**Classical Computationalism vs. Connectionism**

**Is the Mind Best Understood as a Digital Computer?**

Riddhi Patel

Minds, Machines, & Persons

December 7, 2023

## Abstract

Technology is a prevalent and powerful tool that helps us in many aspects of our everyday lives. We use the internet and digital computers in various ways, including accessing information, having diverse forms of entertainment, assisting with work, and communicating with others effectively. But how did we create these technologies, and how do they relate to our minds? This question has intrigued philosophers since the 19th century when computers could perform simple arithmetic functions. It leads us to this simple question: is the mind best understood as a digital computer? To answer this question, I will assess and analyze the two known theoretical models surrounding the mind: Classical Computationalism and Connectionism.

In this paper, I argue that the mind is not best understood as a digital computer but rather as an interconnected, distributed network of neurons that communicate with each other through electrical signals, otherwise known as Connectionism. This theory gains support from empirical evidence from neuroscience and successfully reproduced with the generative A.I. models we know today. I argue that Connectionism can account for the complexity and flexibility of cognition while avoiding the limitations of the computationalist view, which assumes that the mind operates like a digital computer algorithm with symbol manipulation to create meaning and behaviors. First, Connectivism supports neuroscientific evidence, while Computationalism does not. Second, Connectivism is successfully reproduced through Generative A.I. models today, while Computationalism is not. Therefore, the mind is not best understood as a computer.

## Introduction

As technology rapidly evolves throughout the centuries, philosophers ponder how technology– digital computers–correlates with our minds. Various philosophers have made strides to answer this idea, which leads to a singular question: Is the mind best understood as a digital computer? To answer this question from my point of view, I will refer to two theoretical models known in the Philosophy of Mind: Classical Computationalism and Connectionism.

In this paper, I argue that the mind is not understood best as a computer. Better yet, the mind is represented and understood as an interconnected system of neurons that communicate with each other through electric signals–otherwise known as Connectionism. Furthermore, to support this argument, I explain that the Connectionist theory model supports physical neuroscientific research as well as achieving successful replication with well-known generative A.I. models today such as GPT-3. Additionally, I argue that the Connectionist theory model will address cognitive-related ideas and cover the possible limitations of Computationalism– the idea that the mind is a computer-like machine with algorithms that use symbol manipulation. Therefore with these arguments, I conclude that the mind is not best understood as a digital computer.

## Understanding the Basics

Before diving deeper into this paper, some concepts need to be clarified. The following sections will explain the two theoretical models that this paper focuses on to answer the question: Classical Computationalism and Connectionism, as well as small terminology such as computers and representation. The sections will also review the basic background history of the models and other relevant information that is important to understand for the rest of this paper.

## I. Computers and Representation

We may know computers as an everyday technology ranging from desktops and laptops to even our mobile phones–a screen with a keyboard. The components inside a computer contain wires and circuits. However, in the philosophy of mind, a computer is defined as a device that processes representations systematically. It is different from our current understanding of computers. Representations under this context represent an unworldly phenomenon that takes place in our minds to create meaning or stand for something when we observe or interact with our environment (Crane, 2016, p.58). In other words, computers are a way to digest information about our environment methodologically. I will call a regular computer (that we know today as) a digital computer for this reason.

## II. Classical Computationalism

The 19th century was progressive with mathematical advancements: Leibniz proposed a universal language as well as created binary notation, George Boole invented boolean algebra to express elementary logical relationships, and Alan Turing's contribution with the Turing machine, the Turing test, and the concepts for a universal computer. This logic-based math and theory led to the birth of general-purpose programmable computers in the 20th century (Crane, 2016, p.80). As a result, this paved the way for a new idea of how we may think with our minds using mental representations–based on these general-purpose computers and logic. In other words, develop the idea of the Classical Computationalism model theory. Classical Computationalism, by definition,  is the mind being a computing system that is "realized" in our brains. The theory aims to answer what specific cognitive functions are computed at the computational level, what algorithms are used, and how they are implemented. The human mind is constructed similarly in fashion like a digital computer where symbol manipulation (mental representations) is involved to create meaning, behaviors, etc. Classical Computationalism also supports the “Language of Thought” hypothesis re-introduced by Jerry Fodor in 1975, in which we have an unknown internal language for thoughts in our minds called *mentalese* which is different from the natural languages we know today that communicates thought with logical symbols (Sorensen). In essence, Computationalism inherits the idea that our minds systematically process representations, like an algorithm, to create meaning, behaviors, and interact with our environment–similar to how digital computers function.

