CORTICAL CIRCUITRY

Printed in the United States of America
First edition, 2017

Figures are by the author, unless otherwise attributed.

www.corticalcircuitry.com

ISBN 978-0-9994422-1-0

CONTENTS

1. Evolution of Information Processing Capabilities 7
2. Introduction to the Neocortex 11
3. Classic View of Neurons 24
4. New Insights into Neurons 35
5. Functions of Cortical Layers 48
6. Cortex in Action 62
7. Using Insights into the Neocortex in Algorithms 71
8. Hardware for Next-Gen Neural Networks 86
9. The Limits of Our Understanding 92
10. Afterword 96
11. References 99

INTRODUCTION

I wrote this book because I have some new ideas about how the neocortex works. Since this remarkable organ is relevant to all human beings and not just neuroscientists, the first three chapters provide a short introduction to biological information processing and the neocortex – they are designed to provide enough background information for the rest of the book to be readily understood, regardless of your prior knowledge. If you have studied neuroscience, this book is about dendritic computation and the functions of neurons in different cortical layers, and you might want to skip straight ahead to chapter 4.

Enjoy.

1

EVOLUTION OF INFORMATION PROCESSING CAPABILITIES

Every biological organism can be thought of as an information processing mechanism. It is worth looking at the history of information processing capabilities and their evolution to see how the modern mammalian brain differs from other information processing systems.

While little is known about the earliest living organisms, we can be certain that they did not have sensors to examine the outside world. They existed in energy rich environments and gained energy simply by colliding with things from which they could extract it.

Single cell organisms developed sensors somewhere in the evolutionary process. E. coli have the ability to detect whether the environment they are in contains either of the sources of energy they most frequently use: glucose and lactose. When glucose, their preferred energy source, is available, E. coli produce enzymes to metabolize it. If glucose is absent but lactose is present, they produce enzymes that metabolize lactose instead[1]. In terms of information processing, such behavior can be described with a very simple algorithm of the sort that humans have been writing for decades:

```
if (glucose_is_present):
        produce_enzymes_to_metabolize_glucose ()
else:
        if (lactose_is_present):
                produce_enzymes_to_metabolize_lactose()
```

Note that no memory storage is required for the execution of this algorithm, nor are the E. coli sensors particularly advanced. They cannot analyze the entire spectrum of chemical compounds that E. coli encounter and create a list; they can only detect the particular compounds of glucose and lactose (and a few others). This level of information processing requires no nervous system, which a single-cell organism could not possess anyway.

Insects do have nervous systems with up to a million neurons each. They allow insects to perform information-processing tasks of considerable complexity. Insect eyes can analyze any visual information they encounter and recognize familiar shapes. They also utilize advanced sensors. Jumping spiders, for example, have eight eyes. Six of them are like fisheye cameras that survey the environment for moving objects. When such objects are detected, the spider turns its head to image the objects with its two high-resolution, tunnel-like eyes. These confirm if the objects have any nutritional or reproductive value to the spider and elicit fairly complicated behaviors in response, such as jumping at prey or courtship dances[2]. Moreover, insects can navigate real-world environments very well – sand wasps will hunt as far as a mile away from their nests and can reliably find their way back to a tiny hole in the ground.

However, there are clear limits to what insects can do. While they are able to recognize shapes, the shapes they recognize are genetically preprogrammed, as are the insects' behavioral responses to the shapes. Sand wasps would be unable to learn that they can push a lever to receive a reward of food. Insect capabilities are therefore similar in their level of complexity to those of a modern self-driving car: they can examine its surroundings, recognize dangers and opportunities, and select from a set of behaviors those that are the most likely to lead to success. Navigational capabilities do require memory, but in the case of insects, the kinds of memories they can have are fixed. They can remember maps and objects but would not be able to learn symbolic representations. Note that the acquisition of new skills for insects and other organisms of similar complexity happens not within the lifetime of a single organism, but as an evolutionary process.

Mammals[i] have the most advanced brains and, thus, the best information processing capabilities of all animals on Earth. Mammals are able to acquire new skills within their lifetimes, learn symbolic representations such as languages, and, perhaps most importantly, teach and learn new skills through communication. While it is up for debate whether all mammals can do such things, it is largely understood that these capabilities heavily depend on the part of the brain called the neocortex (the cortex for short). It's the part of your brain that is reading this book, as well as the part that can learn math and remember faces[ii].

