Mario Radovan
Fakultet ekonomije i turizma
Pula
email: Mario.Radovan@efpu.hr

UDK: 007.52
Pregledni rad

Mind and Computation

The paper examines basic positions concerning the computational model of the mind, and the background assumptions on which these positions are based. It has been ascertained the question of the relation between human mind and computational machines does not concern so much the obsenter independent phenomena in the world, as it concerns our attitude toward these phenomena. Taken literally, mind is not a programmable machine, but there are pragmatical reasons why assign a computational interpretation to it. However, it has been shown that the paradigm established by such interpretation has also its limitations.

Key words: mind, computability, language of thought, connectionism, consciousness, metaphor, hardware independence.

1 Introduction

There are various positions concerning the computational model of the mind. These positions range from the complete acceptance, to the complete abnegation. For example, while Haugeland holds that "we are, at root, computers" [9, p. 2], and Fodor stresses that mental processes should be computational, Searle holds that his argument "put the final nail in the coffin of the theory that the mind is a computer program" [14, p. xi]. However, it seems that such sharp differences follow primarily from the different backgrounds (starting assumptions and goals) on the basis of which a particular position has been formed and defended. My intention in this paper is to explicate these background assumptions, and to examine the positions based on them. I argue that conscious mind is not a programable machine (computer), but that there are pragmatical reasons why assign a computational interpretation to it. However, a computational model of the mind has also its limitations.

2 The Weaknesses of the Model

The computational model of the mind says that mind is to the brain as the software is to the hardware. However, there are a few reasons why abnegate the validity of such an analogy; let us see them.

2.1 Software and Emergent Property

Mental states are emergent properties of the material structure brain, while software could hardly be conceived of as an emergent property of hardware. Software is imposed on the hardware (from the outside), so that there is no sense even to compare conscious mind to software. In other words, the computational model of the mind is not suitable.

But there are objections to such fast elimination of the model; for example, it has been argued that some kind of software could induce (make to emerge) some kind of mental states on some kind of hardware. In principle, it could. Namely, programs (when active) have direct impact on the primary properties of the hardware: therefore. they could, in principle. induce all sorts of states in/on it. But it would be a miracle if any kind of software (as actually conceived) would induce a mental state on any type of hardware (as actually conceived). Furthermore, it is actually not possible to work on the development of such kind of software because we don't know how the mental states emerge in the human brain. And if it is not possible even to work on the development of such software, it is not wise to expect that such (mental) states could simply happen/emerge.

However, our actual ignorance concerning the nature of mental states is not the main problem here; namely, if we would know everything concerning the human brain/mind, we would very probably also know the fact that it is not possible to replicate mental states by mere computer programs. Indeed, there is no more reason to expect that the emergence of mental states could be caused by computer programs than there is to expect that the growth of grass could be: both events are natural phenomena, and programs can at best simulate such phenomena, but not also replicate them.

2.2 Behavior and Mental State

The next problem concernes the distinction between the functional and the mental. It is argued that machines can simulate functional properties of the human mind, but not the mental ones: they can perform various well defined functions, but not also have/experience mental states. There is an unbridgeable gap between the functional and the mental; functional properties can be described and replicated by the automated formal systems, while mental states are intrinsically first-person and they can be only experienced. In other words, programmable machines can have intelligence (as behavioral disposition), but not feelings (as first-person property).

But there are also objections concerning this argument. For example, Dennett, who pleads for a "version of functionalism", says: "If all the control functions of a human wine taster's brain can be reproduced in

31]. Perhaps; for if you reproduce all the causes, you should reproduce also all the effects. However, functionalism is ipso facto not an approach which could lead us to the real "control functions" (whatever it means) of the brain.

According to functionalism, a mental state can be defined in terms of the (1) sensory inputs, (2) causal effects of other mental states, and (3) behavioral outputs. In this context, two different brain-state tokens are said to be tokens of the same type of mental state if they have the same causal relations to the input stimulus that the system receives, to its other inner mental states, and to its output behaviour. Any system, no matter what its physical realization, could have mental states provided that it had the right causal relations between its inputs, its inner states. and its outputs. Hence, an artificial brain (computer) which would be functionally isomorphic to the natural brain, would have also the same mental states. However, there are a few problems inherent to such approach.

