Oluwagbemiga Odebode

Professor Michel

24.09 Minds and Machines

07 May 2026

The Brain as a Computer:

Meaningful Information from Meaningless Computation

Introduction **3**
Part One: The Computational Theory of Mind **4**

1.1: The computational mind 4

1.2: The causal and the formal 5

Part Two: Machine Functionalism **6**

2.1: The functional kind of computation 6

2.2: The insufficiency of function 7

The Blockhead Thought Experiment 8

The China Brain Objection 8

2.3: The nature of thought 9

The Productivity of Thought 9

The Systemacity of Thought 10

Part Three: Representation, Unification, and Scientific Success **10**

3.1: The search for meaning 10

The Representational Theory of Mind 11

3.2: The applications 11

Conclusion **12**

Works Cited **13**

Introduction

In this essay, I argue that complex cognitive tasks can be decomposed to the point of performance by mindless, syntactic agents. More specifically, I will defend that the computational theory of mind (CTM) provides the strongest philosophical and scientifically actionable explanation of the mind. Further, I will argue that, when combined with the representational theory of mind (RTM)–which asserts that thoughts are represented with internal symbols–a strong, unified computational theory of mind is plausible.

In part one, I explain the argument for, stipulations of, and assumptions made by CTM: that the causal structure of a physical system can mirror the formal structure of a computation, allowing for meaningful computations to arise from meaningless, syntactic processes. I clarify the difference between a syntactic and semantic system, and explain why a strong theory of mind must account for both. Then I analogize complex cognition to Turing machines to demonstrate how exactly a semantically blind system could produce semantic information.

In part two, I outline the ideology of machine functionalism first argued by Hilary Putnam. I urge the reader to consider the mindfulness of a bee, a dog, or even a silicon based Martian and make the argument for a multiple realizable–medium independent–theory of mind. While I concur that the complexity of the brain may lend itself to a functionalist explanation, I demonstrate that Putnam's functionalism cannot explain the infinitely productive and systemic nature of thought, a vital feature of the mind.

Finally, in part three, I elaborate on the representational theory of mind, emphasizing its strength in combination with CTM to explain what machine functionalism cannot. I assert the value of a unified theory of mind (UTM) that combines these two ideologies and briefly demonstrates its scientific success to date.

Part One: The Computational Theory of Mind

1.1: The computational mind

How do you perform the daily, but endless complex cognitive task of making a decision? Each subtask which composes that decision–whether to do *action A* or *action B*–can be further decomposed with subtasks of their own; perhaps there is an *action Aa* which must be evaluated first, an *Ab* that follows, and so on. One can imagine assigning each of these subtasks to a mental agent, who will evaluate their task and pass information upstream. As each task is subdivided, it becomes simpler, and therefore actionable by increasingly less intelligent agents, until intelligence is no longer required. The apparent decomposability of these complex cognitive tasks into mindlessly simple subtasks is the hallmark of the computational theory of mind: cognition from meaningless processes.

One might, reflectively, and fairly, pose the question of how exactly a mindless mechanism can accomplish anything. In order to answer this question, I must assert the difference between syntactic and semantic processing. I first define syntax as the formal structure and the rules that govern the manipulation of symbols–without regard for meaning–and semantics as the meaning of those symbols. It is crucial to understand that these two modalities are not equal. Many symbols are similar in one regard, but vastly different in the other; much like how command shortcuts can differ wildly in syntactic form between processing systems (the CTRL + V of Windows as opposed to the CMD + V of MacOS), but perform identical semantic functions (performing the paste operation). To assert the mind as a machine, as CTM does, is to argue that, at its simplest subagent level, it acts on syntactic properties with no regard for their associated semantic meaning.

1.2: The causal and the formal

This dissociation of processing can be analogized to the Turing machine. The Turing machine is an abstract computational model which manipulates symbols on a set of tape according to a finite list of rules. Despite the simplicity of Turing computation, any computer algorithm, regardless of its complexity (assuming boundless memory), can be executed on this purely syntactic machine. Bringing this analogy back to our exploration of the mind, the path to semantics from syntactics becomes clearer. Meaningful computations arise from meaningless processes when the causal structure of the implementing system mirrors the formal structure of the computation.