In the evolutionary process, the neocortex did not replace the reptile brain. Instead, it evolved on top of it – literally, as the

---

[i] The goal of this chapter is not to delineate precisely the information processing capabilities of various living organisms, but to showcase the general growth of them.

[ii] Other parts of the brain participate in these processes, but the cortex does the heavy lifting.

neocortex is physically located atop the brain, just beneath the skull. Because the neocortex was the last part of the brain to evolve, it is heavily interconnected with all the other, older parts. These other parts of the brain have specific purposes. Some track information about hunger. Some help the animal maintain balance. Some store basic skills that animals are born with, like walking.

The cortex introduces a layer of intelligence on top of the range of abilities that reptiles possess. It adds significant data analysis and memory capabilities for mammals, allowing them to utilize their bodies and other parts of the brain more efficiently. Exactly how it does this has so far been a mystery. We understand perfectly how other, older parts of the brain respond to sensory information. We can predict how ganglion cells that are the output cells of the retina react to light, as they react the same way every time we show the eye the same picture. When the same visual information is transmitted to the cortex, however, even the first part of the visual cortex to receive it responds differently across trials[3]. A simple way to think about it is that thoughts, memories, and the state of mind all come into play at that point. It is the pinnacle of information processing on Earth – let's jump right in to how it works.

2

INTRODUCTION TO THE NEOCORTEX

The cortex is a flat, thin structure on top of mammalian brains. It is about two millimeters thick. If we were to straighten it out, a human cortex would look like a pizza or a dinner napkin. A rat cortex would be the size of a quarter. While crammed into a skull, a cortex is highly convoluted – the image that you have in your head of the brain is actually an image of the neocortex. The other parts of the brain are concealed underneath it.

All mammalian cortices are similar. They all have the same types of neurons and they all have very similar structures (more about both of these later). The reason that humans are smarter than rats is largely a function of the size of our cortex, not because we have any different special cell types[iii].

The cortex is difficult to study because it is huge and largely uniform. The human neocortex contains ~20 billion neurons and offers no immediately obvious way to divide it into parts (which is the first thing scientists try to do when faced with a new object), apart from, perhaps, the division into two hemispheres. That being said, scientists have made incredible advances over the past century,

---

[iii] Humans may have cell types not present in, say, mice, but the vast majority of neuron types are the same in all mammals.

and we now know quite a few things about what the cortex does and how it does it.

It is a computer scientist's immediate instinct to compare the cortex to a CPU of a computer. It is natural to do so – the cortex performs functions, like math, that we usually associate with processing. However, humans have not yet built anything that could be accurately compared to the cortex[iv]. Moreover, we do know that the cortex is, in essence, not a processing system like a CPU, but a memory system. As this is not immediately obvious, some examples are in order.

Imagine a batter on a baseball field. Any second now, another person is going to throw a ball at him. That ball is going to fly at ~100 miles per hour for 0.45-0.6 seconds. In that time, the batter has to decide whether to hit the ball and how to swing the bat if he chooses to do so. It is impossible for a human brain to compute the differential equations that will give the trajectory of the ball and the muscle movements required to hit it in half a second. So, the cortex does not do that. What it does instead is bring up the memories of all the baseballs flying at you (and the preceding movements of the person throwing the ball) and selects a behavioral response that, in the batter's experience, has been the most likely to result in a successful response to the sensory input. For professional baseball players, such memories occupy a relatively high proportion of their cortex, which is why they are so good at what they do. Generally, people become better at doing things with practice. This is a function of a memory system, but not of a processing, computing system.

[iv] As evidenced by a lack of all-out warfare between humans and this object, and following enslavement of the former by the latter. Note to self: joke funny, but unfit for publication. Delete before disseminating.

But what about math? Surely doing math is a computing process, not a memory one! Not necessarily. Try multiplying 13 by 26 in your head. There are a few ways you could do this, but let's assume that you pick the most basic one. First, you multiply ten by twenty-six. You *remember* that multiplying something by ten requires only adding a zero at the end. You hold the number 260 in your *memory*. Then you multiply twenty by three. You *remember* that the answer is sixty. You hold 60 in your *memory* also. You then multiply 3 by 6. You have all single-digit multiplications *memorized*, so you know that the answer is 18. Most of the operations involved in a computational process such as math are, in human brains, memory operations.

