanical) state of a wrist watch to study the mental state of the watch maker. It apparently amounts to studying the nature of the mental states of the watch maker. The watch is a creation of the watch maker and is not autonomous from it.

Now, let us tie the arguments together. Premise 1 is strongly assumed. Premise 2, by the existing fact, subsists on premise 1. Premise 3 subsists upon the assumed autonomy between the machine state and the mental state. From premises 1, 2 and 3, we can deduce the claim that it is not the case, given the understanding of autonomy, that the nature of machine state is autonomous from the nature of mental state. Therefore, from (3) it is inadequate to attempt to use the nature of machine state to study the nature of mental state. The fore-going analysis, again, may lead to another debatable point. It may prove that it is improbable that Nature allows the exercise of using the nature of an entity A to study the nature of another autonomous entity B, though human indulges himself in this. For instance, how much of the nature of the mental state of Bat has man actually understood? Nagel's (1979) question "What is it like to be a bat?" is germane here. Anyway, we have understood that which our nature only permits us to understand and nothing more. But, are what we understand all there is to be understood in the nature of the animal? Besides, are bats also studying the nature human beings? The intuitive response to the second question is a "No", based on the human perceived position of self. But, I presume that the correct answer to both questions, in this case, should be "We don't know."

However, it may be objected that computational functionalism, at its radical form, actually identifies an identity between a machine state and mental state. Therefore, the claim of superfluity only addresses the position of the moderate sense. The radical position seems to establish an equivalence between machine state and mental state. This, again, is a very

questionable position though it requires carefulness to address it. It raises a question whether the nature of machine state exhausts all there is to be the nature of mental state. As it has been asserted by some scholars, human mental nature is so complex to the extent that man himself has not fully understood what it is, (Churchland, 1993: 745). First, due to this complexity of mental state, there are evidences to support the claim that computational property is just a component of the mental state. Computational nature does not exhaust the nature of mental state. (1) Human advancement in science and technology underlines this point. There have been and will still be new discoveries and inventions even in the digital computer world which have nullified and could nullify some of the present assumptions respectively, which actually could have been impossible should machine state exhausts mental state. (2) In programming computers, there are some non-computational/digital capacities required. For example, there is the evaluative capacity which is largely required and demonstrated by man and this has not been sufficiently built into the machine state. There is wide gap of difference between "true or false" and "right or wrong." Machine is a strict rule following automaton even in the true/false identification. (3) Human advancement in non-computational aspects also strengthens the point. There are inventions in every aspect of existence ranging from medicine, arts/humanities, technology, mechanics, abstract and social theory constructions, and so on, yet to be made, which when made may lead to upgrading the current nature of the machine states. While machine state may be upgraded, mental states is just discovering or revealing itself. No scientist has been able to upgrade human cognitive system. Plato's "knowledge as recollection" lends credence to this claim. So, to claim that human nature is only computational is, apparently, to screw it.

Second, equating machine state with mental state implies that the limits, weaknesses and deficiencies of machine state is *ipso facto* true of mental state. But, is this true? The point is that the limits, weaknesses or deficiencies of computer machine is not the limits, weaknesses or deficiencies of human mental states. This is strengthened by the fact that the weaknesses and or deficiencies of computer machine still depends upon human mental capacity for correction. It is still impossible for computer machine, by whatever means, to rectify the deficiencies of human mental state, though bioethicists are making some recent, but debatable assumptions. For instance, how can machine state handle the popular objection, on the possibility of

pretense, against behaviourism? From instance, if I address a colleague with the following statement and mean it; "I want to tell you that I hate you" but while smiling, how is the machine state to handle the weakness? It is only man that can device a means of programming computer to handle that. Computer cannot perform what it is not programmed to perform. Man still retains the power to willfully shut any malfunctioning or erring machine state down, whereas it hasn't got to the point when machine can willfully shut human mind down except as it is programmed by man. Well, it may be objected that the fact that this has not occurred does not mean that this may not be the case in the nearest future. Again, this may be possible when machine state can duplicate mental state instead of the present simulation.

The third is the absurdity which is implied in the assumption. Doesn't it appear absurd to attempt to use the nature of human mental state to study the nature of the entire cosmos? Human mental state is just a token component of the cosmos. Apart from the property of being human, there are countless other properties constituting the cosmos, even those that may never be known or understood by man. Besides, wouldn't amount to using the nature of man to study itself? It is then absurd to argue that the nature of human mental state is the same as, all there is to the nature of the cosmos and that an understanding of one implies an understanding of the other. How impoverished would such a thought be? The same goes for saying that the nature of mental state exhausts the nature of God, for those who believe in God. In view of Royce, (Royce, 1968), man is just a component of the Absolute. The same absurdity goes into the attempt of using the nature of machine state to study and understand the nature of mental state of its creator. On a lighter mood, should the computer be conscious, it would have declined such a fruitless venture. The point which is made is that the claim that machine state is equivalent to mental state is rather too exclusive. It forcefully defines human mental state by and restrict it to just its component part. Whereas, on the list of what the nature of mental state is, the property of computation is just an item.

