The emergence of the human brain through bottom-up and top-down procedures

Orlando Melchor-Alonso

Abstract

The emergent properties of the human brain such as learning and memorization are explored. First, one discusses the short comings of representing the brain using artificial neural networks. Then one proceeds to outline one of the first bottom-up methods used to describe the emergent phenomena of the brain known as the Hopfield network. One then discusses why top-down procedures are more successful in describing the emergent states of the brain. Lastly, one outlines the advantages of top-down methods and potential advances in these methods for the future.

1. Introduction

The emergence of the human brain is one of the most perplexing occurrences in evolutionary history. How did modulations in the gene pool give rise to the human brain and what models exist today to describe its complexity? In the same way that the basic building blocks of various materials are the particles that describe them, it is understood that the basic building blocks of the brain are the neurons that promote particular chemical reactions. The current research on this topic seems to indicate that non-interacting neurons are not suffice to exhaust all possible phenomena associated with the brain. Interactions are necessary in order to explain phenomena like memory and learning.

The emergent properties of the human brain have inspired computer scientists, physicists and biologists to model modern day machinery and methodologies after the functionality of the human brain. The question then arises can such computational tools also give insight into the mechanisms behind learning and memorization? In this paper, one explores the effectiveness of artificial neural networks in describing the properties of the human brain. An example of an artificial neural network is shown in Figure 1.

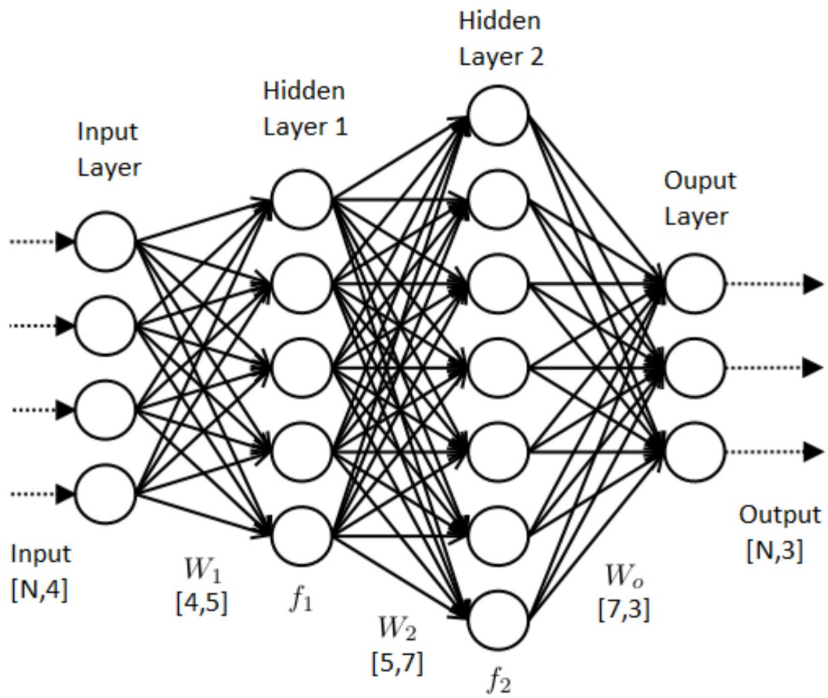

Figure 1: The basic outline of an artificial neural network. The goal of a neural networks is to segment the input data into smaller sub-classes that can interpreted by more segments and later classified. These segments are passed down to hidden layers until they reach the final output. Hidden layers get their name because of their black-box nature. A major problem with neural networks is interpreting which weights affect the output the most. Due to the vast number of connecting edges this is often an impossible task. Therefore, one often abstracts away the problem.

The discussion will begin with an overview of the complexity of neurons that modern methods are attempting to address as well as their inevitable shortcomings. The discussion will continue by highlighting the first bottom-up approaches taken by J.J. Hopfield who created a method of describing learning and memorization using a model reminiscent of the 2D Ising chain. Next, one will discuss the difficulties of bottom-up approaches and their implicit assumptions.

The focus will then shift to the infrastructure behind top-down approaches and why they show more promise than previous bottom-up approaches. Lastly, one will conclude with the progress of modern day top-down approaches and the potential they hold for the future.