The type of memory that the neocortex uses for most purposes is sequence memory. Information comes into animal brains in a continuous stream, rather than in discrete chunks. While awake, mammals generally do not stop seeing or hearing. Moreover, information always comes into your brain sequentially, and it is frequently important which parts come after others.

To illustrate this with an example, try recalling the third line of your favorite song. Most likely you will have to go mentally through the preceding two lines first. Your brain has always heard the third line following the first two. Sure, you remember the lyrics of the whole song – but you remember them as a sequence. Another example would be trying to recall the location of a certain lost object, such as your car keys. To try to get to that location, you recall the prior sequence of behaviors that involved you and, potentially, that could have included the object as well.

The cortex (and another part of the brain called the hippocampus, which we'll get to later) always record the sequence of events that happen to you. Not all of that memory is stored permanently; some is either purposefully deleted or gradually degrades over time[4,5].

By contrast, computers don't generally store sequences of events. Most computer memory storage systems involve discrete chunks of data, such as images or numbers. If we want to build computers that are more like humans in their capabilities, we will need to address the problem of storing memories in the form of sequences.

The most important reason for storing sequences in the cortex is that they allow mammals to make predictions. When we hear a certain set of piano chords and the phrase "Imagine there's no heaven," our brain immediately predicts that the next phrase we hear will be "It's easy if you try." Moreover, if the next phrase does not match our expectation, our attention immediately kicks into high gear. Recognizing sequences in the sensory stream, making predictions about the next steps of these sequences, and comparing predictions with the actual input that follows are fundamental activities of the neocortex.

While visual and auditory information may seem very different to you, your cortex approaches both of these data types in the same way: recognize sequence, make prediction, and verify prediction. One of the biggest revelations about the neocortex was made by Vernon Mountcastle when he suggested that the cortex has a common algorithm[6] for dealing with different types of information. He based that suggestion on the fact that the parts of the cortex that process auditory and visual information look exactly the same. In fact, ALL of the parts of the cortex look very similar[v].

Let's dwell on this, for it is very important. We have now experimentally verified that if you take a baby ferret and rewire its brain, connecting its eyes to the part of the cortex that normally handles hearing, the ferret will develop functioning visual pathways in that part of the cortex[7]. It's the same way with humans. If you take a human who has lost their sight and connect a camera to their tongue with a chip in the middle that encodes visual information into electric pulses shocking the tongue, the person will be able to functionally "see" the world, albeit at a lower resolution than people regularly do[8]. There is nothing fundamentally different about different parts of the cortex, only the sources of information they are connected to.

Francis Crick, one of the discoverers of the shape of the DNA molecule, turned to neuroscience later in life and achieved quite a bit of prominence in the field. Once, he was delivering a lecture to people unversed in biological sciences. A woman asked him – "Professor Crick, what is the most important discovery in neuroscience?" He thought for a while and then replied, "The most important discovery I know of is that the brain is plastic". The woman promptly fainted. While it would be more amazing if the brain was literally made of plastic, the ability of the cortex to rearrange itself to adapt to new circumstances is nothing short of amazing. That is why mammals are able to learn new skills, while animals without a neocortex are mostly limited by the set of behaviors that is genetically preprogrammed in them[vi].

That being said, there is a fair bit of structure in the cortex

---

[v] There are known differences, but they are easily eclipsed by the similarities.

[vi] Birds can also learn new skills, but do not have a neocortex. Birds evolved a different structure on top of the reptilian brain than mammals did. Theirs is also plastic.

that is maintained across humans[vii]. Barring extraordinary circumstances, different people will process speech in the same parts of the cortex. In other parts of the cortex, they will plan reaching motions. And yet all of these parts have the same structure – and can perform functions they do not perform naturally if rewired! That means that the question of how the different parts of the cortex are connected to each other is very important.