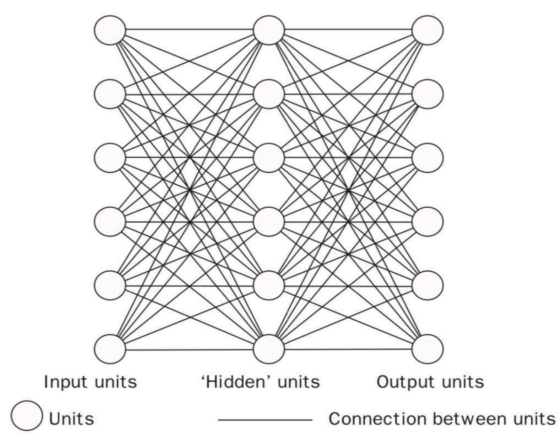
## III. Connectionism

In the 1970s, Classical Computationalism became a dominant theory for the philosophy of mind due to the lack of alternatives that could represent the mind back then. As a result, philosophers sought to find a better alternative to Classical Computationalism by searching for possible limitations within the theoretical model. One primary limitation was the “Language of Thought” hypothesis,  supported by Classical Computationalism. There were beliefs that *mentalese* has some limitations, ranging from how this internal language produces meaning and that these logical rules found in Computationalism do not successfully apply to human thought (unpredictable behaviors) (Crane, 2016, p.105).

This limitation resulted in the revival of the Connectionism theoretical model. Connectionism, by definition, is a type of theory that studies human cognition with the use of mathematical models. These models, otherwise most commonly known as Neural Networks, contain an interconnected set of neurons or “nodes” that act as input/output (Waskan). In addition, synapses connect these neurons (Buckner & Garson, 2019). Each of these connections has different weights associated with them. Neurons communicate with each other through electrical signals, traveling through various nodes and synapses. By communicating through electrical signals, they create meaning. Each neuron or node does not create meaning itself, but together they do. The mental representation is highly distributed within these nodes (Sorensen). Furthermore, this type of mathematical model led to the birth of Artificial Neural Networks (ANNs) and made successes in face recognition, reading, and detecting grammar syntax. Essentially, this theory gained support to help explain cognition and behavior, which is not clarified enough in Classical Computationalism.

**Figure 1.**

*Connectionist Model*




*Note.* From *The Mechanical Mind: A Philosophical Introduction to Minds, Machines and Mental Representation*. [Image], 2016, Routledge, Taylor & Francis Group

**Connectionism and Neuroscience**

Neuroscience is a subfield of Biology where scientists focus on studying the human nervous system–brain, spinal cord, and neurons. Neuroscience aims to answer how the nervous system works to produce human behaviors and keep the human body functioning. To do so, scientists in this discipline study nerve cells, nerve networks, and brain cells to understand how these components interact. It is a physical science that utilizes brain scans, imaging devices, inverted microscopes, and brain tissues (U.S. Department of Health and Human Services).

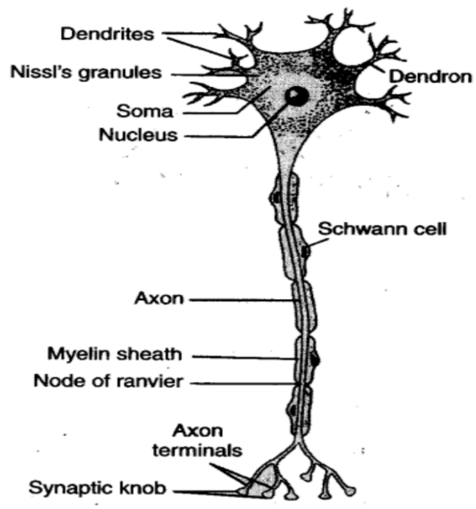
The scientist known as the "Father of Neuroscience" is Ramon y Cajal; he paved the way for modern Neuroscience due to his theory called the *Neuron Doctrine*. With this doctrine, Cajal states that the nervous system consists of microscopic cells, or neurons, that are distributed and can make complex connection patterns. This idea came when Cajal was studying the histology slides (microscopic study of brain tissue, where a section is stained and observed upon a microscope). Furthermore, Cajal states that the information flows within different neurons from the dendrites to the axons. In essence, this theory has gained support as Neuroscience continues to evolve (De Castro, 2019).