First, by raising the discussion to the level of abstract functional structure, functionalism neglects the fact that mental states are primary qualities of the specific physical structure brain; and there is no basis for believing that abstract states which play the same functional role in a different medium would have the same primary qualities as the human brain. Or, as Schweizer put it: "while there is a good reason to believe that consciousness results from the physical processes that take place in the brain, there seems to be no reason to conclude that different material implementations of the same computational structures will reproduce these same internal effects" [12, p. 272]. In other words, based on the idea of functional isomorphism, functionalism could be the right way to the smartest machines. but not to the conscious ("enjoying") ones.

2.3 Reality and Description

Functionalists pass in silence over some basic problems, which makes their positions less convincing. First, isomorphism is not identity, and isomorphic entities are not supposed to have the same properties but only the same structure: hence, it is not sufficient to speak of functional isomorphism with human brain when aiming to replicate the human mental properties. Further, even the very idea of functional isomorphism is problematic, as long as we don't have a clear criterion on the basis of which we could decide when an artificial structure can be said to be functionally isomorphic with the human brain. Without such a structure can be said to be functionally isomorphic with the human brain. Without such a criterion, all depends on the way one describes the brain: describe it in poor (reduced) terms, and you will easy construct an artificial system isomorphic to such description! However, reality does not care much about our descriptions: hence, we could hardly replicate a phenomenon without knowing its real, and not merely "abstract", structure. Of course, the question of the real structure of (a part of) the reality is not an easy one; moreover, it seems to be open-ended. and will probably remain such forever. It is immanent to the science (in general) to search for the conceptual system which would "carve nature at its systematic joins" [4, p. 279], but there is not much hope that such result could ever be obtained. Consequently, every description of a phenomenon necessarily depends on the conceptual system of the beholder. But that does not change the fact that expressions such as "all the control functions of the human brain" (from the Dennett's quotation) do not say enough as long as we don't even approximately know what would count as "all control functions". In other words, if "reproduction" (in that quotation) meant identical system, then the quoted assertion is trivial; if not, it says far less than it seems to.

3 The Opennes of the Model

There are formal results in the theory of computability which open new possibilities for the computational model of mind; the most important among these results are Universal Turing Machine and Church 's Thesis.

Roughly speaking, Universal Turing Machine is a formal system which consists of a minimal number of symbols, states, and operations, in terms of which processes (algorithms) can be defined. Universal Turing Machine can in principle be implemented on digital computer. Church's Thesis says that every computable function is Turing computable. In other words, it means that any precisely defined process of symbol manipulation can be expressed by means of Universal Turing Machine, and be carried out by a digital computer. Church's Thesis has not been proved because the concept of "precisely defined process" is not formally defined, but "it has been supported by evidence, much as any empirical scientific theory might be" [1. p. 66].

3.1 Simulation and Reality

The above results open at least a new speculative space in the scope of the cognitive science. Namely brain processes are natural phenomenon, and they could be described by a scientific theory; further, every scientific theory is essentially a syntactic system: consequently, it should be possible to describe brain processes in purely syntactic (computational) fashion. And according to Church's Thesis, that means that these processes could be carried out by a programmable machine, which should then also have (or be in) the same mental states as humans.

Opponents of such interpretation of Church's Thesis will call our attention to the difference between simulation and reality: computer simulation of a process is based on some symbolic representation of reality, and it can produce only symbols, but not real entities. For example, no computer simulation of the processes in the cow's udder could produce the real milk, and the same holds for mental states: they can be simulated but not replicated by computational processes. However, with the mental processes, things could be at least partially different. Namely, some human mental processes are algorithmic, and consequently also literally reproducible by a computational machine. For example, there needn't be any difference between the human process of carrying on an arithmetical operation, and the machine implementation of the same process: both, man and machine, can follow the same algorithm. Therefore, when mental processes are concerned, there needn't be any difference between the "reality" and the "simulation".

3.2 The Language of Thought

The given example is too simple to justify the ambitions of to the above line of thought. However, it is possible to suppose (and to take as a hypothesis) that there are many algorithmic processes in human brains at the unconscious level, which (if known) could be explicitly described in some formal language, and then also replicated by computer. And if all the brain processes are algorithmic, the human brain - and ipso facto, human mind - is completely machine (re)producible.