The answer? The different parts are connected hierarchically[9]. A human cortex can be thought of as a hierarchy of cortical regions that all operate using the same principles and are connected in an ordered fashion, wherein some of the regions are more high-level than others[viii]. Here is a high-level scheme of the cortical hierarchy:

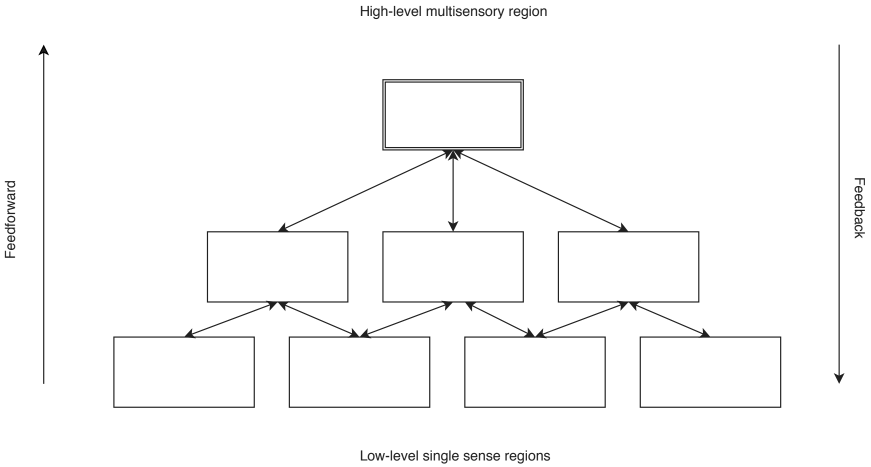


Figure 1 – Hierarchy of the Neocortex

[vii] The structure of the cortex is maintained in other mammals, but each specie has their own structure. The structure of the human cortex is, expectedly, more similar to that of apes than that of mice.

[viii] This idea is the basis of the "hottest" machine learning technology – deep learning.

There are a few interesting things about this.

First of all, the lower regions in the hierarchy are single-sense regions. They find small patterns in information, like edges and large blobs of color in the visual field. Higher regions find patterns in the edges and blobs of color – squares, circles, etc. Regions higher yet might find patterns in squares and circles to identify objects as tables or cats. Most objects that humans encounter affect multiple different senses. For example, chairs affect touch and vision, bananas affect touch, vision, smell and taste, pianos affect vision, touch and hearing, etc. Hence, higher-level concepts representing real-world objects are generally found in multi-sensory regions.

The second important aspect of the hierarchy is that information flows both ways – up and down it. Generally, sensory information flows up the hierarchy, while predictions and goals flow down the hierarchy[10]. If a cortical region identifies a cat, it sends that identification upwards. Down comes information about this particular cat – that it's named Nokizaru, and that it generally meows when you don't feed it for a while, which is the case currently. Consequently, this region might send a prediction of meowing to the auditory region beneath it in the hierarchy. Another prediction might be that the cat bites when it's very hungry.

Finally, a crucial feature is that cortical regions have both sensory and motor components. Visual regions mostly process information that comes from the retina, but they also control eye movement and receive information about the position of the eyes. Similarly, regions that receive information from mouse whiskers control whisker movement[11], and so on. It makes sense – to be able to predict what you are going to see the next second, you need to know where your eyes are going to move.

It was mentioned earlier that the cortex has the same structure in all of the different regions. That structure is a laminar one – each part of the cortex consists of layers, about six of them[ix]. Generally, they are referred to as layers 1 through 6. They are ordered by proximity to the skull – layer 1 is the closest to the skull, while layer 6 is the deepest (the motor component mentioned above is layer 5). There are variations on how thick different layers are in different parts of the cortex, but the basic structure is remarkably similar. Moreover, each layer has cell types specific to it, and that specificity is maintained across the neocortex.

The higher regions of the cortex (higher in terms of hierarchy of information flow, not physical location) deal with higher level concepts, and so their layer 5 is slightly different from that of the sensory regions. Their layer 5 may not send motor commands directly to the parts of the nervous system that control muscles. Instead, it can send "high level" motor commands – like "go to the bath" – to the lower levels of the cortex, which translate those into sequences of smaller motor behaviors.

In a sense, the cortex is always building a sensory-motor model of the world. It makes predictions about what sensory information will come in if you pursue different courses of action, and it chooses the course with the most desirable outcome.

To give you an example of how that works, my lower visual regions currently see looming green circles. My higher-level visual regions recognize them as cat eyes. My high-level cortical regions start paying attention to my tactile regions that verify that I still have

---

[ix] Six is the classical number, although many scientists divide the cortex into layers differently.

bite marks from the last time I tried to postpone feeding the cat. The high level cortical regions make a prediction that unless I feed my cat right now, I will be bitten again shortly. Hence, the high level cortical regions decide that I should feed the cat and send the command down the hierarchy to go feed her. My lower level regions recall the sequence of movements that end up in a fed cat and, consequently, lack of bite marks, and begin to execute them.