2. Inevitable Shortcomings

To begin, one notes that artificial neural networks over-simplify the model for cognition. Many models depends on abstracting away the complicated function of the neuron. The value of the state of a neuron is often taken to be either a 1 if it is fired or a 0 if is not fired. The firing of a neuron is assumed to only be dependent on the activation energy required to fire a neuron. Most models represent this as a threshold that must be surpassed from all incoming connections in order to be fired. The plasticity of the brain, however, can affect the structure of neurons in the brain over time. This can lead to misfiring and is often difficult to model. In fact, few models exist today that have unfixed neuronal morphologies. [2]

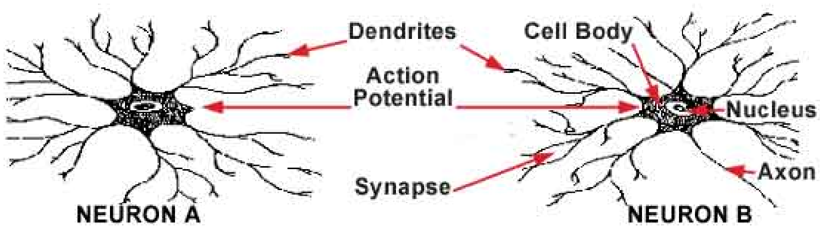

Figure 2: A basic outline of the anatomy of neurons. Communication between neurons occur through the dendrites connecting the cell bodies. If the action potential exceeds a certain threshold, then a signal is fired between neurons.

Physical neurons also consist of multiple branches known as dendrites that connect to other neurons. One of the greatest weaknesses of artificial neural networks is the computational cost required to represent every neuronal connection in the brain. Assuming that neural networks constitute complete graphs, the space complexity of every connection in the brain is at most $\mathcal{O}(n^2)$. Here, n represents the number of neurons. This number is found to be anywhere on the order of approximately 100 billion neurons per person. This fact results in neural networks becoming easily saturated by too many connecting weight edges. Furthermore, a neural network consists of multiple layers known as convolutional layers, pooling layers, and dense layers. Each layer affects the run time of neural networks in its own way. The larger the system size, the greater the impact on the time complexity from these layers.

Lastly, all bottom-up methods greatly suffer from an incomplete picture of neuroscience. The incomplete picture is due to incomplete data on the brain, a lack of predictability in current models, limitations in experiments, and a lack of understanding behind every function of the brain.[1] Any method presented must address these problems, hence why it is difficult to devise a comprehensive bottom-up procedure. Now that one understands the limitations of developing a model for the brain, one can begin to discuss modern bottom-up approaches.

3. Bottom-Up Approaches: The Hopfield Network

The Hopfield network is a method developed by J.J. Hopfield in 1982 that encapsulates the idea of applying an artificial neural network to the brain. The Hopfield network is essentially a glorified Ising model where the state of spins are replaced by the states of a neuron.[3] The weights of a Hopfield network are analogous to site dependent magnetic fields with more than nearest-neighbor interactions. In the Ising model, the network update rule is to randomly flip a bit if it decreases the energy. For the Hopfield network, this analogy can be extended to image reconstruction because of its energy minimization principle. Since the Ising model converges to an energy minimizing state, one can use this method to reconstruct a state that has been perturbed k-times.

In situations where the model has to remember m-memories, it performs better the less memories the network has to ”remember.” This is reminiscent to the behavior in human brains. A human brain is less likely to remember a larger sequence of memories than it is likely to remember a smaller sequence. In addition, the network performs best the less k-perturbations that are applied to the initial image. In other words, the less a memory is corrupted, the more likely one is to remember the image. One can show this result by plotting the hamming distance between the actual result and the the reconstructed image.

One can reconstruct random memories using this method. Following the procedure outlined, one produces the results shown in Figure 3.

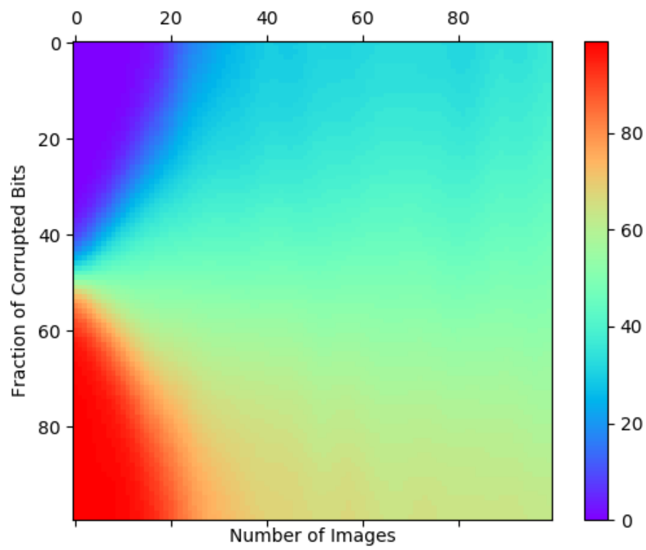

Figure 3: The hamming distance as a function of the number of corrupted bits and the number of random memories. A total of 100 20x20 images are used and the Hamming distance varies from 0 if the images reproduced are identical and 100 if they are completely dissimilar.

