ROCZNIKI FILOZOFICZNE
Tom LXVI, numer 3 – 2018
DOI: http://dx.doi.org/10.18290/rf.2018.66.3-3

MARCIN MIŁKOWSKI

OBJECTIONS TO COMPUTATIONALISM: A SURVEY*

1. INTRODUCTION

Computationalism has been a fruitful and broad research tradition in cognitive research. Its main commitment is to claim that the brain is a kind of information-processing mechanism, and that information-processing is necessary for cognition. Notably, the radical view that computation is both sufficient and necessary for cognition is rarely accepted[1] because it implies that any computer is already a cognitive system, and this view is counterintuitive, particularly since such a machine may be engaged only in utterly trivial operations. In this paper, however, I do not wish to enter the debate on the issue of whether there is a single correct view on the nature of cognition for all cognitive science disciplines (but see: MIŁKOWSKI 2013; AKAGI 2017).

The positive view will not be developed here. In particular, the full account of physical computation will be set aside because it has already been elucidated in book-length accounts (FRESCO 2014; MIŁKOWSKI 2013; PICCINI 2015). For the purposes of this paper, it suffices to say that the current consensus among realists about physical computation is, roughly, that it occurs always and only in physical mechanisms whose function is to compute, or to process vehicles of information (understood as degrees of freedom of a physical medium); in the case of cognitive systems, these vehicles may also

Dr hab. MARCIN MIŁKOWSKI — Zakład Logiki i Kognitywistyki w Instytucie Filozofii i Socjologii Polskiej Akademii Nauk; adres do korespondencji: ul. Nowy Świat 72, 00-330 Warszawa; e-mail: mmilkows@ifispan.waw.pl

* The work on this paper was funded by a National Science Centre (Poland) research grant under the decision DEC-2014/14/E/HS1/00803.

[1] Even Alan Newell required that the computation has to be of sufficient complexity. (See NEWELL 1980, 135–83).

bear semantic information. Instead of elaborating on this account further, a review of objections is offered below, as no other comprehensive survey is available. However, the review presented below may not encompass all kinds of objections to computationalism. Hopefully, the most popular kinds of objections are included.

The survey suggests that the majority of objections fail because they make computationalism a straw man. A number of them are, at best, red herrings. Some of them, however, have shown that stronger versions of computationalism are untenable, as well. Historically, they have helped to shape the theory and methodology of computational modeling. In particular, a number of objections show that cognitive systems are not *only* computers, or that computation is not the sole condition of cognition. This objection is already incorporated in the formulation of the main commitment above; no objection, however, establishes that there might be cognition without computation. The objections are ordered in the following way: (1) as related to the (supposed) nature of computers; (2) as related to semantics; (3) against computational functionalism; (4) as related to physical computation. In the conclusion, I briefly summarize the state of the debate.

2. OBJECTIONS RELATED TO THE (SUPPOSED) NATURE OF COMPUTERS

COMPUTER METAPHOR IS JUST A METAPHOR

A computational description of a cognitive system is sometimes described as *a computer metaphor*. The use of the term suggests that the proposed description is rough and highly idealized; thus, it cannot be treated *literally*. For example, Karl Friston writes about the mathematical formulation of the free-energy principle in the following way: "Crucially, under some simplifying assumptions, these variational schemes can be implemented in a biologically plausible way, making them an important metaphor for neuronal processing in the brain." (FRISTON 2012, 2101; see also, e.g., EKMAN 2003). As such, this is not an objection to computationalism. Obviously, all kinds of scientific models use idealizations. However, by using the term, others suggest that no computational model may be treated seriously; all are mere metaphors (DAUGMAN 1990, 9–18).

A defender of computationalism might concede this and weaken his position to say that idealizations inherent in computational modeling are misleading. But the position is also tenable in the stronger version (NEWELL & SIMON 1972; PYLYSHYN 1984, xiv–xvi). This is because computer metaphors cannot really be tested and rejected, whereas computational models can. The objection, in other words, fails. It is also confused: all scientific models are idealized, which does not make them metaphorical.

SOFTWARE IS NOT IN THE HEAD

This objection is that there is no simple way to understand the notions of *software* and *hardware* as applied to biological brains. But the software/hardware distinction, popular in the slogan "the mind is to the brain like software is to hardware" (BLOCK 1995; PICCININI 2010), need not be applicable to brains at all for computationalism to be true. This objection is probably based on a conflation of stored-program computers with all possible kinds of computers. There are non-program-controllable computers: they do not load programs from external memory to internal memory in order to execute them. A mundane example of such a computer is a logical AND gate. In other words, while it may be interesting to inquire whether there is software in the brain, even if there were none, computationalism could still be true. This objection, therefore, fails.

COMPUTERS ARE JUST FOR NUMBER-CRUNCHING

Another intuitive objection, already stated (and defeated) in the 1950s, is that brains are not engaged in number-crunching, while computers compute over numbers. But if this is all computers do, then they don't control missiles or send documents to printers. After all, printing is not *just* number crunching. The objection rests, therefore, on a mistaken assumption that computers can only compute numerical functions. Computer functions can be defined not only by integer numbers but also through arbitrary symbols (NEWELL 1980), and, as physical mechanisms, computers can also control other physical processes.