Note that every behavior is a result of some motivation – in this case, ~~fear of being bitten~~ the unconditional love I have for my cat. Most such motivations are connected to emotions, such as fear, hunger, or love, which reside in subcortical structures (the older parts of the brain that are not the cortex). The cortex is tightly interconnected with those structures, and emotions have important effects on cortical functioning. They help select goals, but they also highlight certain memories as important. Think of the last emotionally intense situation you encountered – you are likely to remember it vividly.

Going back to goals for a second – this book is going to be talking a lot about goals. While that word has a lot of meanings in the human world, this book will be largely talking about immediate animal goals – hunger, cold, fear, etc. Interestingly enough, all but one of the different goals that humans have come from emotions and, thus, from subcortical structures. The only goal that is native to the cortex is curiosity. For the cortex to be useful, it has to make correct predictions about the world. Hence, it must have a motivation for learning to make correct predictions. That motivation has to force the cortex to build better models of the world and resolve inconsistencies through experiments – it's curiosity! When we build true artificial intelligence (AI, or AGI as people like to call it today), curiosity will be its only inevitable motivation. Such an entity will have no need for the evolutionary baggage of hunger or fear of snakes (the common sci-fi premise of “artificial humans” is very unlikely), but intelligence simply does not work without curiosity. What other motivations such a being might have will be left up to its creators.

One of the main features of information processing systems is how they store information in memory. Computers store information in binary code – 101010 would, for example, denote the number 42. Note that this notation only works well if each bit is reliable. If we change the leftmost bit to 0, the meaning of the notation changes to ten. The cortex utilizes a rather different form of memory storage. A useful abstraction to help people understand it is called an SDR – sparse distributed representation. The word “sparse” means that a small proportion of bits are active a (i.e. are 1s and not 0s; within the context of the cortex, a neuron firing represents a 1) at the same time – maybe 2%. By comparison, regular binary notations are dense – roughly 50% of bits are ones in them. The word “distributed” means that the activation of many bits is required to represent something.

An SDR might look like this:

0100010000010000000100000100001000000010000000000

Note that an SDR does not assign more meaning to the bits further to the left. In fact, the same SDR can look like this:

01000100000100
00000100000100
00100000001000
0000000

The position of each bit does not matter. No bit is more “important” than another. However, each bit has some semantic meaning associated with it. Let's say that this SDR is representing an animal. Some of the bits might have the following meanings:

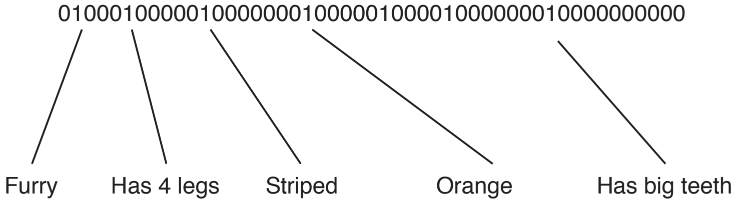


Figure 2 – Tiger Encoded in an SDR

All of these bits denote different aspects of a tiger. Note that if we change a single active bit to 0, the meaning of the representation as a whole will not change that much (even less so given that in brains these representations consist of thousands of bits, as a single neuron might have anywhere between 1000-30000 synapses). Such an arrangement is useful in identifying objects with incomplete information. For example, if you can't see the tiger's legs, and you can't be sure that there are four of them, you can still identify it as a tiger because the other evidence is overwhelming.

Naturally, within the cortex, the semantic meaning of each neuron is dependent on where the neurons are located and what they are connected to. Some neurons in the visual cortex have meanings like edges or large blobs of color. Some neurons in the auditory cortex might correspond to various musical notes. In the higher levels of the cortical hierarchy, individual neurons have more abstract meanings, like "noun", "eloquent" or "Jefferson Airplane". Note that when I say that neurons have meanings – it's not that individual neurons precisely encode concepts, but rather that the "Jefferson Airplane" neuron fires when the brain thinks of something related to that name.