The best way to understand casual and formal structure is through example. The logical operator "AND" has the functional goal of combining two inputs. Formal transitions are represented by binary code to produce a summative output. It can also be mirrored by the purely physical system of a logic gate. The "AND" logic gate allows physical transitions between input and output information to mirror transitions between numbers–that is, by definition, meaningful addition. In this way, the "AND" gate illustrates how a purely syntactic system can formally implement meaningful operations. This class of systems, those with mirroring between the causal and the formal, demonstrate a possible mechanism by which the human syntactic engine could give rise to semantic content. According to philosopher Ned Block, this mirrored relationship between the causal and the formal "explains how it is that our syntactic engine can drive our semantic engine (Block, 1995)." Rather than ask for credulity in CTM, it is valuable to examine specific subsets of the theory and measure their success as explanations of the mind.

Part Two: Machine Functionalism

2.1: The functional kind of computation

Machine functionalism is a computational theory of mind which asserts that mental states like pain are specifically defined by their function. To understand machine functionalism, it is helpful to imagine an extraterrestrial being, as put forth by Hilary Putnam. Putnam's alien exhibits all of the diagnostic criteria for pain but in place of a brain, it is composed purely of an intricate system of valves and tubes. Although it lacks the carbon-based biological substrate of a brain, it seems bizarre to deny this otherworldly creature its experience of pain. One does not even have to appeal to outer space. Simply examining the oceans or the sky will reveal a host of completely "alien" creatures with disparate physiology to our own. This raises the question of what exactly pain, or any other mental state, is, and who exactly can experience them. To answer this, Putnam put forward the idea of multiple realizable mental states–that there may be many different ways of having these mental properties because rather than being defined by their form, they are defined by their function. Instead of denying our earthly and extraterrestrial companions their pain, Putnam posits that so long as they are in an analogous version of the pain state, we are all in pain the same.

Putnam recognized that if mental states were multiple realizable, then something about their fundamental nature must lend them to be. Functional kinds–properties which are defined by their common function–are multiple realizable, because function can always be realized in numerous ways.

Through this line of thought, Putnam concluded that mental states are simply functional states. By the purely functionalist definition, the mind behaves as follows: an input stimuli results in a functional state, which may then trigger a behavioral output and the transition into another subsequent functional state. This idea should look familiar, because this is the Turing machine analogy of the broader CTM.

2.2: The insufficiency of function

However, adopting this conclusion opens the theory of mind to a whole new host of problems. Function alone, as the following thought experiments will demonstrate, is not sufficient to a mind. Imagine a creature which is functionally identical to a person, but seemingly mindless, first put forth by titular philosopher Ned Block, coined the Blockhead. The Blockhead is simply a large database of stimuli mapped to responses. For every possible conversational input, there is a precomputed response. Whether you ask question *A*, prompt opinion *B,* or conversate on topic *C*, the Blockhead simply searches the table and returns the corresponding reply. Despite functionally mirroring the behavior of a mindful creature, it is implausible for the Blockhead to have mental states. Block's argument can be formalized as follows:

Premise 1: Machine functionalism states that mental states are functional states
Premise 2: The Blockhead has functional states
Premise 3: The Blockhead does not have mental states
Conclusion: Machine functionalism therefore must be false

The Blockhead Thought Experiment

In response, the functionalist may reject this conclusion by arguing that simple input/output mapping of the Blockhead is insufficient for mental states because these mental states are additionally recursively defined by their relations to one another. Even if we accept this relational necessity put forth by the functionalist, the fact that functionalism implies multiple realizability means that, in theory, we could create a mindful computer out of anything.