COMPUTERS ARE ABSTRACT ENTITIES

Some claim that, because symbols in computers are, in some sense, abstract and formal, computers — or at least computer programs — are abstract as well (LAKOFF 1987; BARRETT 2016; BARRETT, POLLET, & STULP 2014). In other words, the opponents of computationalism claim that it implies ontological dualism (SEARLE 1990). However, computers are physical mechanisms, and they can be broken, set on fire etc. These things may be difficult to accomplish with a collection of abstract entities. Computers are not *just* symbol-manipulators. They do things, and some of the things they do are not computational. In this minimal sense, computers are physically embodied, not unlike mammal brains. This objection is, therefore, false. It is, however, a completely different matter whether the symbols in computers have any meaning.

PEOPLE ARE ORGANISMS, COMPUTERS ARE NOT

Louise Barrett (2016), among others, presses the point that people are organisms. It is trivially true, but irrelevant without further specification: physical computers are physical, and they may be built in various ways. A computer may be built of DNA strands (ZAUNER & CONRAD 1996), so why claim that it is metaphysically impossible to have a biological computer?

Another way to spell out this objection is to say that biological brains, thanks to their unique biological structure, give rise to features necessarily absent in computers.[2] In other words, features of life are somehow necessary for cognition, maybe because wherever there is cognition, there is life (THOMPSON 2007). It could be argued, for example, that only processes occurring in living organisms can lead to the emergence of normativity, which, in turn, is required in order to evaluate the accuracy of mental representations. Some authors claim, for instance, that biological functions are required to account for such normativity, and that artifacts, such as computers, necessarily lack it, because these artifacts are not far from their thermodynamic equilibrium; hence, they need not maintain their own far-from-equilibrium state (BICKHARD 2009; DEACON 2012). This is an interesting claim, but it does not constitute a strong objection to the version of computationalism assumed widely in the cognitive science. Only its very radical proponents would say that computation is both necessary and sufficient for cognition. If computation is just necessary, then biological functionality may also be required, and this objection is not at all plausible. To

[2] This much more cogent version of this objection was not included in the previous, shorter version of the paper.

make it plausible, one would have to show that life excludes computation, which is absurd. People really can compute and perform mathematical proofs without dying.

Moreover, it is far from clear that one cannot build a *biological* computer, in the sense that it would be far from its thermodynamic equilibrium. Preserving the far-from-equilibrium state does not require cognition, either. A wax candle, while burning, does maintain such a state, and it remains dubious that a more complex organization, which could organize the system to maintain its far-from-equilibrium state (making it *ultrastable*, as cybernetics calls it (FROESE & STEWART 2010)), cannot occur in systems engaged in computational processing.

The objection can be, however, understood in a slightly different vein, as implying that the very structure of biological brains gives them particularly important powers that are lacking in computational artifacts. For example, Daniel Dennett has recently interpreted Terrence Deacon's objection against computationalism as implying that one should look at biological features such as the lack of clear-cut hierarchies and bureaucratic divisions of labor in the brain. Instead of positing simplistic and biologically unrealistic "politburo" hierarchies, one should include the role of the competition and noise in the brain (DENNETT 2017). Note, however, that Dennett does not consider this as an objection to computationalism, unlike Deacon. Nothing logically precludes building such computers if they turn out to be, for example, energetically more efficient (even if the price for the efficiency is a lack of thermodynamic stability).

Yet another way of putting this objection is to say that computationalism has to ignore the embodiment of cognition. Although it is not entirely clear that embodied cognition is as fecund an approach to cognition as many claim (GOLDINGER et al. 2016), the mechanistic approach to computation implies that cognitive mechanisms have to be understood as embedded in a larger environment (MIŁKOWSKI 2017) and may be realized using bodily structures, which may go beyond the neural system (MIŁKOWSKI 2012; NOWAKOWSKI 2017). Again, computationalism does not say that the body is not required for cognition. It only says that computation is necessary, and it is plausible to think that biological information-processing mechanisms mesh with bodily processes and structures. In short, this objection is just a red herring.

Computer models ignore time

Proponents of dynamical accounts of cognition stress that Turing machines do not operate in real time. This means that this classical model of computation does not appeal to real time; instead, it operates with the abstract notion of a computation step. There is no continuous time flow, just discrete clock ticks in a Turing machine (Bickhard & Terveen 1995; Wheeler 2005). This is true. But is this an objection to computationalism?

First, some models of computation appeal to real time (Nagy & Akl 2011), so one could use such a formalism. Second, the objection seems to confuse the formal model of computation with its physical realization. Physical computers operate in real time, and not all models of computation are made equal; some will be relevant to the explanation of cognition, and some may only be useful for computability theory. A mechanistically-adequate model of computation that describes all relevant causal processes in the mechanism is required for explanatory purposes (Miłkowski 2014). This objection is, again, a red herring; the debate about computationalism does not depend on modeling time in computers.

Brains are not digital computers

Universal Turing machines are crucial to computability theory, and these are digital machines. One could, however, maintain that brains are not digital computers (Edelman 1992; Lupyan 2013). But computationalism can appeal to models of analog computation (for example, Siegelmann & Sontag 1994), or even more complex kinds of computation (Piccinini & Bahar 2013), if required. These models are still understood as computational in computability theory, and some theorists indeed claim that the brain is an analog computer, which is supposed to allow them to compute Turing-incomputable functions. Thus, one cannot dismiss all kinds of computationalism by saying that the brain is not a digital computer. There are analog computers, and an early model of a neural network, Perceptron, was analog (Rosenblatt 1958). The contention that computers have to be digital is just dogmatic or purely verbal. This objection is, therefore, mistaken (and sometimes also a red herring); computationalism does not assume digitality.