Of course, real neurons aren't "bits". The concept of SDRs is simply useful to think about how complex concepts are encoded with large assemblies of neurons. The concepts of sparsity and information being encoded by large groups of neurons are both present in the cortex[x].

If this is a little bit confusing, consider how the neocortex encodes the memory of an elephant. The visual cortex has neurons that encode blobs of gray and the shape of an elephant. The auditory cortex has neurons that encode trumpeting. Somewhere high up in the hierarchy there are neurons that encode abstract concepts like "mammal" and "herbivore", as well as the word "elephant". These neurons together encode an elephant.

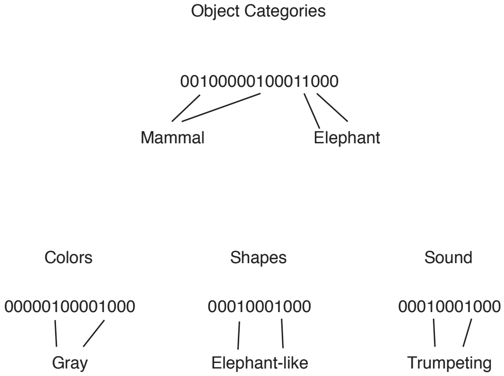


Figure 3 – Elephant Encoded in a Hierarchy of SDRs

It is tempting to think that only the neurons that are labelled "Elephant" encode an elephant, but that view is wrong. An elephant

---

[x] SDRs have many interesting properties; to learn more about them, check out Numenta's work.

is encoded by all of the neurons active above. Moreover, neurons that are labelled “Elephant” can only be activated if the lower-level neurons that describe features of an elephant are active. That being said, the example above only deals with sensory information and object recognition, which largely stems from it. The cortex has neurons that encode other things – movement, goals, predictions and so on.

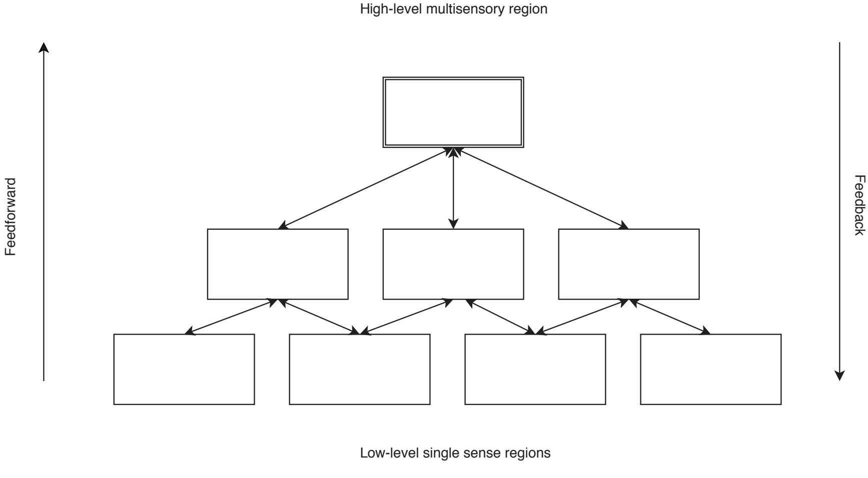
The question people most frequently ask me about the brain is whether it’s true that we only use 10% of our brain at any given time. The SDR concept provides an explanation – sure, we only have a certain percentage of neurons active at the same time. But if we had ALL neurons active at the same time – all meaning would be lost. We would see all of the colors and hear all of the sounds at the same time and would have an epileptic fit. But is it possible to improve our intelligence by having, say, 30% more neurons active than we normally do? Yes and no. I imagine that it would feel a lot like being on LSD – you would see more patterns, but there is no guarantee that the patterns you see exist in the world outside your head.

With that out of the way, we can proceed with our examination of how the cortex actually works.