To complete the above hypothesis we must suppose the existence of some language in which the basic (unconscious) processes take place, and in which they could be described. A candidate for such language could be Fodor's Language of Thought (LOT), which is common to all humans. Fodor takes LOT as a kind of machine language: computers use one or more input/output languages by means of which they communicate with their environment. and a machine language in which they run their computation; LOT is intended as a machine language. [7, p. 66].

With LOT, Fodor introduced an intermediate level between the physiological (hardware) and the conscious (input/output) level: cognitive processes are taken to be completely definable on this (intermediate) level; they are algorithmic, so that human thinking can be conceived of as computation over basic units (atoms) of LOT. Needless to say, if such brain/mind model would be the right one, then human cognitive abilities could be not only replicated but far exceeded (in speed and scope/range) by programmable machines. However, some things concerning the model have remained open; first, Fodor's hypothesis imply the existence of a set of context-free atoms on which all our thoughts (i.e. computations) are based, but we cannot say what these basic atoms/units could look like. Clark describes such (hypothetical) context-free atom as a "syntactic item" which "plays a fixed representational role", and as "an inner state which makes the same semantic contribution to each of the larger states in which it figures" [3, p. 31]. Such descriptions don't seem to be enough (in the operational sense); however, our actual ignorance concerning the "details" need not invalidate Fodor's hypothesis concerning the basic nature of the cognitive processes.

3.3 The Language of the Mental

LOT is concerned primarily with thought processes; Fodor says: "Nothing can be expressed in natural language that can't be expressed in the language of thought" [7. p. 84]. But mental states are not really expressible in natural language. Hence, following Fodor's line, let us make a further (speculative) move and postulate a Language of the Mental (LOM) with a set of innate mental atoms as basic data items, by means of which all mental processes can be defined in the algorithmic fashion. Now, we can paraphrase Fodor's above assertion by the following words: Nothing can be experienced by a human being that can't be computed in the Language of the Mental. The LOM hypothesis would make the human mind completely definable at the syntactic level, and then, by Church's Thesis, also artificially reproducible by a programmable machine.

Of course, I cannot say how LOM (and mental atoms) would look like; but we are in the same ignorance also with Fodor's LOT. However, we are often constrained to presuppose the existence of hidden or unobservable entities, structures, and processes, and then to judge the validity of such hypotheses by evaluating their formal consequences and empirical effects. Therefore, the fact that we cannot actually prove the existence of the language/atoms of some basic mental level need not be the reason to abandon the very idea of the computational definability of the conscious human mind. Instead, if there is any possibility that such hypothesis could open the way to the secret of the mental, we should proceed with it. Yes, it seems that the LOM hypothesis violate the general principle that content/semantics cannot be reduced on the mere form/syntax, but - confronted with the mystery of the mental - we are in fact constrained to try even some wild speculations. In fact, the real breakthroughs in science have always started exactly with such kind of speculation. (Which doesn't mean that every inconsistency opens a way to a great discovery.)

3.4 Beyond the Natural Mind

Syntactic definition of the basic mental processes would have fascinating consequences. First, it would render possible the replication of any human mental state possible by running a computer program that implements a process which is type-identical to that process (at the human brain's basic mental level) which causes the given mental state in the human being. However, it would become possible not only to replicate (imitate) human mental states, but also to produce new kinds of mental states, completely unknown to human beings. Namely, natural evolution (the selection principle) surely hasn't favoured the development of all possible (i.e., computable) kinds of mental states; on the other hand. an artificial brain/mind (computer) could explore a virtually open-ended combinatorial space, and so create completely new results (mental states). But even only thanks to the enormously greater processing speed, such brain-machines would far exceed man's mental abilities. Therefore, if we ever succeed to develop an artificial human-like conscious mental state based on pure computation, we shall soon be faced with machines (artificial brains/minds) whose intellectual, emotive, and creative abilities will far exceed those of the best among human beings. Some of the questions we should put to ourselves while moving in this direction could be: 'Who shall be then called to settle the measure of High and Low, of Good and Bad?': 'Could such machines make a man better?'; and finally, 'Will not the natural mind, besides such artificial beings, become superfluous, or servant to the proper product?'. (Un)fortunately, it seems that we have still time enough to think about such questions.

