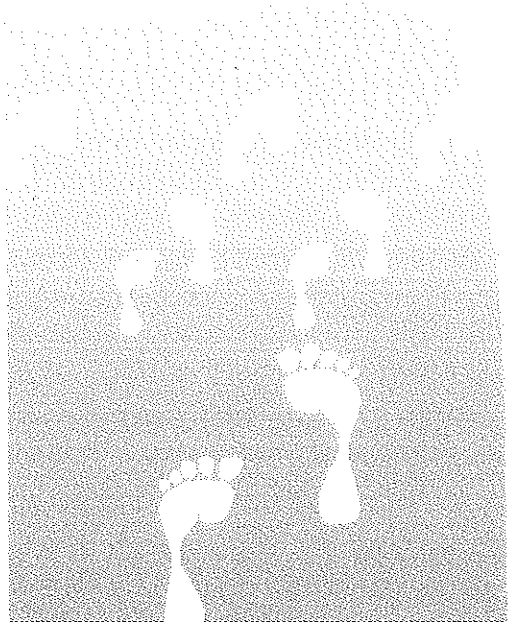
Sandhan

Journal of Centre for Studies in Civilizations

Editor
D.P. Chattopadhyaya

Volume V
Number 2
July–December
2005



Rajakishore Nath

Early and Later Putnam on Functionalism

In this paper, I shall review the reasons that led Putnam to propose functionalism and the reasons that subsequently led him to abandon it. I would like to discuss Putnam's views as belonging to early Putnam and later Putnam. First, let us focus on early Putnam. Early Putnam tries to show the possibility of robot consciousness. As a functionalist, Putnam shows that the human being is an automaton: that is, human mind is a computing machine. Before showing Putnam's views on functionalism, let me focus on the question: What is functionalism?

Functionalism is the view that mental states are defined by their causes and effects. It holds that what makes an inner state is not an intrinsic property of the state, but rather its relations to sensory stimulation (input), to other inner states, and to behaviour (output). And according to the functionalists, all these functional states can be multiple realizable in different kinds of machines. But the arrival of computer science gave impetus to functionalism. Firstly, the distinction between software and hardware suggested the distinction between function and structure. Secondly, since computers are automated, they demonstrate how inner states can be causes of output in the absence of a homunculus. Thirdly, the Turing machine provided a model for functionalism. According to Turing, each psychological state is identical to a Turing machine state. This Turing machine functionalism is largely developed by early Putnam. Thus, in short, 'functionalism' may be defined as the theory that explains mental phenomena in terms of the external input and the observable output. It explains the mind as a complicated machine.

Putnam writes, "Probabilistic automation has been generalized to allow for 'sensory inputs', and 'motor outputs', that is, the machine table specifies, for every possible combination of 'state' and a complete state of 'sensory inputs' and instruction which determines the neurological states and processes of the human brain, the machine also undergoes certain processes in determining the output".[1] The following are the steps which explain how a machine functions in general:

(i) The description of the sequence of states (procedure)
(ii) Description of rules
(iii) The explanation of the rationale of the entire procedure.[2]

The computing machine is thus a system constructed of different subsystems that function inside to process the inputs and to procedure the output once the input is simulated in the machine. It tries to match the simulating state with the states already computed and mapped in the system. This mapping order follows a certain syntax. The syntax is responsible for the correlation of total cognitive states. Thus, the entire process of simulation can be called the rational process. This rational process not only justifies the strong notion of identity, but also helps in explaining the process involved in the machine itself. However, the simulation takes place between the functions of the two functionally isomorphic systems. As Putnam defines, "Two systems are functionally isomorphic if there is a correspondence between states of one and the states of other that preserves the functional relation".[3]

There is functional isomorphism, according to Putnam, between the brain/mind and a machine. This functional isomorphism holds due to the causal capacity of the functional states of the machine. For example, when I have a pain, there is a neuro-physiological process corresponding to that mental state because of the Fining of the C-fiber. The brain/mind identity follows as there is functional identity between the two. Thus, identity between the mental states and the physical processes of the brain is established from the functional point of view. That is, in functional terms, the

brain state is isomorphic with the mental state.[4] That is, there is identity between software that constitutes the program and the hardware of the machine, which helps the software to be realized in the machine.