3

CLASSIC VIEW OF NEURONS

The picture of the cortex painted so far is tidy. There are regions that have functions, they are arranged in a hierarchical order, and information flows neatly between them up and down the hierarchy. This is a neat diagram:



Unfortunately, that's not what the cortex looks like. The cortex looks like this:

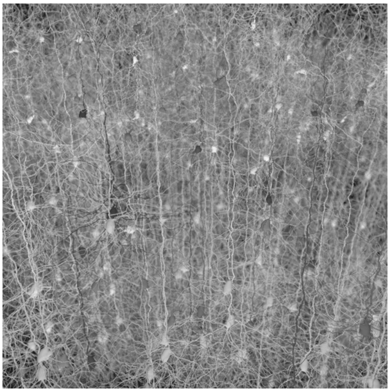
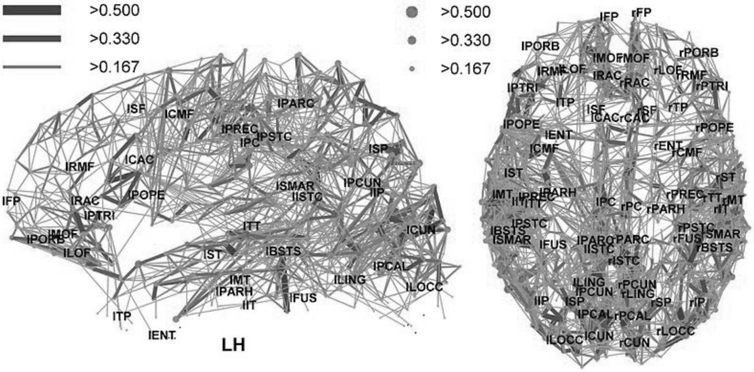


Figure 4 – Image of the Cortex

(Forest of synthetic pyramidal dendrites grown using Cajal's laws of neuronal branching by Hermann Cuntz CC-BY-2.5 https://creativecommons.org/licenses/by/2.5/ Original - http://journals.plos.org/ploscompbiol/article?id=10.1371/image.pcbi.v06.i08*)*

And when we try to create a map of how the different regions of the cortex are connected with each other and with other parts of the brain, it looks like this:



(Dorsal and lateral views of the connectivity backbone of human brain. Labels indicating anatomical subregions are placed at their respective centers of mass. Nodes (individual ROIs) are coded according to strength and edges are coded according to connection weight by Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns CC-BY-3.0 https://creativecommons.org/licenses/by/3.0/deed.en Original - https://commons.wikimedia.org/wiki/File:Network_representation_of_brain_connectivity.JPG*)*

There are many different maps like this. There is no agreement on how to divide the brain into parts, or which of those parts are connected (what connection weight means “connected”?). Moreover, cortical regions are not the discrete entities computer scientists would like them to be. It is just a lot of neurons[xi] intertwined together like shown above.

There are different types of neurons, although there is no agreement on how many types there are, or even on how to divide neurons into types. People have tried to classify neurons by the genes they express, the cortical layer their cell body is in, and even their shape. One distinction is obvious and discrete, though, which is that some neurons are excitatory, while others are inhibitory. When excitatory neurons fire, they make other neurons fire. When inhibitory neurons fire, they prevent other neurons from firing. The biological difference between these two types of neurons is the type of neurotransmitter, or signaling molecule, that they use. Excitatory neurons primarily use a neurotransmitter called glutamate, while inhibitory neurons use other neurotransmitters, like the one called GABA[xii]. About 80% of neurons in the cortex are excitatory, while 20% are inhibitory. Moreover, the inhibitory neurons are a lot more diverse than excitatory neurons, whereas most excitatory neurons look very much like each other[xiii]. It follows that to understand the cortex, one would have to understand what excitatory neurons do, given their prevalence.

[xi] There are also glia cells, but information processing doesn’t seem to be their main function. Glia cells provide nutrition and other auxiliary services to neurons.

[xii] There are other excitatory and inhibitory neurotransmitters, but these are by far the most prevalent.

[xiii] The exception is stellate cells.

This is what an excitatory neuron in the cortex looks like[13]:

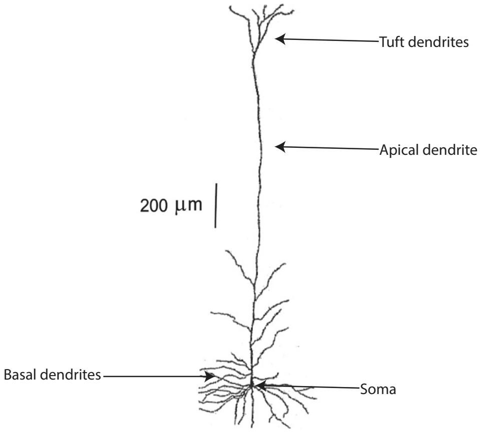


Figure 6 – Excitatory Neuron in the Cortex

Adapted from Kasevich RS, LaBerge D (2011) Theory of Electric Resonance in the Neocortical Apical Dendrite. PLoS ONE 6(8): e23412. https://doi.org/10.1371/journal.pone.0023412*)*