The next objection put forth by Block is called the China brain objection. Take, for example, humans completing a simple task, like flipping a coin. With a large enough population, say, the population of China, it is perfectly conceivable to simulate the inner workings of the human brain with people alone, but no one would attribute to this glorified field day, a mind. Similarly, the China brain objection can be written as:

> Premise 1: Machine functionalism states that mental states are functional states
> Premise 2: The China brain simulates the functional states of the brain
> Premise 3: The China brain does not have mental states
> Conclusion: Machine functionalism therefore must be false

The China Brain Objection

Being in the right functional state may be necessary to having a mind, but the thought experiments put forth by Block elucidate the insufficiency of this condition alone.

2.3: The nature of thought

Another issue for machine functionalism highlighted by Block and American philosopher Jerry Fodor (1972) involves the productivity and systematicity of thought. The productivity of thought is the ability of humans to entertain a potentially infinite quantity of propositions over a lifetime. Machine functionalism identifies mental states with machine states, of which there are finitely many. Although a real human will only ever entertain a finite number of thoughts due to constraints on lifespan and memory, unless the machine functionalist proposes some psychological law restricting the set of entertainable thoughts, there are simply not enough machine states to satisfy the injective mapping between the infinitely entertainable and finitely machinable. The productivity objection is:

> Premise 1: Humans possess a capacity for indefinitely extensible thought.
> Premise 2: There exists only a finite number of machine states
> Premise 3: Machine functionalism identifies mental states with machine states
> Conclusion: Machine functionalism is incompatible with human productivity

The Productivity of Thought

The systemacity of thought argues that entertaining one thought may be correlated with the ability to entertain another. To the machine functionalist, the thought *A action B,* and *B action A* are unstructured machines states without systemic relation to one another. The thought that *A* is able to act on *B*, and that, in the same way, *B* can act on *A*, intuitively, shares some systemic relation, one which machine functionalism neglects to explain. Likewise, the systematicity objection can be understood as:

> Premise 1: There exists some intuitive system relation between certain thoughts
> Premise 2: Machine functionalism asserts that machine states are unstructured
> Premise 3: Machine functionalism asserts that these thoughts are machine states
> Conclusion: Machine functionalism is incompatible with human systemacity

Although the machine functionalist may reject this burden of explanation, the best possible theory of mind would not neglect such an essential feature of thought. These objections provide a strong impetus to not only pursue an alternative theory over machine functionalism, but to unify it with a representation theory that accommodates the productivity and systematicity fundamental to the mind.

Part Three: Representation, Unification, and Scientific Success

3.1: The search for meaning

Advanced by Fodor, the representational theory of mind calls attention to the symbols manipulated in Turing analogous computation. RTM argues that thoughts are represented via internal symbols or mental representations. Fodor postulates a system of both primitive and complex–built from primitive–representations. These representations are compositional; the meaning of representation is not only the sum of its parts (primitives) but also the ways in which they are combined.

RTM readily accommodates the productivity and systematicity gaps of machine functionalism. By postulating a finite set of primitive representations, combinable in infinity of ways, RTM provides a medium for the dimensionality and abundance of mind. Additionally, empowered by the compositional nature of the theory, RTM allocates meaning not only in symbols, but in the ways they are combined in accordance with rules. By treating thoughts as relations to these complex mental symbols, RTM robustly explains both productivity and systematicity. The thought *A action B,* and *B* *action A* now share the relational structure that machine functionalism failed to provide. The formal argument is:

Premise 1: Human thought is both productive and systemic
Premise 2: Thoughts are represented by primitive and complex internal symbols
Premise 3: Thought is productive via infinite possible combinations of primitives
Premise 4: Thought is systemic because there is meaning in the way primitives are combined
Conclusion: RTM is compatible with human systematicity and productivity

The Representational Theory of Mind

When combined with CTM, RTM allows for the creation of a unified theory of mind, one in which mindful activity is enacted by Turing computations over representational symbols. These symbols are stored in memory and are manipulated in accordance with a set of rules. Mental states are therefore computationally characterized but manipulate mental symbols in a variety of specific ways.