Moving forward, a recent study conducted in 2018 by Neuroscientists at the Massachusetts Institute of Technology (MIT) discovered more about dendrites in the human brain. This study was conducted based on their observation of a small section of human brain tissue. Dendrites are long, thin-like neural structures that act as extensions that receive electrical signals from other neurons. Just like axons, dendrites are a distinct part inside a neuron. Furthermore, the scientists researched that dendrites have diverse electrical properties. This means that electrical signals can weaken when it's within the dendrite, and when they do, these dendrites behave independently from the neuron. The discrepancy between electrical signals, according to the neuroscientists at MIT, believes that it may contribute to the understanding of the "enhanced computing power of the human brain". In other words, if an electrical signal within the dendrite weakens, the dendrite will act independently from the neuron and this phenomenon relates to our mind's advanced computing power (Trafton, 2018).

Cajal's doctrine alongside a recent neuroscientific study from MIT suggests that neurons exist and that electrical signals are involved for neurons to communicate with each other to create human behaviors and thoughts. The two studies support the Connectionist theoretical model because, despite the complexities, they illustrate the same basic structures founded in the Connectionist framework. For example, while the MIT study may go more in-depth with the specific complexities of the neuron– such as dendrites, it still shows that the mind, to function and perform specific behaviors- such as exhibiting more computational skills involves electrical signals to travel through the neuron for a phenomenon like this to occur. Additionally, Cajal's doctrine states that information flows from different neurons from their axons to dendrites. In other words, it explains information is connected through a distributed network of neurons. Connectionism may be broadly classified and abstract, founded with their ANNs in many uses such as Artificial Intelligence and Machine Learning; it has solid support and presence within neuroscientific studies starting from its first modern study to even current research. In other words, we have a theoretical model that can be broken down and reduced to physical scientific explanations (reductionism).

**Figure 2.**

*Basic Neuron Structure*




*Note.* From *Microscopic Structure of a Neuron* [Image], by Prasanna, Learn CBSE Forum (https://ask.learncbse.in/uploads/db3785/original/1X/b9362e4848d74f27dfa3b1d41fbacdd22dcab719.png)

## Successful Implementations Using Connectionism

In May 2020, an Artificial Intelligence startup company called OpenAI released their new language model, Generative Pre-trained Transformer 3 (GPT-3 for short). It is one of the first successful language models created by OpenAI. After being trained intensively through text-based data under 175 billion parameters, the following model could generate essays within different writing styles, answer questions, summarize texts, translate texts to different languages, and complete missing texts (Rees, 2022). However, there were significant limitations that OpenAI failed to address in the language model, including its responses containing potential biases and making toxic comments– which was found similarly with their predecessors GPT-1 and GPT-2 (Imamguluyev, 2023). Due to the slight success with GPT-3's performance, with a few more tweaks, GPT-3.5 was released in late November 2022 as a website known today as

ChatGPT;  the website is currently serving as a public testing platform to observe the chat-bot performance among users.

As a result, ChatGPT became significantly popular among diverse audiences, profitable from a business standpoint, and heralded the race for producing more AI-based technologies. That said, the question is, how do these language models work? What is the underlying framework or mechanism for these massive models training under billions of data before being released to us, the public? The answer may sound complicated, but I will break it down. The language models, notably GPT-3 and GPT-3.5, are developed through concepts surrounding "Natural Language Processing" or NLP for short. It is a subdivision subject in Artificial Intelligence that aims to let a computer comprehend, process, and generate in human language. NLP began in the 1950s but made significant progress in the 2000s when neural networks and deep learning models were introduced. GPT-3 and GPT-3.5 rely on these deep learning models and ANNs to comprehend and generate text as an NLP (Imamguluyev, 2023). Therefore, Connectionism laid the groundwork for these mathematical models to thrive, as the theoretical model's original intent is to help us understand how the human mind works.

When Classical Computationalism came to be, it was influenced by the development of the first digital computers. Philosophers thought that our minds may be similar in characteristics to the machines built by humans and that our brain runs similarly to an algorithm to process representations and create meanings. In other words, a new way of thinking about how the mind works developed because of humanity's ability to invent a machine that can calculate and perform various feats as accurately and efficiently as possible. Meanwhile, Connectionism is introduced to address limitations found in Classical Computationalism and gain a sudden revival. Instead of being influenced by the development of a machine created at the time, Connectionism acted as a building block where the model replicated successfully onto new technological advancements (GPT-3 & 3.5) as it can learn and generalize from data and generate new content based on its observations. In other words, to replicate how we may learn through new technologies we make and not the opposite way.