3. OBJECTIONS RELATED TO SEMANTICS

SYMBOLS IN COMPUTERS MEAN NOTHING

One of the most powerful objections formulated against the possibility of Artificial Intelligence is associated with John Searle's (1980) Chinese Room thought experiment. Searle claimed to show that running a computer program was not sufficient for semantic properties to arise, and this was in clear contradiction to what had been earlier advanced by proponents of artificial intelligence (AI), who assumed that it was sufficient to simulate the syntactic structure of representations for the semantic properties to appear. As John Haugeland quipped: "if you take care of syntax, the semantics will take care of itself" (HAUGELAND 1985, 106). But Searle replied that one can easily imagine a person with a special set of instructions in English who could manipulate Chinese symbols and answer questions in Chinese without understanding it at all. Hence, understanding is not reducible to syntactic manipulation. While the discussion around this thought experiment is hardly conclusive (PRESTON & BISHOP 2002), the problem was soon reformulated by Stevan Harnad (1990) as "the symbol grounding problem" (SGP): How can symbols in computational machines mean anything?

If the SGP makes sense, then one cannot simply assume that symbols in computers mean something simply by being part of computers, or at least they cannot mean anything *outside* the computer so easily (even if they arguably contain instructional information, at least in stored-program computers — FRESCO & WOLF 2013). Many (though not all — SHAGRIR 2010; 2011; DRESNER 2010) defenders of computationalism admit that representational properties do not necessarily exist in physical computational mechanisms (EGAN 1995; FRESCO 2010; PICCININI 2008; MIŁKOWSKI 2013). So, even if Searle is right and there is no "intrinsic" semantics in computers, the brain might still be a computer, as computers need no semantics to be computers. Perhaps something additional to computation is required for semantics. But admitting this does not undermine computationalism.

There *is* an important connection between the computational theory of mind and the representational account of cognition: they are more attractive when both are embraced. Cognitive science frequently explains cognitive phenomena by referring to semantic properties of mechanisms capable of information-processing (SHAGRIR 2010). Brains are assumed to model reality, and these models can be utilized in computations. While this seems plausible to many, one can remain computationalist without assuming representationalism (the claim that cognition requires cognitive representation). At the same time, a plausible account of cognitive representation cannot be couched merely in computational terms, as long as one assumes that the symbol grounding problem makes sense at least for some computers. To make the account plausible, most theorists of representation appeal to notions of teleological function and semantic information (MILLIKAN 1984; DRETSKE 1986; BICKHARD 2009; CUMMINS & ROTH 2012), which are not technical terms of computability theory, nor can be reduced to such. However, processing of semantic information is still processing of information. Hence, computation is necessary for manipulation of cognitive representation. Instead of undermining computationalism, Searle's objection seems to vindicate it.

Computationalism has been strongly connected to cognitive representations by the fact that it offered a solution to the problem of what makes meaning causally relevant. Many theorists claim that because the syntax in computer programs is causally relevant (or efficacious), so is the meaning. While the wholesale reduction of meaning to syntax is implausible, the computational theory of mind makes it clear that the answer to the question includes the causal role of computational vehicles' syntax. Still, the fact that it does not offer a naturalistic account of meaning is not an objection to computationalism itself. That would indeed be too much. At the same time, at least some naturalistic accounts, such as Millikan's and Dretske's, can be used to solve the SGP.

This objection is, therefore, plausible in some sense, but does not go against computationalism. It is a red herring, even if the debate over representation that it started is serious. Indeed, representational properties of computation should not be taken for granted. The semantics will not take care of itself when there is only syntax in place.

COGNITION IS NOT ALL SYMBOLIC

This objection might as well be true. Note that computationalism need not say anything about symbols. Moreover, the notion of the symbol in cognitive science has been notoriously ambiguous, since it has been used to mean various things, including tokens over a finite alphabet of the Turing machine, LISP data structures, pointers to data structures, variables standing for targets of natural language nouns, mental representations whose meaning is not related to the form of their vehicles, and more (MIŁKOWSKI 2013, chap. 4). But there is no special reason for a computationalist to defend the notion that cognition has to be symbolic in either of the above senses. It suffices that there are vehicles of information (understood in terms of degrees of freedom). This objection is a red herring concerning computationalism.

COMPUTERS CAN ONLY REPRESENT WITH ALL DETAIL

The debate over meaning in computers and animals abounds with other red herrings, however. One recent example is Robert Epstein (2016). His most striking mistake is the assumption that computers always represent everything with arbitrary accuracy. Epstein cites the example of the way in which people remember a dollar bill. He assumes that computers would represent it in a photographic manner with all available detail, in contrast to how people do it (but he seems to ignore the existence of people with excellent and apparently unbounded memory—LURIJA 2002). This is obviously false: representation is useful mostly when it does not convey information about all properties of the represented target. If Epstein is correct, then there are no JPEG files in computers, as they are not accurate, because they are based on lossy compression. Moreover, no assumption of the computational theory of mind says that memory should be understood in terms of the von Neumann architecture, and it is controversial to suggest that it should, as some theorists claim (GALLISTEL & KING 2010). This objection is, therefore, a red herring.

PEOPLE DON'T PROCESS INFORMATION

Ecological psychologists stress that people do not process information, but merely pick it up from the environment (GIBSON 1986; cf. CHEMERO 2003). To understand this, one should make the meaning of *information processing* in the computational theory of mind more explicit. What kind of information is processed? The information in question need not be semantic, as not all symbols in computers are *about* something. The minimal notion that could suffice for our purposes is one of information as a degree of freedom of a physical medium, called *structural information* by D. MacKay (1969). The number of degrees of freedom, or yes-no questions required to exactly describe its current state, is the amount of structural information. As long as there are vehicles with multiple degrees of freedom, and as long as they are part of causal processes that cause some other vehicles—just as some models of computation describe these processes (MIŁKOWSKI 2014)—there is information processing. This is a very broad notion, as all physical causation implies information transfer and processing in this sense (COLLIER 1999).