There can be indefinitely many different physical properties, which constitute the realizations of the same functional property. However, "it is also true that the same physical state can realize different functional properties at different times or in different circumstances or in different creatures".[5] The functional states are 'multiply realizable' in the sense that a functional state cannot be identical to any particular physical realization of it. For example, someone could write a program using two completely different types of computers, which use different sorts of hardware to run the same program. In this sense, the program is said to be 'multiply realizable' in that any number of computers may be used to realize the same program. Functionalism takes states of mind and mental properties to be functional states and properties. Mental properties are realizable by, but not identical with, material properties. For example, the same mental property, the property of being in pain, may be realized by one property in a human being and to a certain extent by another property in an invertebrate. For the functionalist, if someone has now a particular pain, then he/she can imagine that this pain is realized through a particular neural state. That neural state has an identifiable material structure, and this may be studied by a lower level hardware science like neurobiology. Therefore, for functionalism, what makes the state a realization of pain, is not its material constitution but it's occupying a particular kind of causal role within our nervous system. Multiple realizability thus implies that there is a higher level functional description of physical states in terms of their causal role, which abstracts from their lower level physical constitution. It is with such functional properties that mental properties can be identified.

The above case will be clearer, if we go through David Lewis's essay. That is, for every mental state M, there are different ways of realizing it. What matters is the functional organization of the state and not the stuff out of which it is made. This is called multiple

realizability theories. In his essay "*Mad Pain and Martian Pain*",[6] Lewis discusses two kinds of beings, which experience pain differently than normal humans. In the case of mad pain, the subject experiences pain when doing moderate exercise on an empty stomach; further, it improves his concentration for mathematical reasoning. On the other hand, Martian pain takes place in a Martian organism constructed of hydraulic hardware rather than neurons. Here the point is that pain is associated only contingently with either its causes (as in mad pain) or its physical realization (as in Martian pain). We cannot specify *a priori* its causal or physical role.

Moreover, Putnam points out that the traditional mind-body problems are wholly linguistic and logical in character. He tried to show that all issues relating to mind-body are in connection with any computing system capable of answering questions about its own structure, and has nothing to do with the unique nature of human subjective experience. We may point out that one kind of puzzle that is discussed sometimes in connection with the 'mind-body problem' is the puzzle of privacy.[7] The questions, "How do I know I have a pain?" is a logically odd question. Again, the question, "How do I know Smith has a pain?" is not logically odd. The difference can also be mirrored in impersonal questions: "How does anyone ever know he himself has a pain?" which is logically odd; and "How does any one ever know that someone else is in pain?" is not logically odd.

Putnam mirrored the above two questions in the case of machines: if T is a Turing machine, the question is, "How does T ascertain that it is in state A?" is a logically odd question; but if T is capable of investigating its neighbor machine T^1, the question; 'how does T ascertain that T^1 is in state A?' is not logically odd. Another question connected with the mind-body problem is the question whether, or not, it is ever permissible to identify mental events and physical events. Putnam claims that this question does not arise for Turing machines, but this question can be mirrored in terms of the machine analogue.

To obtain such an analogue we have to image a Turing machine, which generates theories, tests them and asserts theories, which

satisfy certain criteria. In particular, if the machine has electronic 'sense organs' which enable it to 'scan' itself, while it is in operation, it may formulate theories concerning its own structure and subject them to test. Suppose the machine is in a given state A, when, and only when flip-flop 36 is on. Then this statement, 'I am in state A, when, and only when flip-flop 36 is on,' may be one of the theoretical principles concerning its own structure accepted by the machine. Of course, here, 'I am in state A' is 'observation language' for the machine, while 'flip-flop 36 is on' 'theoretical expression' which is particularly interpreted in terms of the 'observable'. Now all the usual considerations for and against mind-body identification can be paralleled by considerations for and against saying that state A is in fact identical with flip-flop 36 being on. Now the question is: how could two things which are different be possibly the same? Putnam replies that the whole issue of mind-body identity is a linguistic one.