## Objections

I have argued so far that in comparison to Connectionism, Classical Computationalism lacks support with physical scientific evidence as well as no reproduction of this model in a setting within machines to understand how we learn, which is a crucial characteristic of the human mind besides exhibiting other behaviors such as thought. These two crucial ideas should be present to understand the mind, based on how we study and replicate other inventions. However, it may be possible to undermine these crucial arguments, primarily if Classical Computationalism has substantial scientific evidence and the currently known limitations present within the Connectionist point of view. These objections will be the primary ones that I believe are the most important to address within the scope of this question I aim to answer.

### I. The Best Game in Town

The Best Game in Town is a philosophy paper focused on the "reemergence" of the LoT hypothesis. The authors of this paper are Jake-Quility Dunn, Eric Mandelbaum, and Nicolas Porot. In this paper, they aim to argue that there is substantial evidence found in various fields of Cognitive science, that shows LoT is still a possible relevant explanation for how our minds work, thus, a form of revival in the Classical Computationalism theory. The most notable evidence includes Bayesian computational modeling in Cognitive Psychology, studies done on infants in Developmental Psychology demonstrating complex reasoning ability in infants, and research done in Social Psychology on automatic intuitive cognition in adults. The main argument in question:

> While acknowledging the possibility of other theoretical models (ANNs) that explain the mind and how it works, Dunn, Mandelbaum, and Porot argue that the progressive advancement of LoTs (Classical Computationalism structures) through cognitive science-based evidence will be well-supported over time and illustrate the mind more effectively (Quility-Dunn, Mandelbaum, & Porot, 2022).

I object to this statement. Before moving forward with my reasoning, I acknowledge that there may be scientific evidence to support Classical Computationalism concepts such as LoT, which may come from other sciences and studies that I needed to gain awareness of. However, from the evidence presented in this paper, I question whether the evidence is directly plausible and can conclude in forming the LoT hypothesis. In other words, I did not fully grasp a clear indication of how LoT is involved in these studies other than what the authors are claiming. The authors may have made connections to how LoTs can be modified and better interpreted in different studies ranging from infants in Developmental Psychology to adults in Social Psychology, but these studies do not indicate from the actual research itself the connection. Do we truly use probability and mathematical-related logic in our thoughts to form a decision? Humans, ranging from babies to grown adults, are unpredictable. If we had the probabilities that we did not know about in our minds, we would make better decisions and not be influenced by outside environmental factors easily in our lives. Within the Connectionism model, as we interact constantly with our environmental factors, the neurons in our brains interact with electrical signals that may result in different weight strengths when traveling with the synapses. I believe that depending on the weight of the signal underlies our cognitive behaviors and the unpredictable responses we may make. In essence, it is not perfect, or mathematical. Back to LoTs, we may know some implications that can work in Bayesian models (PLoTs)  and their applications, but that does not demonstrate the characteristics of the mind and how we think, nor is there reduced biological plausibility in this case. The goal is to answer whether or not the mind is understood best as a digital computer; in my perspective, I believe the way to do so is to have physically reducible scientific evidence to explain such an idea is possible.

## II. The Biological Implausibility

Another well-known objection to the Connectionist framework is the Biological Implausibility case:

> Real neurons are much more complex than Artificial Neurons (ANNs). We only have one type of neuron in ANNs (input/output based) but in the human brain, it has been discovered that there are different variations of neurons. ANNs are also claiming to lack other similar characteristics to what is present in the human brain including neurotransmitters, glial cells, and hormones (Sorensen).

I admit that Connectionism is an abstract theoretical model concept. Connectionism intends to have this broad road map-based idea of the mind's workings. If we incorporate every small missing detail involved that is present in the mind into this model, we may be adding unnecessary details if we are aiming to replicate this form of learning to other machines or software. The end goal of this is to answer the question of whether the mind is best understood as a digital computer or not. To best understand how the mind works, we need to understand the foundations of what makes up the mind and how it learns, which is an important characteristic. Disciplines in Artificial Intelligence and Machine Learning have developed their software using ANNs, where the neurons inside these mathematical models may have different activation functions within hidden layers, these will act similarly and address the behavior of how our minds have different neuron characteristics (Foote, 2022). These two disciplines will eventually advance and gear towards developing more complex ANNs that may act even closer in similarity with our minds.