The Gibsonian notion of information pickup requires vehicles of structural information as well. There needs to be some information out there to be picked up, and organisms have to be structured so as to be able to change their state in response to information. Gibsonians could, however, claim that the information is not processed. It is unclear what is meant by this: for example, Chemero seems to imply that processing is, in effect, adding more and more layers of information, like in Marr's account of vision (CHEMERO 2003, 584; cf. MARR 1982). But information processing need not require multiple stages of adding more information. In summary: the Gibsonian account does not invalidate computationalism at all. So, the objection is a red herring.

4. OBJECTIONS AGAINST COMPUTATIONAL FUNCTIONALISM

CONSCIOUSNESS IS NOT COMPUTATIONAL

Some find (some kinds of) consciousness to be utterly incompatible with computationalism, or at least, unexplainable in purely computational terms (CHALMERS 1996). The argument is probably due to Leibniz's thought experiment in *Monadology* (LEIBNIZ 1991). Imagine a brain as large as a mill. Now, enter it. Nowhere in the interplay of gears could you find perceptions, or qualitative consciousness. Hence, you cannot explain perception mechanically. Of course, this Leibnizian argument appeals only to some physical features of mechanisms, but some still seem to think that causation has nothing to do with qualitative consciousness.

The argument, if cogent, is applicable more broadly, not just to computationalism; it is supposed to defeat reductive physicalism or materialism. For this reason, the objection might be dismissed as attacking any scientific project which explains consciousness reductively (or functionally). Thus, if the objection is correct, there is nothing special about computationalism that is under attack in this case. For this reason, any successful defense of a reductive or functional account of consciousness should focus on the issue of consciousness as reductively explainable, rather than on the issue of computation itself, and this goes beyond the scope of this paper.

Interestingly, however, virtually all current theories of consciousness are computational, even those that appeal to quantum processes (HAMEROFF

2007). For example, Bernard Baars offers a computational account in terms of the global workspace theory (BAARS 1988; cf. also DENNETT 2005), David Rosenthal gives an account in terms of higher-level states (ROSENTHAL 2005; cf. CLEEREMANS 2005), and Giulio Tononi (2004) explains it in terms of minimal information integration. Is there any detailed theory of consciousness that is not already computational?

John Searle (1992), however, suggests that only a non-computational theory of consciousness can succeed. His claim is that consciousness is utterly biological (so this resembles the objection discussed above in section 2). How does this contradict computationalism, given that there might be biological computers? Moreover, Searle fails to identify the specific biological powers of brains that make them conscious. He just passes the buck to neuroscience, which often offers computational accounts. This objection is, therefore, a red herring: if computationalism fails, the whole project of a scientific theory of consciousness fails with it (elsewhere, I have shown that some apparently dualist theories are not only functionalist but also fail to be dualist — MIŁKOWSKI 2011).

GENUINE ARTIFICIAL INTELLIGENCE IS IMPOSSIBLE

There are a number of arguments of a form:

People ψ.
Computers will never ψ.
____
So, artificial intelligence is impossible (or computationalism is false).

This argument is enthymematic. The conclusion follows with a third assumption: if artificial intelligence is possible, then computers will ψ. The plausibility of the argument varies from case to case, depending on what you substitute for ψ. For years, it was argued that winning in chess is ψ (DREYFUS 1979), but it turned out to be false. So, unless there is a formal proof, it is difficult to treat premise 2 seriously.

What could be plausibly substituted for ψ? There are many properties of biological organisms that seem irrelevant to this argument, including, exactly the same energy consumption, having proper names, spatiotemporal location, etc. The plausible candidate for substitution is some capacity for information-processing. If there is such a human capacity that computers do not possess, then the argument is indeed cogent.

ONLY PEOPLE CAN SEE THE TRUTH

This classical anti-computational argument points to the human ability to recognize the truth of logical statements that cannot be proven by a computer (LUCAS 1961; PENROSE 1989). It is based on the alleged ability of human beings to understand that some statements are true, which is purportedly impossible for machines. This argument is based on the Gödel proof of incompleteness of the first-order predicate calculus with complexity sufficient to express basic arithmetic. The problem is that this human understanding has to be both non-contradictory and certain to qualify as mathematically true. But Gödel has shown that, in general, it cannot be decided whether a given system is contradictory or not. So, either it is mathematically certain that human understanding of mathematics is non-contradictory, which makes the argument inconsistent, since it is undecidable and, as such, cannot ever be mathematically certain; or the argument simply assumes non-contradiction of human understanding, which makes the argument unsound because people make contradictions unknowingly (KRAJEWSKI 2007; PUTNAM 1960).

COMMON SENSE CANNOT BE FORMALIZED

Another similar argument points to common sense, which is a particularly difficult capacity. The trouble with implementing common sense on machines is sometimes called (somewhat misleadingly — SHANAHAN 1997) *the frame problem* (DREYFUS 1972; 1979; WHEELER 2005). Inferential capacities of standard AI programs do not seem to follow the practices known to humans, and that was supposed to hinder progress in such fields as high-quality machine translation (BAR-HILLEL 1964), speech recognition (held to be immoral to fund — WEIZENBAUM 1976), and so on. Even if IBM Watson wins in *Jeopardy*!, one may still think it is not enough.