4 Reality and Explanation

Analogies and speculations aside, it has been argued that human mind and programmable machines do not, in fact, have much in common. In this section I put forward arguments for such position, but I put forward also some reasons for the popularity of the computational model of the mind/brain.

4.1 A Commonsense Argument

Man and computer are completely different on the physical (hardware) level as well as on the psychic (software) level. Concerning the physical level, brain consists of a collection of neural networks which don't have anything in common either with serial (von Neumann's) or with parallel (Darwinian) type of hardware. Computers of the serial type consist of some basic units (processor, working memory, etc.), and there are no such components in the human brain; on the other hand, parallel hardware consists of a bunch of serial computers which work in parallel mode, while brain's (sub)networks do not form anything similar to the bunch of (parallelly connected) serial computers. (More about it in the third argument.)

Concerning the psychic level, computer program (serial and parallel) defines/creates a set of context-free data structures and a process (algorithm) over (the contents of) these structures, but there is no trace of mental states in all of this. On the other hand, it is well known that the human mind has very poor abilities of memorizing and performing algorithms. For example, many of us are not able to carry out a mental multiplication of two three-digit numbers.

It seems that these differences (complete discrepancy!) should be sufficient reasons for abandoning the computational model of the mind. But there are further, "philosophical" reasons.

4.2 Physic, Syntax and Semantics

One of the most discussed arguments against the computational model of the mind is Searle's Chinese Room thought experiment [13]. With this experiment Searle intended to show that programs do not understand what they do/produce, independently of how intelligent their products (answers) may seem to an observer. Hence it should follow that mere programming cannot lead to the machine understanding or mental states. I hold that such conclusion is no less obvious without experiments than it is with them. Besides, Searle has interpreted wrongly his own experiment; namely, the experiment shows that the processor (Searle in the Room) does not understand, and not that the program (which Searle- processor execute) does not understand; but Searle will offer us a new (better) argument. There are many replies to the Chinese Room argument; however, I hold that all these replies can be qualified as "arguments from ignorance". The most frequent among them (System Reply) emphasise that there is not only a processor/program there, but a whole system: and it is possible that in some systems some degree of understanding somehow simply emerges, although we actually don't know how and where. Of course, it is possible; however, it is equally possible that all actual computers are self-denying beings, which often suffer in silence and sometimes make fun of us. But something of that kind doesn't seem plausible: hence, there is no much sense to put forward such kind of "arguments".

The Chinese Room argument argues that a program, as purely syntactic system, cannot constitute the mind. This argument "rests on the simple logical truth that syntax is not ... sufficient for semantics" [14, p. 200]. And in [14], Searle put forward a new argument against the computational model of the mind, based on the fact that "syntax is not intrinsic to the physics" [14, p. 208]. The argument runs like this: (1) "computation is defined syntactically" (in terms of symbol manipulation); (2) "syntax is not intrinsic to physics" (an assignment of syntactic properties to physical phenomena is relative to an observer): consequently, (3) computational processes are not intrinsic to the physical world, and ipso facto not to the brain. In other words, it cannot be discovered/shown (as an empirical fact) that brain is intrinsically a digital computer; computational interpretation can be only assigned to the brain, as well as to anything else. And then it means that the assertion 'The brain is a computer' is not "simply false", but is "ill defined" and without a "clear sense" [14, p. 225]. Namely, if we interpret it as an assertion about the discovery of some intrinsic property of the brain, it is trivially false, while if we interpret it as a decision to assign the computational interpretation to the brain, it is trivially acceptable.

Summing up, the results of Searle's two arguments seem to be the following: The 'syntax - semantics' gap is fatal for the attempt to define the mind in terms of software. and the 'syntax-physics' gap is fatal for the attempt to qualify the brain as computational hardware. In other words, we don't have stronger formal reasons to conceive of the human mind/brain as a kind of computer than we have it for anything else: therefore, the computational model of the mind/brain tells us, in fact, nothing essential/intrinsic about the mind/brain.

4.3 Connectionism:Effects and Explanation

The computational model of the mind is essentially of Fodors' type: it presupposes the existence of a stock of context-free syntactic atoms (as physical tokens and content bearers) upon which the cognitive (algorithmic) operations are performed. Atoms form the combinative base of all potential thoughts, and remain unchanged through all cognitive processes. Humans are supposed to inherit a fixed set of such atoms, while the learning and inference processes consist in the recombination and redeployment of the preexisting context-free representational primitives.