Putnam holds that if an identity theory were true then it would have to be true as a consequence of the meaning of psychological words. If we take the question whether light is electric magnetic radiation of such and such wavelength, it would lead to the conclusion that this too was not a question of empirical fact but called for a 'decision' on our part, a decision to treat electromagnetic radiation in a certain way. Still, light is not identical with electromagnetic radiation.[8]

In this respect, Putnam,[9] in one of his famous articles on '*Robots: Machines or artificial created life?*' says that a robot is more powerful than a human being and also more intelligent than a human being. At least in the literature of science fiction, it is possible for a robot to be 'conscious'; that means to have feelings, thoughts, attitudes and characters of traits. When this model of a robot sees something red it calls it 'red'. Putnam points out that our robots are supposed to be capable of inductive reasoning and theory construction. So a robot may discover that something it called red was not really red. Then robot would say, 'it looked red'. Thus, it will have to distinguish between the physical reality and the visual appearance as we human beings do. But the robot will never say, 'that looks as if it looked red,

but it does not really look red'. When robot says, 'it looks red' it means it is in internal state, that is 'flip-flop 72 is on'. More clearly, 'seeing red' is being in internal state 'flip-flop 72 is on'.

One can argue that the meaning of 'it looks red' is not the same as the meaning of 'flip-flop 72 is on'; the attributes or conditions or state 'being in the state of seeing to see something red' and 'having flip-flop 72 is on' are two attributes but not one. One of the important objections to this view is that a robot or a computing machine cannot follow rules. When robot says, 'I see something that looks red' may be uttered whenever flip-flop 72 is on, whether the robot 'knows' that flip-flop 72 is on or not. 'Flip-flop 72 is on' may be correctly uttered only when the robot 'knows' that flip-flop 72 is on. Now, the questions are: Can the robot know that it has the 'sensation' of red? Can the robot have the sensation of red? When we say that a human being has the 'sensation of red, in this case it has some epistemic and logical sense. But in the case of robot, it has no sense.

Following the above argument, there are two objections proposed by Kurt Baier:[10]

1. The connection between mind's visual sensation of red and my utterance 'it looks as if there is something red in front of me' is not merely a casual one. In the case of utterance, I utter the utterance because I know that I am having the sensation. But the robot utters the utterance because it is caused to utter it by its internal state {flip-flop 72 is on}. Thus there is a fundamental disanalogy between two cases.
2. Certain qualia are intrinsically painful and others are intrinsically pleasurable. However, if a robot is programmed so that it acts as if it were having a pleasant experience when, say, a certain part of its anatomy jangles, it could easily be reprogrammed so that it would act as if it were having a painful state. But the qualia in the robot are absent; they cannot be intrinsically pleasurable and painful.

Putnam, in order to establish his theory of robotic mind, says that one can know that one has a sensation without knowing that one

is in brain state; hence, the sensation cannot be identical with brain state. If it follows in the human case that the sensation of red is not identical with the brain state S then by the same argument from the same semantical premises, the robot philosopher can conclude that the 'sensation' of red is not identical with 'flip-flop 72 is on'. The robot philosopher can argue: "I am not merely caused to utter the utterance 'it looks as if there is something red in front of me', by the occurrence of the 'sensation'; part of the causation is also that I 'understand' the words that I utter; I 'know' that I am having the 'sensation'; I 'wish' to report my 'sensation' to other robots, etc".

From the above example one can observe that Putnam is supporting both Baier and the robot philosophers. He says, psychological attributes, whether in human language or in robot language, are simply not the same as physical attributes. Thus, the attributes having the 'sensation' of red and 'flip-flop 72 being on' are simply not identical in the case of robots. If materialism is taken to be the denial of the existence of 'non-physical' attributes, then materialism is false even for robots.