However, it is not sufficient to say that common sense has still not been implemented on a computer in order to reject computationalism as a whole. Granted, human-level performance in perception, motor control, or speech recognition is still a pipe dream. But this does not mean that these domains are, in some sense, non-computational. Although, introspectively, these capacities may seem effortless and simple, they may be based on fairly complex information-processing.

Even if the proponent of computationalism need not require that genuine AI be based on a computer simulation of human cognitive processes, he or she still must show that human common sense can be explained computationally. Whether it can or not is still a matter of debate; however, it is important to note that the CTM would still remain true even if it was impossible to implement common sense on any artificial computer. In other words, the issue of the possibility of genuine AI is separate from the issue of whether computationalism is true. The objection is, therefore, again a red herring.[3]

5. OBJECTIONS RELATED TO PHYSICAL COMPUTATION

COMPUTERS ARE EVERYWHERE

At least some plausible theories of physical implementation of computation lead to the conclusion that all physical entities are computational (this stance is called *pancomputationalism* — cf. MÜLLER 2014). If this is the case, then the computational theory of mind is indeed trivial, as not only are brains computational, but also cows, black holes, cheese sandwiches, etc. However, a pancomputationalist may reply by saying that there are several kinds (and levels) of computation, and that brains do not execute all kinds of computation at the same time (MIŁKOWSKI 2007). So not just any computation, but some *non-trivial kind* of computation, is specific to brains. Only the kind of pancomputationalism which assumes that everything computes all kinds of functions at the same time is catastrophic, as it makes physical computation indeed trivial (PUTNAM 1991; SEARLE 1992). But there is no reason to believe that such pancomputationalism is true.

THERE ARE NO COMPUTERS

Another more radical move is to say that computers do not really exist; they are just in the eye of beholder. According to John Searle, the beholder decides whether a given physical system is computational, and, therefore, may make this decision for virtually everything. Nothing, intrinsically, is a computer. But the body of work on physical computation in the last decade or so has been focused on showing why Putnam and Searle were wrong in some

[3] In the shorter version of this paper, presented during Cognitive Science Annual Conference, this objection was considered as the most important and successful. However, for the reasons given above, I now consider this to be yet another red herring.

sense (CHALMERS 2011; PICCININI 2015; MIŁKOWSKI 2013; SHAGRIR 2011; SCHEUTZ 1996; CHRISLEY 1994; COPELAND 19961). The contemporary consensus is that computational models can be used to adequately describe causal connections in physical systems, and that these models can also be *falsely* ascribed. In other words, computational models are not different in kind from any mathematical model used in science. If they are mere subjective metaphors and do not describe reality, then mathematical models in physics are subjective as well (MCDERMOTT 2001).

Intuitively, arguments presented by Searle and Putnam are wrong for a very simple reason: why buy a new computer instead of ascribing new software to the old one? We know that such ascriptions would be extremely cumbersome unless they are systematically justified. Therefore, there must be a flaw in such arguments, as the point of ascriptions in this case is that they are arbitrary. In contrast, while in computability theory, it is common to describe a physical computer, even if it implements a finite state machine with a sufficiently large memory, as a universal Turing machine, such idealization is systematic and non-arbitrary. Even if the technicalities involved are indeed interesting, the arguments presented by Searle and Putnam fail to establish a conclusion. At the same time, they have sparked a lively discussion on the issue of physical computation that makes computation not just a matter of arbitrary descriptions.

6. CONCLUSION

In this paper, I have listed and summarized a number of arguments against computationalism. Objections in the second section mostly assume that all computers have the properties of common, industrially produced computers, which is simply incorrect and which underestimates the sheer variety of computational machines. The objections in the third section are related to an important issue in cognitive science, which is mental representation. But computationalism is logically compatible with antirepresentationalism, so these objections miss the point, even if the debate about representation remains important in the research on cognition. The objections listed in the fourth section associate computationalism with functionalism, which assumes that AI models will produce all cognitive capacities. However, because computationalism need not assume that computation is sufficient for cognition, the issue of genuine AI or computational simulation of consciousness is logically independent from the issue of whether computationalism is true. Lastly, the fifth section brings general arguments meant to show that either computers do not exist, or that everything is a computer. Both claims have recently been shown to be very contentious or to demonstrate that a stronger version of computationalism was assumed in the debate.

Hence, the take-home lesson is that there is no good reason to think that the brain, or, more broadly speaking, the nervous system is not a computer. But it is not a *mere* computer: It is physically embedded in its environment and interacts physically with its body, and, for that, it also needs a peripheral nervous system and cognitive representations. Yet there is nothing that denies computationalism here. Most criticisms of computationalism therefore fail. Adhering to them is probably a matter of ideology rather than rational debate.

ACKNOWLEDGMENTS

The author wishes to thank Ron Chrisley, Tomasz Korbak, Martin Hughes, Panu Raatikainen, Błażej Skrzypulec, and anonymous reviewers for their comments. A shorter version of this paper was included in the Proceedings of the 39th Annual Meeting of the Cognitive Science Society in London (see footnotes 2 and 3 for significant changes in this paper from its original version).

REFERENCES

AKAGI, Mikio. 2017 "Rethinking the Problem of Cognition." *Synthese*, April 9, 2017, 1–24. DOI: https://doi.org/10.1007/s11229-017-1383-2.

BAARS, Bernard J. 1988. *A Cognitive Theory of Consciousness*. Cambridge, New York: Cambridge University Press.