A completely different cognitive model has been developed by the connectionist approach, inspired by the results of neurophysiology. Human brain is a set of neural networks; such networks learn (acquire knowledge) by repeated exposure to a training environment (with some form of feed-back effects); a network starts with an innate/random distribution of its hidden unit weights, while exposure to an environment causes it to permanently adjusts its connection weights in a way which tends to produce the best output.

The network hardware differs not only from the hardware of classical (serial) computational systems, but also from that of parallel (PDP) systems; namely, although it is common to describe the human brain as "a massively parallel processor" [4, p. 156], there are, in fact, no "processors" in the neural networks, neither in the natural nor in the artificial ones. Further, there are no "programs" in connectionist systems: such systems are trained and not programmed. Finally, there are no context-independent data atoms/records in connectionist systems, because these systems are based on superpositional representations of knowledge/data. The representation of knowledge Kl and K2 is said to be superposed if Kl and K2 are represented by the same resources. In superpositional systems there are no context-independent data records which stand for ordinary semantic units (concepts and propositions) because every piece of knowledge is stored holistically in the sense that its "record" can be distributed throughout the network; namely, any weight (in the network) can take part in the encoding of any piece of knowledge contained in the network. And if a given (trained) network is later trained to learn/accept a new piece of knowledge, the existing weights (although preserving previous knowledge) will be changed: hence the context- dependence of the knowledge representation in superpositional systems.

Flanagan holds that the connectionist models "have called the distinction between software and hardware into question" [6, p. 180], while Clark says that connectionist models are characterized by "the lack of a firm data/process separation" [3, p. 14] because such systems do not involve program-driven computation over a fixed set of symbols. Indeed, there are in fact neither "symbols" nor "programs" nor "processors" in connectionist systems. But why then at all use the standard computational taxonomy here? And although virtually all authors emphasize the essential difference between standard computational models and connectionist models, they keep on using standard computational taxonomy (see, for example, [3, 4, 6]). It seems that there are two main reasons for such praxis.

The first reason could be the custom; namely, in spite of all the differences, connectionist systems still have some inputs and outputs, and some processes in between; therefore, one is inclined to use the same old computational terminology. In this context, rather than change too closely with their particular incarnations in classical systems". He holds that "it may be more productive to seek less restricted understandings of such concepts - understandings which can cut across many types of computational device (connectionist, classicist, and types as yet undreamed of)" [3, p. 122].

Therefore, connectionists want to preserve the computational terminology; however, they are also constrained to do so. Namely, with the superpositional knowledge representation, a weight (or set of weights) cannot be identified with any fixed semantical content because each weight contributes to the representation of many semantical units, and could be changed with the acquisition of new knowledge. Analogous problems characterise the attempts to explain processes which take part in the network. its abilities and the results it produces. Hence, connectionists need something by means of which they could form a kind of top-level enplanation of what is going on in the network, because without such explanation their results and working methodology would be "obscure" [3, p. 49], or at least not of the scientific kind. And the standard computational terminology can be an appropriate means for all such explanation. There are various techniques/methods by which such top-level explanatory schemes are developed (see, for example, [3]); however, such top-level descriptions are only post hoc semantical explications of what the network knows/does, and not of what is really going on inside the connectionist system. In other words, we use/need more levels of functional descriptions; but we should not confuse reality with descriptive models; it can be useful to assign the standard computational model to the natural systems of neural networks as well as to artificial ones, but there is all the difference between the "deep reality" of such systems and their top-level algorithmic descriptions.

5 Toward the Unknown

We have concluded that human mind is not intrinsically a programmable machine; however, this conclusion leaves open the real problem we are here concerned with, and that is: Can the (unknown) relation between the mental and the physical be studied/explained on the model of the (known) relation between software and hardware? And in the context of this question, the mere fact that "nothing is intrinsically a digital computer" [14, p. 212] counts rather little. Finally, we should ask why is the computational model of the mind so widely used if it is not suitable. But the answer to this question seems rather simple: we need some cognitive model, and we don't have a better one. Namely, "no one has much of a clue" [3, p. 224] about the nature and the ways of emerging of conscious mental states: consequently, we are constrained to speak "in figures", and the most appealing set of figures seems to be the one offered by the computational technology.