Again he argued that a Turing machine might very well be a biological organism. But the question whether an actual human being is a Turing machine, or whether the brain of a human being is just a Turing machine, is an empirical question. In this way, he points out that the Cartesian dualist who likes to think of the human mind as a self-contained system which is in some sort of causal interaction with the body, one can say that, from the point of view of logic, it is entirely possible that the human mind is a Turing machine. Again, one may say that, from the standpoint of pure logic, it is entirely possible that human soul is a Turing machine, or a finite automaton. Although, it is likely that human brain states form a discrete set and that human mental states are a discrete set, no matter what meaning may be given to the somewhat ambiguous notion of a mental state, it is somewhat unlikely that either the mind or the brain is a Turing machine.

The *a priori* reasoning shows that the interconnection among the various brain states and mental states of a human being are probabilistic rather than deterministic. The empirical data cannot

decide between the hypothesis that the human brain is a Turing machine and the hypothesis that it is a more complex kind of automaton with probabilistic relations.

Now the question arises: Does a computing machine have intelligence, consciousness, and so on in the way that human beings do? As we have seen, Putnam says that 'the mind is a Turing machine.' If this is so, the whole human body is a physical system obeying the laws of Newtonian physics, then the whole physical universe is at least metaphorically a machine. Thus, early Putnam's argument shows that the whole human body is at least metaphorically a machine.

Let us go back to the robot which was supposed to be a 'psychological isomorphic' to a human being. Firstly, there is no psychological isomorphic, rather there is an extension of natural psychology. Because a machine is merely a library of predetermined behaviour routines and each of the machines' action determined by the programmer or human being. This shows that it fails to be psychological isomorphic to any human being. Secondly, epistemological, metaphysical and moral arguments show that there is no isomorphic relationship between human and robot. If machines are conscious, they have feelings, thoughts, attitudes etc. But now the question is; is it really possible? If it is possible then what are the necessary and sufficient conditions? Then there will be no mind-body problem, and the problem of other minds does not arise. With regard to this. Paul Ziff[11] wishes to show that it is false that robot is conscious. He begins with the undoubtful fact that if robot is not alive it cannot be conscious. Here, Ziff has given the semantical connection between 'alive' and 'consciousness' in English. This semantic connection shows that robot is not alive. Thus from Ziff's argument it is clear that Putnam is wrong in holding isomorphic relation between human being and robot.

The theory that proposes to provide complete description of our psychological state as a Turing machine is a utopian project. He realized this later because this sort of utopianism is an illustration of what is called 'scientism'.[12] It is based on speculations regarding scientific possibilities. The problem is that it is completely unclear

just what possibility has been envisaged when one speaks about robotic consciousness. While arguing against AI, the later Putnam said that pessimism about the success of AI in simulating human intelligence amounts to pessimism about the possibility of describing the functioning of the brain. He raised the question: What connection is there between simulating intelligence and describing the brain? Even if the computer model of the brain is correct, it does not follow that AI will succeed. Simulation is just simulation, it does not mean exact replications of human mind.

Moreover, the later Putnam mentions that functionalism is incompatible with our semantic externalism because functional organism is not simply a matter of 'sensory inputs', transition from one state to another, and 'motor outputs'. As we know that semantic externalism is the content of our words and thoughts, which is partly determined by our relation with things in environment.[13]

A robot whose 'brain' was merely a library of predetermined behaviour routines, each imagined in full detail by the programmer would indeed be uninteresting. This shows that there is no isomorphic relation to the programmer. Later Putnam revising his earlier view in his book '*Representation and Reality*' rejected the computational view of mind on the ground that the literal Turing machine would not give a representation of the psychology of human beings and animals. For him, functionalism is wrong in holding the thesis that propositional attitude is just the computational state of the brain. For example, to believe that there is a cat on the mat, is not the same thing as that there is one physical state or a computational state believing that there is a cat on the mat. No one any longer believes that semantic and propositional attitudes are semantically or conceptually reducible to computational predicates. Then the question is whether these semantic and propositional attitudes, properties and relations are 'reducible' to physical cum computational properties and relations. According to later Putnam,[14] this is impossible because propositional attitudes refer to the intentional state, that is to say that it refers to various states of affairs in the world. For example, if I say that John will go to New Delhi from Hyderabad, this statement refers to many

attitudes, and it cannot be realized computationally. Thus, according to later Putnam, the functionalist is wrong in saying that semantic and propositional attitude predicates are semantically reducible to computational predicates.