BAR-HILLEL, Yehoshua. 1964. "A Demonstration of the Nonfeasibility of Fully Automatic High Quality Translation." In *Language and Information*, 174–79. Reading, Mass.: Addison-Wesley.

BARRETT, Louise. 2016. "Why Brains Are Not Computers, Why Behaviorism Is Not Satanism, and Why Dolphins Are Not Aquatic Apes." *The Behavior Analyst* vol 39, issue 1: 9–23. DOI: https://doi.org/10.1007/s40614-015-0047-0.

BARRETT, Louise, Thomas V. POLLET, & Gert STULP. 2014. "From Computers to Cultivation: Reconceptualizing Evolutionary Psychology." *Frontiers in Psychology* 5 (January 2014): 867–867. DOI: https://doi.org/10.3389/fpsyg.2014.00867.

BICKHARD, Mark H. 2009. "The Interactivist Model." *Synthese* vol. 166, no. 3: 547–91. DOI: https://doi.org/ 10.1007/s11229-008-9375-x.

BICKHARD, Mark H., & Loren TERVEEN. 1995. *Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution*. Amsterdam, Lausane, New York, Oxford, Shannon, Tokyo: North-Holland.

BLOCK, Ned. "The Mind as the Software of the Brain." In *An Invitation to Cognitive Science*, edited by Daniel N. Osherson, Lila Gleitman, & Stephen Kosslyn. Cambridge, Mass.: MIT Press, 1995.

CHALMERS, David J. 2011. "A Computational Foundation for the Study of Cognition." *Journal of Cognitive Science*, no. 12: 325–59.

CHALMERS, David J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

CHEMERO, Anthony. 2003 "Information for Perception and Information Processing." *Minds and Machines* vol. 13 577–88.

CHRISLEY, Ronald L. 1994. "Why Everything Doesn't Realize Every Computation." *Minds and Machines* vol. 4, no. 4: 403–20. DOI: https://doi.org/10.1007/BF00974167.

CLEEREMANS, Axel. 2005. "Computational Correlates of Consciousness." *Progress in Brain Research* vol. 150: 81–98. DOI: https://doi.org/10.1016/S0079-6123(05)50007-4.

COLLIER, John D. 1999. "Causation Is the Transfer of Information." In Howard Sankey (ed.). *Causation, Natural Laws and Explanation*, 279–331. Dordrecht: Kluwer.

COPELAND, B. Jack. 1996. "What Is Computation?". *Synthese* vol. 108, no. 3: 335–59.

CUMMINS, Robert, & Martin ROTH. 2012. "Meaning and Content in Cognitive Science." In Richard Schantz (ed.). *Prospects for Meaning*, 365–82. Berlin & New York: de Gruyter.

DAUGMAN, John. 1990. "Brain Metaphor and Brain Theory." In Eric L. SCHWARTZ (ed.). *Computational Neuroscience*, 9–18. Cambridge, Mass: MIT Press, 1990.

DEACON, Terrence W. 2012. *Incomplete Nature: How Mind Emerged from Matter*. New York: W.W. Norton & Co.

DENNETT, Daniel C. 2017. *From Bacteria to Bach and Back: The Evolution of Minds*. London: Allen Lane.

DENNETT, Daniel C. 2005. *Sweet Dreams. Philosophical Obstacles to a Science of Consciousness*. Cambridge, Mass.: MIT Press.

DRESNER, Eli. 2010. "Measurement-Theoretic Representation and Computation-Theoretic Realization." *The Journal of Philosophy* vol. 107, no. 6: 275–92.

DRETSKE, Fred I. 1986. "Misrepresentation." In Radu BOGDAN (ed). *Belief: Form, Content, and Function*, 17–37. Oxford: Clarendon Press.

DREYFUS, Hubert. 1972. *What Computers Can't Do: A Critique of Artificial Reason*. New York: Harper & Row Publishers.

DREYFUS, Hubert. 1979. *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge Mass.: MIT Press.

EDELMAN, Gerald M. 1992. *Bright Air, Brilliant Fire: On the Matter of the Mind*. New York, N.Y.: BasicBooks.

EGAN, Frances. 1995. "Computation and Content." *The Philosophical Review* vol. 104, no. 2: 181–181. DOI: https://doi.org/10.2307/2185977.

EKMAN, Paul. 2003. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York: Times Books.

EPSTEIN, Robert. 2016. "The Empty Brain." *Aeon*, May 18, 2016. https://aeon.co/essays/your-brain-does-not-process-information-and-it-is-not-a-computer.

FRESCO, Nir. 2010. "Explaining Computation Without Semantics: Keeping It Simple." *Minds and Machines* vol. 20, no. 2: 165–81. DOI: https://doi.org/10.1007/s11023-010-9199-6.

FRESCO, Nir. 2014. *Physical Computation and Cognitive Science*. (Studies in Applied Philosophy, Epistemology and Rational Ethics). Berlin, Heidelberg: Springer-Verlag.

FRESCO, Nir, & Marty J. WOLF. 2013. "The Instructional Information Processing Account of Digital Computation." *Synthese* vol. 191, no. 7: 1469–92. DOI: https://doi.org/10.1007/s11229-013-0338-5.

FRISTON, Karl. 2012, "A Free Energy Principle for Biological Systems." *Entropy (Basel, Switzerland)* vol. 14, no. 11: 2100–2121. DOI: https://doi.org/10.3390/e14112100.