5.1 Metaphor as a First Move

One of the basic "categories" in the treatises concerning conscious mind is mystery. For example, Clark says it is a "mystery how conscious content is possible at all" [3, p. 224], and Dennett qualified human consciousness as "just about the last surviving mystery"; he defines 'mystery' as "a phenomenon that people don't know how to think about" [5, p. 21]. And hence, figurative speech enters here the stage as the tool of creative imagination. Namely, if theory can be described as "the conceptual vehicle with which we ... come to grips with the world" [4, p. 117], we can say that metaphor is a vehicle with which we come to grips with the inexpressible: it is the first move toward a scientific theory.

Metaphor is a mapping between the two domains, and as such it forms a cognitive model which give us an opportunity to speak of one (unknown) domain in terms of another (known) domain. Black [11] speaks of the isomorphism between two domains of the metaphor, while Lakoff stresses that the mapping defined by metaphor must "preserve the cognitive topology of the source domain, in a way consistent with the inherent structure of the target domain" [10, p. 215]. Let us note that isomorphism by itself does not guarantee the preservation of the cognitive topology because an isomorphic mapping could be defined in a way which is not in accordance with the cognitive topology. Further, most metaphorical mappings are not isomorphic, but partial and/or incomplete in the sense that not every entity from the source domain has its counterpart in the target domain, and not every entity from the target domain has a counterpart in the source domain. In this context, the degree of the preservation of cognitive topology and of the isomorphism of structures can be taken as the criteria of the strength/validity of the metaphor: the strength of metaphor rests on its "systematic structural match between the two domains" [8, p. 453]. According to Boyd, the computational metaphor of the mind has an "indispensable role" in the formulation of theoretical positions in cognitive science [2, p. 487]. The impact of figurative speech on cognitive science seems to be immense; indeed, greater part of the ideas are expressed in figurative fashion, and disputations are often (only) wars with metaphors. As an indicative example, let us mention the concluding paragraph of Dennett's book; he has admitted that his explanation of consciousness was "far from complete"; namely, he has not proposed a new scientific theory, but only a new metaphor. "All I have done". he says, "is to replace one family of metaphors and images with another" [5, p. 455]. Therefore, faced with the mystery of conscious mental states we are, in fact, still on the first move. But there have been made attempts to go further, fast and far.

5.2 Limitations of the Metaphor

The (alleged) independence of the software from the hardware brought to the idea that the mind should be independent of the brain. And thence, it seems to follow that: (1) mind can be studied independently of the brain, and (2) mind can be realized by means different than the human brain. I hold that both hypotheses are worthy of research: however, the independence of software from hardware should be well understood before putting too great expectations on it; let us see an example. Dennett says: "if what you are is the program that runs on your brain's computer, then you could in principle survive the death of your body as intact as a program can survive the destruction of the computer on which it was created and first run" [5, p. 431]. As it is often the case with Dennett's arguments, I must say "perhaps"; namely, to see the real strength/weakness of his argument we should clarify (1) under which conditions a program can "survive the destruction of the computer", and (2) how we could afford the same conditions for the human mind (understood as a program).

In its source form, a program does not depend on hardware, but it depends on the compiler on top of which it was written: the destruction of the compiler would make a source program "dead". Namely, a source program for which there is no compiler is but a heap of signs without meaning, because it is the compiler that defines ("gives life" to) the syntax and semantics of a programming language, and with it, to all programs written in that language. On the other hand, after being compiled (linked and loaded), the program no more needs (depends on) the compiler, but it is now dependent on the hardware on/in which it has been loaded. Moreover, when a program is loaded and linked, it could be conceived of as "a part" of hardware; namely, what in the source program were words and sentences (instructions) now are energetic (tensional) states of some points (bits) of the hardware. But couldn't a program, even in such form, be copied on a new hardware (and so outlive the old one)? Perhaps; but without additional adjustments, only on an "nearly identical" one. Therefore, for the survival of the human mind (after the death of his brain-hardware) we should have a "nearly identical" new brain, and "a version of functionalism" (Dennett's position) will not afford us anything of that kind.