This diagram is in (m, k)-space where m is the number of memories to be remembered and k is the number of corrupted bits. This plot shows that the images reproduced are in agreement with what one would expect for a human brain. However, what one also notices is that as the number of corrupted bits increases and the number of memories decreases one is more likely to obtain an inverted image. This results from the fact that the descent direction tends toward two possible minima. One of the minima occurs at the original image and the other at the inverted image. Despite this, the behavior for large k and m show that the network randomly assigns a reconstructed image as the minimum as is expected. This is reminiscent to the difficulty in remembering all of the images. What one discovers from the Hopfield network is that it reproduces images but only if the number of images is small and the number of perturbations are not too large. [4]

Although the Hopfield network does an accurate job of reproducing the behavior of a human brain it is only as good as its implicit assumptions. For instance, the model assumes that the network is not plastic and that it is not subject to any external change. The model also assumes all connections are linear despite the fact that many important phenomena depend on loops and parallelization of the neural connections. Often these problems are obscured by taking the average of non-linear results despite their importance. One also notices that a fundamental break down of the model occurs at large k and small m that is unexpected. The assumption that pathways are undirected also leads to an issue as it seems to imply that the network evolution is reversible. A single connection going one way has the same weight placed on it as a signal going in reverse which seldom is the case. Bottom-up processes for the brain are difficult to construct as it is nearly impossible to cover all possible phenomena that can occur in the brain. In the case of the human brain, the emergent phenomena is too difficult to decipher. Modern approaches seem to acknowledge these difficulties. In the past few years one has seen a shift towards top-down analysis as one must reconstruct the phenomena the brain exhibits.

4. Bottom-Up Approaches: Modern Improvements

The modern approach to brain modeling appear to be hierarchal rule abstraction. In this lens, one attempts to explain an emergent phenomenon by explaining it in terms of other emergent phenomena. This procedure is successful in that the result must reproduce the emergent state. This is accomplished by producing a mathematical correlation between low-level rules and abstract properties of an emergent state. [5]

The difficulty of this method is that it requires a significant level of human intuition to find correlations. It is currently difficult to automate the procedure of rule abstraction as one must have a sufficient enough understanding of the problem to advance. However, in the same manner that one can use supervised learning techniques for bottom-up procedures, one can do the same for top-down analyses as well. Many classification methods can be performed to find correlation mappings between multiple variables and an emergent phenomenon. This procedure is simply a classification task that can be used to uncover meaningful properties of brain processes.

In the case of the human brain, the abstract property is the arising emergent states of the mind such as learning and the low-level rules are the neuron firings. From this point it is difficult to break down these properties and find correlations between the low-level rules so one creates mid-layers and learns the correlation between those instead. These mid-layers could be a result of other emergent properties like language, memory, and reasoning. The hope is to approach a point where one can successfully find the correlation between the lowest sub-layer and its respective parent node.

5. Conclusion

In conclusion what one observes is that old bottom-up methods are ineffective in describing the emergent phenomena in the brain. Although one could implement these methods using complicated neural networks many of the implicit assumptions in the models may be too damaging. Ultimately, one is often left with a limited scope for which a model can apply, leaving the problem unresolved. Other methods like rule abstraction show more promise as they are required to reproduce emergent phenomena. In fact, it was shown that a bottom-up process for the brain is really just a top-down process in disguise. There still remain difficulties, however, in top-down processes namely the need for human intuition. However, there exist supervised learning methods that can provide insight into low-level rules and abstract properties of an emergent state. It also always possible that rule abstraction may not suffice in situations where the system is too complex. Lastly, if no midpoint exists between low-level rules then there are no meaningful reductions in abstract layers.

6. References

[1] Markram, Henry. " Seven Challenges for Neuroscience ." Functional Neurology 28.3 (2013): 145 – 151 . Print.

[2] Ooyen, Arjen van, and Markus Butz-Ostendorf. "The Rewiring Brain - A Computational Approach to Structural Plasticity in the Adult Brain." Egyptian Journal of Medical Human Genetics, Elsevier, 30 June 2017.

[3] Hopfield, J J. "Neural Networks and Physical Systems with Emergent Collective Computational Abilities." Proceedings of the National Academy of Sciences of the United States of America 79.8 (1982): 2554–2558. Print.

[4] Manandhar, S. K., and R. Sadananda. "Effect of Hamming Distance of Patterns on Storage Capacity of Hopfield Network - IEEE Conference Publication." Design and Implementation of Autonomous Vehicle Valet Parking System - IEEE Conference Publication, Neural Information Processing, 2002, ieeexplore.ieee.org/document/1202172/citations.

[5] Miner, Don and Pickett, Marc. "Understanding the Brain's Emergent Properties." 2009

[6] Breakspear, Michael, et al. "Computational Models of the Brain: From Structure to Function." Egyptian Journal of Medical Human Genetics, NeuroImage, 26 May 2010, doi.org/10.1016/j.neuroimage.2010.05.061