FROESE, Tom, & John Robert STEWART. 2010. "Life After Ashby: Ultrastability and the Autopoietic Foundations of Biological Autonomy." *Cybernetics & Human Knowing* vol. 17, no. 4: 7–50.

GALLISTEL, C.R., & Adam Philip KING. 2010. *Memory and the Computational Brain*. Chichester: Wiley-Blackwell.

GIBSON, James J. 1986. *The Ecological Approach to Visual Perception*. Hove: Psychology Press.

GOLDINGER, Stephen D., Megan H. PAPESH, Anthony S. BARNHART, Whitney A. HANSEN, & Michael C. HOUT. 2016. "The Poverty of Embodied Cognition." *Psychonomic Bulletin & Review* vol. 23, no. 4: 959–78. DOI: https://doi.org/10.3758/s13423-015-0860-1.

HAMEROFF, Stuart R. 2007. "The Brain Is Both Neurocomputer and Quantum Computer." *Cognitive Science* vol. 31, no. 6: 1035–45.

HARNAD, Stevan. 1990. "The Symbol Grounding Problem." *Physica D* 42: 335–46.

HAUGELAND, John. 1985 *Artificial Intelligence: The Very Idea*. Cambridge, Mass.: MIT Press.

KRAJEWSKI, Stanisław. 2007. "On Gödel's Theorem and Mechanism: Inconsistency or Unsoundness Is Unavoidable in Any Attempt to 'Out-Gödel' the Mechanist." *Fundamenta Informaticae* vol. 81, no. 1–3: 173–81.

LAKOFF, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

LEIBNIZ, Gottfried Wilhelm. 1991. *The Monadology*. Translated by Robert Latta. Raleigh, NC; Boulder, Colo.: Alex Catalogue; NetLibrary.

LUCAS, J.R. 1961. "Minds, Machines and Gödel." *Philosophy* vol. 36, no. 137: 219–27.

LUPYAN, Gary. 2013. "The Difficulties of Executing Simple Algorithms: Why Brains Make Mistakes Computers Don't." *Cognition* vol. 129, no. 3: 615–36. DOI: https://doi.org/ 10.1016/ j.cognition.2013.08.015.

LURIJA, Aleksandr Romanovič. 2002. *The Mind of a Mnemonist: A Little Book about a Vast Memory*. Cambridge, Mass.: Harvard University Press.

MACKAY, Donald MacCrimmon. 1969. *Information, Mechanism and Meaning*. Cambridge: MIT Press.

MARR, David. 1982. *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: W.H. Freeman and Company.

MCDERMOTT, Drew V. 2001. *Mind and Mechanism*. Cambridge, Mass.: MIT Press.

MIŁKOWSKI, Marcin. 2007. "Is Computationalism Trivial?" In Gordana Dodig CRNKOVIC & Susan STUART (eds.), *Computation, Information, Cognition — The Nexus and the Liminal*, 236–46. Newcastle: Cambridge Scholars Press. http://marcinmilkowski.pl/downloads/is-computationalism- trivial.pdf.

MIŁKOWSKI, Marcin. 2011. "Jak udawać dualistę, wprowadzając epicykle do funkcjonalizmu." *Przegląd Filozoficzny – Nowa Seria*, no. 1 (77): 27–45.

MIŁKOWSKI, Marcin. 2012. "Limits of Computational Explanation of Cognition." In Vincent C. MÜLLER, *Philosophy and Theory of Artificial Intelligence*, 69–84. Berlin, Heidelberg: Springer. http://www.springerlink.com/content/k6w34j70459wv782/.

MIŁKOWSKI, Marcin. 2013. *Explaining the Computational Mind*. Cambridge, Mass.: MIT Press.

MIŁKOWSKI, Marcin. 2014. "Computational Mechanisms and Models of Computation." *Philosophia Scientiae* vol. 18, no. 18–3: 215–28. DOI: https://doi.org/10.4000/ philosophiascientiae. 1019.

MIŁKOWSKI, Marcin. 2017. "Situatedness and Embodiment of Computational Systems." *Entropy* 19, no. 4 (April 7, 2017): 162. https://doi.org/10.3390/e19040162.

MILLIKAN, Ruth Garrett. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, Mass.: MIT Press.

MÜLLER, Vincent C. (2014). "Pancomputationalism: Theory or metaphor?" In Ruth HAGENGRUBER & UWE V. RISS (eds.). *The relevance of philosophy for information science* (213–222). London, New York: Routledge.

NAGY, Naya, & Selim AKL. 2011. "Computations with Uncertain Time Constraints: Effects on Parallelism and Universality." In Cristian CALUDE, Jarkko KARI, Ion PETRE, & Grzegorz ROZENBERG (eds.). *Unconventional Computation*, 6714:152–63. (Lecture Notes in Computer Science). Berlin, Heidelberg: Springer. DOI: http://dx.doi.org/10.1007/978-3-642-21341-0_19.

NEWELL, Allen. 1980. "Physical Symbol Systems." *Cognitive Science: A Multidisciplinary Journal* vol. 4, no. 2: 135–83. DOI: https://doi.org/10.1207/s15516709cog0402_2.

NEWELL, Allen, & Herbert A. SIMON. 1972. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

NOWAKOWSKI, Przemysław Robert. 2017. "Bodily Processing: The Role of Morphological Computation." *Entropy* vol. 19, no. 7: 295. DOI: https://doi.org/10.3390/e19070295.

PENROSE, Roger. 1989. *The Emperor's New Mind*. London: Oxford University Press.