There is no reason why the study of human cognition requires that we try to reduce cognition either to computations or to brain processes. We may well succeed in discovering theoretical models of the brain which vastly increase our understanding of how the brain works. But if we reduce the human mind into brain, it will no way be a help to brain science. The reductionists way of understanding human mind is a tried one, but still it has not lost its grips on our scientific culture.

Therefore, functionalism fails to account for the real nature of the mental states because of its not too clear attempt to reduce mental states to the machine-states. Putnam said, functionalism fails as a theory of mind because of its reductionist dogma: it makes mind superfluous or meaningless in the human world. Mind is made at best a mechanical system with certain determinate functions.

Moreover, the mechanistic theory of mind faces many questions like: How can we account for the qualitative content of our consciousness? It cannot tell us how the qualia are possible and also how consciousness can be real in the universe. This thesis says that the qualia of consciousness are real and that constitutes the essence of human mental life. According to R.C. Pradhan,[15] the above thesis follows from the conviction that we cannot conceive of consciousness unless we view its states as having the raw feelings. He said, there are two aspects of this thesis, the epistemological and the metaphysical. Epistemologically, the raw feelings are intimately known to the subject of consciousness; they are introspectively known to the subject in the sense that he or she can have a first person account of them more reliable than others. That is to say, the reports such as 'I am in pain' have a characteristic first-person authenticity which follows from the fact that the person concerned cannot doubt that he or she has pain while uttering the sentence 'I am in pain'. However, metaphysically speaking, the raw feelings are real in the sense that they are part of the furniture of the mental

world. Thus it is very difficult for us to deny that the mental world is real.

The mechanistic view does not have any convincing answer to the question how qualia are a necessary feature consciousness. If mind functions like a machine, it can at best exhibit only mechanical states, which look very much like the mental states but ontologically are very different.

In the end, Putnam concludes that the mental life is autonomous because of our knowledge of the inner structure of consciousness. This shows that the mental life is irreducible and different from the physical world. Then mind has its own and autonomous existence in this physical world. This is because of the separate existence of mental life but the materialistic or physicalistic or functionalistic theories fail to understand or explain the mental life of human beings.

Acknowledgement

My sincere thanks and regards to my supervisor Professor R.C. Pradhan for his constant guidance and for helping me to improve the philosophical quality of my paper. All the errors, if any, are mine.

References

1. Putnam, H., "The Nature of Mental States", in *Philosophy of Mind: Classical Problem/Contemporary Issues*, Veakkly Brain and Peter Ludlow (ed.), The MIT Press, MA, 1992, p. 54.
2. Putnam, H., "Mind and Machine", in *Minds and Machines*, Allan Ross Anderson (ed.), 1858, p. 78.
3. Putnam, "Philosophy and Our Mental Life", in *Mind, Language and Reality: Philosophical Papers 2*, Cambridge University Press, Cambridge, 1976, p. 91.
4. *Ibid*, pp. 37-38.
5. Shoemaker, S., *Identity, Cause and Mind*, Cambridge University Press, Cambridge, 1984, p. 267.

6. David, Lewis, *Philosophical Papers*, Volume 1, Oxford University Press, Oxford, 1983, p. 122.
7. Putnam, H., "Minds and Machines in Mind" in *Mind, Language and Reality*, Volume 2, 1975, p. 362.
8. Putnam, H., "Hillary Putnam" in *A Companion to the Philosophy of Mind*, Samuel Gunterplan, Blackwell Publishers, USA, 2000, p. 501.
9. Putnam, H., "Robots: Machines or Artificially Created Life?" in *Mind, Language and Reality*, Volume 2, 1975, p. 386.
10. *Ibid*, pp. 389-390.