To sum up, it is of little avail to try to map the hardware independence from the domain of computer system on the domain of human mind/brain system as long as we don't know (notably) more about the lower levels of the latter. Hardware independence has been developed with an essentially bottom-up approach: one must know the below level to develop an interface which makes the higher level entities independent from those of the lower level. On the other hand, to deal with the mind on the abstract/functional level, while leaving aside the physical idiosyncrasies of the brain, means to follow the top-down approach. And it is hard to expect that such an approach could lead to some spectacular results concerning the independence before we touch the bottom/brain. In other words, the computational metaphor/model of the mind is of a limited power; it could be useful inside some limits, but when we try to step by it further than it can lead us, we bind our efforts to failure.

6 Conclusion

In connection with the background assumptions, we can argue that mind is more than any machine could be; namely, machines are (by definition) functional/syntactic systems, while conscious mental states cannot be described and even less replicated by merely syntactically defined processes. However, there are such theoretical results which open the ways to speculation on the possibility of the development of such computational machines which would far exceed the humble human mental abilities.

Human mind is a product of biological/physical processes which take part in the brain, and which are no more computational than those in the liver are. However, there are pragmatical reasons why assign the computational interpretation to mind/brain processes. Namely, conscious mind is still a mystery, and literal speech cannot express essential features of mysterious things; hence, we are constrained to use metaphorical speech as the only possible tool of thought.

Finally, theories/paradigms are (in general) constructions rather than discoveries; the phenomena can be described in many ways, The Truth (if it exists) is unattainable, so that the pragmatic value (supposing we can recognize it) remains our basic guide in the scientific enterprise. It is a common practice to try to apply a known models to a new (unknown) domains; such is also the attempt to explain the human mind/brain on the basis of the computational model. I hold that such approach can give good results of the functional type, but concerning the phenomena of consciousness we are still in the scope of speculations. And it seems that there can be no real progress on the way to artificial mental states as long as we don't know more about the ways the mental states emerge in the human brain.

Literatura

[1] Boolos, S.G., C.R. Jeffrey: Computability and Logic, Cambridge University Press, 1980.

[2] Boyd R.: 'Metaphor and theory change: What is "metaphor" a metaphor for?', in A. Ortony (ed).

[3] Clark, A.: Associative Engines: Connectionism, Concepts, and Representational Change, The MIT Press, 1993.

[4] Churchland, M. P.: A Neurocomputational Perspective: The Nature of Mind and the Structure of Science, The MIT Press, 1992.

[5] Dennett, D. C.: Consciousness Explained, Penguin Books, 1993.

[6] Flanagan, O.: The Science of the Mind, The MIT Press, 1992.

[7] Fodor, J. A.: The Language of Thought, Harvard University Press, 1979.

[8] Gentner, D., M. Jeziroski: 'The shift from metaphor to analogy in Western science'. A. Ortony (ed).

[9] Haugeland, J.: Artificial Intelligence: The Very Idea, The MIT Press, 1986.

[10] Lakoff, G.: 'The contemporary theory of metaphor', in Ortony, A. (ed).

[11] Ortony, A. (ed): Metaphor and Thought, Cambridge University Press, 1993.

[12] Schweizer, P.: 'Intentionality, Qualia, and Mind/Brain Identity', in Mind and Machines, Vol. 4, No. 3, pp. 259-282., 1994.

[13] Searle, J.: 'Minds, Brains, and Programs', in D. Hofstadter and D. C. Dennett (eds): The Mind's I, Penguin Books, 1981.

[14] Searle, J.: The Rediscovery of the Mind, The MIT Press, 1992.

Primljeno: 1995-06-01

Radovan M. Um i računanje

SAŽETAK

Članak analizira temeljne stavove o kompjutacijskom (računalskom) modelu ljudskog uma, i polazne pretpostavke na kojima su ti stavovi zasnovani. Ustvrđeno je da se pitanje odnosa ljudskog uma i kompjutacijskih strojeva ne odnosi toliko na same od promatrača nezavisne pojave (u svijetu) koliko na naš stav prema tim pojavama. Doslovno govoreći, um nije kompjutacijski stroj, ali postoje pragmatički razlozi da mu se dodijeli kompjutacijska interpretacija. Međutim, pokazano je da paradigma koju takva interpretacija uspostavlja ima i svoja ograničenja.