Conclusion

The paper identified that the bone of contention between Putnam and Searle is whether machine state is sufficient to study the nature of mental state. The main finding of this paper is that; (1) to be qualified for such study, machine states must be autonomous. (2) Searle's

intriguing challenge against Putnam's computational hypothesis, through the Chinese room experiment, is based upon an assumed autonomy of the machine state from mental state. The paper found out that machine state does not possess such autonomy and then may not be sufficient for the study of the nature of mental state.

Moreover, the paper further identified and objected to the extreme position of computational

functionalism. The paper presented arguments showing the impossibility as well as the absurdity contained in the belief that machine state is equivalent to mental state. It therefore concluded that the view that intends to use the nature of machine state to study the nature of mental is superfluous and the view that equates machine state with mental state is trivial.

References

- Anderson D.L. Searle's Chinese Room Argument, Part Two: The Robot Reply, CCSI – Consortium on Cognitive Science Instruction, Supported by Mind Project.http://www.mind.ilstu.edu/curriculum/searle_chinese_room/searle_robot_reply.php?modGUI=2: Accessed on 15th December, 2013, at 6:09 pm
- Barrat, J. (2013). *Our Final Invention: Artificial Intelligence and the End of the Human Era*. New York: St Martin's Press
- Block N. (1993). *The Computer Model of the Mind*. In: Alvin Goldman (ed) *Readings in Philosophy and Cognitive Science*. Cambridge, Mass: MIT Press
- Boden, M. A. (1990). *Escaping from the Chinese Room*. In: Boden M. A. (ed) *The Philosophy of Artificial Intelligence*. Oxford: Oxford University Press
- Boden M. A. (1990). (ed) *The Philosophy of Artificial Intelligence*, Oxford, Oxford University Press
- Brown, R (ed) (2014). *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*. New York: Springer Science+Business Media
- Churchland, P. (1993). *The Co-evolutionary Research Ideology*. In: Alving Goldman (ed), *Readings in Philosophy and Cognitive Science*. Cambridge, Mass: MIT
- Goldman, Alvin (ed) (1993). *Readings in Philosophy and Cognitive Science*. Cambridge, Mass: MIT Press
- Huttenlocher, J. (1973). *Language and Thought*. In: George A. Miller (ed) *Communication, Language and Meaning*. New York: Basic Books
- Milikan, R. G. (2014). *An Epistemology for Phenomenology*. In: Brown, R (ed) *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*. New York: Springer Science+Business Media, pp 13-26
- Nagel, T. (1979). *Mortal Questions*. Cambridge: Cambridge University Press
- Newell A. and Simon H. (1990). *Computer Science as Empirical Enquiry: Symbols and Search*. In: Boden M. A. (ed) *The Philosophy of Artificial Intelligence*, Oxford: Oxford University Press
- Oyelakin, R. T. (2019). *Why Did the Machine Work? A Functional Theistic Interpretation from Computational Functionalism*, *Philosophy & Theology*, 31, 1 & 2, 133-149
- Putnam H. (ed) (1975). *Mind, Language and Reality: Philosophical Papers, Volume 2*; London, Cambridge University Press
- Putnam, H. (1975). *The Nature of Mental States*. In: Putnam H. (ed) *Mind, Language and Reality: Philosophical Papers, Volume 2*; London, Cambridge University Press
- Putnam, H. (1988). *Representation and Reality*. Massachusetts: The MIT Press
- Royce, J. (1968). *The Problem of Job*. In: Abernethy, G. L. and Langford, T. A. *Philosophy of Religion: A Book of Readings*. New York: Macmillan Publishing Co., Inc.
- Searle, J. (2008). *Philosophy in a New Century: Selected Essays*. New York: Cambridge University Press
- Searle, J. (1993). *The Critique of Cognitive Reason*. In: Goldman Alvin (ed) *Readings in Philosophy and Cognitive Science*. Cambridge, Mass: MIT Press.
- Turing, A.M. (1950). *Computing Machinery and Intelligence*, *Mind* 59, No 2236, 433-460
- Turing, A. M. (1990). *Computing Machinery and Intelligence*. In: Boden M. A. (ed) *The Philosophy of Artificial Intelligence*, Oxford, Oxford University Press