3.2: The applications

Furthermore, the UTM has been scientifically fruitful. In addition to its ability to explain high-level human thought, UTM has been successful in the study of nonhuman animals. In a series of experiments conducted in 1964, neurophysiologists David Hubel and Torsten Wiesel demonstrated that neuronal signals carry mental representations in cats. Certain groups of neurons would only be selectively active for certain regions–perceptual fields—of space, or quality of stimuli–only horizontal bars, for example. Because these neuronal signals carry information about the outside world, they are, by definition, mental representations: internal structures that stand for other things. This discovery was extremely influential. Beyond providing testimony to the quality of UTM, the work of Hubel & Wisel inspired decades of similar experiments which now serve as the basis of computational cognitive neuroscience.

Conclusion

In this essay, I set out to accomplish three broad goals. First, to introduce the computational theory of mind and exemplify it with analogy, then to pay respects to machine functionalism as specialized form of computational mindfulness, and finally to integrate the structure of CTM with the representational theory of mind to create a comprehensive and successful unified theory of mind. I first demonstrated the framework by which CTM explains semantic processes arising from syntactic agents via the analogy of the Turing machine and introduced the class of systems CTM explains: those with mirrored causal and formal structure. The strategy I used to approach the second goal comes in the form of multiple realizable mental states and the introduction of machine functionalism as a specialized computational theory of mind. In my review of Putnam's machine functionalism, I elucidated the weaknesses of machine functionalism alone to explain the productive and relational structure inherent to thought. Finally, I elaborated on RTM and argued in favor of its joining with CTM to create a comprehensive, and scientifically successful, unified theory of mind. Through the execution of these three goals, I have defended CTM as a strong contender for the theory of mind.

Works Cited

Block, N. (1995). The mind as the software of the brain. In E. E. Smith & D. N. Osherson (Eds.), Thinking: An invitation to cognitive science (2nd ed., pp. 377–425). The MIT Press.

Block, N. and J. Fodor, 1972, "What Psychological States Are Not", The Philosophical Review, 81: 159–181.

Levin, Janet, "Functionalism", The Stanford Encyclopedia of Philosophy (Summer 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), https://plato.stanford.edu/archives/sum2023/entries/functionalism/.

Michel, M. (2026). Lecture 4: Functionalism [PowerPoint slides]. Massachusetts Institute of Technology Minds and Machines Canvas: https://canvas.mit.edu

Michel, M. (2026). Lecture 5: The Computational Theory of Mind [PowerPoint slides]. Massachusetts Institute of Technology Minds and Machines Canvas: https://canvas.mit.edu

Place, U.T. (1956), Is Consciousness a Brain Process? British Journal of Psychology, 47: 44-50. https://doi.org/10.1111/j.2044-8295.1956.tb00560.x

Putnam, Hilary. "The nature of mental states." Readings in philosophy of psychology 1 (1980): 223-231.

Rescorla, Michael, "The Computational Theory of Mind", The Stanford Encyclopedia of Philosophy (Fall 2025 Edition), Edward N. Zalta & Uri Nodelman (eds.), https://plato.stanford.edu/archives/fall2025/entries/computational-mind/

Sipser, Michael (2012) [1997]. "Chapter 3: The Church–Turing Thesis". Introduction to the Theory of Computation (Third ed.). Cengage Learning. pp. 165–192. ISBN 978-1-133-18779-0.

Smart, J. J. C. "Sensations and Brain Processes." The Philosophical Review 68, no. 2 (1959): 141–56. https://doi.org/10.2307/2182164.

Smart, J. J. C., "The Mind/Brain Identity Theory", The Stanford Encyclopedia of Philosophy (Fall 2025 Edition), Edward N. Zalta & Uri Nodelman (eds.), https://plato.stanford.edu/archives/fall2025/entries/mind-identity/

Stone, Harold S. (1972). Introduction to Computer Organization and Data Structures (1st ed.). New York: McGraw–Hill Book Company. ISBN 0-07-061726-0.