**Conclusion**

The Classical Computationalism framework is a type of approach that aims to answer that the mind is understood best as a computer. However, some limitations may disprove this notion. These limitations range from lacking physically reduced scientific evidence, failing to answer unpredictable human cognitions with representation, and unknown replications of these models to new machinery or software if they do represent the human mind and its fundamental behaviors. Meanwhile, Connectionism can support these limitations that are addressed where the framework is physically reducible within Neuroscience, explaining how our mind is a parallel distribution of units when communicating together to create representations, and replicating a similar behavior of the human mind (learning) using such a model onto developing Generative AI models like GPT-3. Besides that, there is doubt about the Connectionist framework accounting for every cognitive aspect of the mind, as well as possible evidence found in

Cognitive Sciences and Psychology that may support the LoT hypothesis for how the mind works. However, the framework's support rooted in science and successful applications in today's AI technology provide a stronger foundation for how the mind works currently, in comparison to Classical Computationalism. As our understanding of the mind continues to grow, Connectionism will offer a promising direction for future research and development.

**References**

Buckner, C., & Garson, J. (2019, August 16). *Connectionism*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/connectionism/

Crane, T. (2016). *The Mechanical Mind: A Philosophical Introduction to Minds, Machines and Mental Representation*. Routledge, Taylor & Francis Group.

De Castro, F. (2019, April 12). *Cajal and the Spanish Neurological School: Neuroscience would Have been a Different Story Without Them*. Frontiers. https://www.frontiersin.org/articles/10.3389/fncel.2019.00187/full#:~:text=Santiago%20Ram%C3%B3n%20y%20Cajal%27s%20%E2%80%9Cneuron,of%20connections%20(Figure%202A)

Foote, K. D. (2022, January 20). *Artificial Neural Networks: An overview*. DATAVERSITY. https://www.dataversity.net/artificial-neural-networks-overview/

Imamguluyev, R. (2023, March). *The rise of GPT-3: Implications for Natural Language Processing and Beyond*. ResearchGate. https://www.researchgate.net/publication/369775744_The_Rise_of_GPT-3_Implications_for_Natural_Language_Processing_and_Beyond

Prasanna. (n.d.). *Microscopic Structure of a Neuron* [Image]. Learn CBSE Forum. Retrieved https://ask.learncbse.in/uploads/db3785/original/1X/b9362e4848d74f27dfa3b1d41fbacdd22dcab719.png.

Quilty-Dunn J., Mandelbaum E., & Porot N., (2022, December 6). *The Best Game in Town: The Reemergence of the Language-of-Thought Hypothesis Across the Cognitive Sciences: Behavioral and Brain Sciences*. Cambridge Core. https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/best-game-in-town-the-reemergence-of-the-languageofthought-hypothesis-across-the-cognitive-sciences/76F46784C6C07FF52FF45B934D6D3542

Rees, T. (2022, May 1). *Non-human words: On GPT-3 As a Philosophical Laboratory*. MIT Press.https://direct.mit.edu/daed/article/151/2/168/110618/Non-Human-Words-On-GPT-3-as-a-Philosophical

Trafton, A. (2018, October 18). *Electrical Properties of Dendrites Help Explain Our Brain's Unique Computing Power*. MIT News | Massachusetts Institute of Technology. https://news.mit.edu/2018/dendrites-explain-brains-computing-power-1018

Sorensen, D. (n.d). *Minds, Machines, Persons - Classical Computationalism* [PowerPoint slides]. RU-Canvas.https://rutgers.instructure.com/courses/239733/files/33893253?module_item_id=8985938

Sorensen, D. (n.d). *Minds, Machines, Persons - Connectionism* [PowerPoint slides]. RU Canvas. https://rutgers.instructure.com/courses/239733/files/33893226?module_item_id=8985940

U.S. Department of Health and Human Services. (n.d.). *About Neuroscience*. Eunice Kennedy Shriver National Institute of Child Health and Human Development. https://www.nichd.nih.gov/health/topics/neuro/conditioninfo

Waskan, J. (n.d.). *Connectionism*. Internet encyclopedia of philosophy. https://iep.utm.edu/connectionism-cognition/