PICCININI, Gualtiero. 2008. "Computation without Representation." *Philosophical Studies* vol. 137, no. 2: 205–41. DOI: https://doi.org/10.1007/s11098-005-5385-4.

PICCININI, Gualtiero. 2010. "The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism." *Philosophy and Phenomenological Research* vol. 81, no. 2: 269–311. DOI: https://doi.org/10.1111/j.1933-1592.2010.00356.x.

PICCININI, Gualtiero. 2015. *Physical Computation: A Mechanistic Account*. Oxford: Oxford University Press.

PICCININI, Gualtiero, & Sonya BAHAR. 2013. "Neural Computation and the Computational Theory of Cognition." *Cognitive Science* vol. 37, no. 3: 453–88. DOI: https://doi.org/10.1111/cogs.12012.

PRESTON, John, & Mark BISHOP. 2002. *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*. Oxford; New York: Clarendon Press.

PUTNAM, Hilary. 1960. "Minds and Machines." In Sidney Hook (ed.). *Dimensions of Mind*. New York University Press.

PUTNAM, Hilary. 1991. *Representation and Reality*. Cambridge, Mass.: The MIT Press, 1991.

PYLYSHYN, Zenon W. 1984. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, Mass.: MIT Press.

ROSENBLATT, Frank. 1958 "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain." *Psychological Review* 65, no. 6: 386–408. DOI: https://doi.org/10.1037/h0042519.

ROSENTHAL, David. 2005. *Consciousness and Mind*. Oxford, New York: Oxford University Press.

SCHEUTZ, Matthias. 1996. "When Physical Systems Realize Functions…." *Minds and Machines* vol. 9, no. 2: 1–34. DOI: https://doi.org/10.1023/A:1008364332419.

SEARLE, John R. 1980. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3, no. 03 (February 1980): 1–19. https://doi.org/10.1017/S0140525X00005756.

SEARLE, John R. 1990. "Is the Brain's Mind a Computer Program?". *Scientific American*, January: 26–31.

SEARLE, John R. 1992. *The Rediscovery of the Mind*. Cambridge, Mass.: MIT Press, 1992.

SHAGRIR, Oron. 2010a. "Brains as Analog-Model Computers." *Studies In History and Philosophy of Science Part A* vol. 41, no. 3: 271–79. DOI: https://doi.org/10.1016/j.shpsa.2010.07.007.

SHAGRIR, Oron. 2011. "Towards a Modeling View of Computing." In G. DODIG-CRNKOVIC & M. BURGIN (eds.). Information and Computation (381–391). Singapore: World Scientific Publishing. DOI:10.1142/9789814295482_0014.

SHANAHAN, Murray. 1997. *Solving the Frame Problem: A Mathematical Investigation of the Common Sense Law of Inertia*. Cambridge, Mass.: MIT Press.

SIEGELMANN, Hava T., & Eduardo D. SONTAG. 1994. "Analog Computation via Neural Networks." *Theoretical Computer Science* vol. 131, no. 2: 331–60. DOI: https://doi.org/10.1016/ 0304-3975(94)90178-3.

THOMPSON, Evan. 2007. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, Mass.: Harvard University Press.

TONONI, Giulio. 2004. "An Information Integration Theory of Consciousness." *BMC Neuroscience* vol. 5, no. 1. https://doi.org/10.1186/1471-2202-5-42.

WEIZENBAUM, Joseph. 1976. *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: W.H. Freeman, 1976.

WHEELER, Michael. 2005. *Reconstructing the Cognitive World*. Cambridge, Mass.: MIT Press.

ZAUNER, Klaus-Peter, & Michael CONRAD. 1996. "Parallel Computing with DNA: Toward the Anti-Universal Machine." In Hans-Michael VOIGT, Werner EBELING, Ingo RECHENBERG, & Hans-Paul SCHWEFEL (eds.). *Parallel Problem Solving from Nature – PPSN IV*, 1141:696–705. (Lecture Notes in Computer Science). Berlin, Heidelberg: Springer. DOI: http://dx.doi.org/10.1007/3-540-61723-X_1033.

OBJECTIONS TO COMPUTATIONALISM: A SURVEY

Summary

In this paper, the Author reviewed the typical objections against the claim that brains are computers, or, to be more precise, information-processing mechanisms. By showing that practically all the popular objections are based on uncharitable (or simply incorrect) interpretations of the claim, he argues that the claim is likely to be true, relevant to contemporary cognitive (neuro) science, and non-trivial.

ZARZUTY WOBEC KOMPUTACJONIZMU — PRZEGLĄD

Streszczenie

W artykule Autor przyjrzał się typowym zastrzeżeniom przeciwko twierdzeniu, że mózgi to komputery, a ściślej — mechanizmy przetwarzania informacji. Pokazując, że praktycznie wszystkie popularne obiekcje są oparte na nieżyczliwych (lub po prostu niepoprawnych) interpretacjach tego twierdzenia, uznaje, że twierdzenie to prawdopodobnie będzie prawdziwe, istotne dla współczesnej (neuro)kognitywistyki i nietrywialne.

Key words: computationalism; computational theory of mind; representation; computation; modeling.

Słowa kluczowe: komputacjonizm; komputacjonalna teoria umysłu; reprezentacja; komputacja; modelowanie.

Information about Author: Dr hab. MARCIN MIŁKOWSKI — Department of Logic and Cognitive Science at the Institute of Philosophy and Sociology of the Polish Academy of Sciences; address for correspondence: ul. Nowy Świat 72, 00-330 Warszawa; e-mail: mmilkows@ifispan.waw.pl