In real brains, of course, neurons are three-dimensional, so this neuron would look a little bit like a pyramid; hence they are frequently called "pyramidal neurons." Each pyramidal neuron[xiv] is unique, but they all share a few characteristics. First, they have a cell body called a soma. From the cell body, dendrites protrude. They look like tree branches - dendrite means "treelike" in Greek. On each dendrite, there are many synapses – entities that connect this neuron to other neurons. Synapses can be chemical or electric. Chemical synapses send molecules from one neuron to another. These molecules, known as neurotransmitters, increase the charge of the recipient neuron. Electric synapses work as any electric connection does. This book considers only chemical synapses, which is a common ap-

---

[xiv] Unless otherwise noted, hereafter the word "neuron" is used to talk about pyramidal cortical neurons, all of which are excitatory.

proach in neuroscience (though not necessarily a correct one)[xv]. The long thing that looks like a stem protruding upwards from the cell body is known as the apical dendrite. The dendrites that branch out from the cell body are called basal dendrites. The dendrites at the top of the apical dendrite are called tuft dendrites.

The apical dendrite can end in different layers, depending on which layer the cell body is located in. This relationship is quite strict and persists across the cortex. Layer 6 neurons send their apical dendrites to Layer 4; Layer 5 neurons send their apical dendrites to Layer 1, etc. The apical dendrites of excitatory cells always go upwards (towards the lower-numbered layers, or the skull) from the cell body. A neuron's tuft dendrites spread out around the end of the apical dendrite in the layer that it ends in.

From the bottom of the neuron protrudes the axon, the neuron's output link (not pictured). A neuron's axon is what connects to other neurons' dendrites with synapses. Technically, each synapse is a combination of an output (presynaptic) part on an axon and an input (postsynaptic) part on a dendrite. Unlike dendrites, which in cortical excitatory neurons have a similar pattern (a bunch of them near the soma and some branching out at the top of the apical dendrite), axons even of neighboring neurons can have wildly varying shapes[14].

Axons frequently extend over long distances. Sometimes they go to higher or lower levels of the cortical hierarchy. Other times they go to the same part of the cortex in the opposing hemisphere (called the contralateral cortex). They can extend to subcor-

---

[xv] We don't know what electrical synapses do, so at this point adding them to theories increases complexity without increasing the amount of information communicated. There is a slightly more detailed discussion of electrical synapses in chapter 8.

tical structures, too. Like the apical dendrite, axonal targets are also specific by layer. For example, axons from layer 4 neurons generally go to layer 3. Axons from neurons located in layers 2/3 generally go to layer 5, etc. However, axons from the same layer in the same part of the cortex can have vastly different destinations. Chapter 5 explains in detail how the layers and axonal targets work together.

Another trait that all excitatory neurons have in common is that they spike (also known as firing, also known as action potential, also known as nerve impulse). When multiple excitatory neurons that are connected to a given neuron by synapses fire, they make the cell body of that neuron electrically charged. The more connected neurons fire together, the higher is the charge of the cell body. If the cell body's charge passes a certain threshold, the neuron itself spikes. That means that a signal spreads down the axon rapidly and activates output synapses, thus sending signals to other neurons. In a way a, neuron can be thought of as a leaky bucket balancing on an edge. Water keeps coming in – if a lot of it comes at once, the bucket turns over and spills into other buckets. But some water is naturally dripping out without overturning the bucket. So, if a neuron receives inputs separated by long periods of time, it won't fire – the inputs have to be somewhat concurrent.

Firing is the prevalent method of information transmission in the brain. Moreover, there is broad agreement that the basic function of a pyramidal neuron is pattern recognition. A neuron fires when it recognizes a pattern in the firing of other neurons around it. This explains some of the universality of the cortex – no neuron "cares" whether it's recognizing a pattern in visual or auditory data. It just fires when many neurons connected to it fire at the same time.

That allows for the aforementioned common cortical algorithm. Recall the ferret who had his brain rewired. The neurons in the part of its cortex that used to be auditory do exactly the same thing now that they are connected to its eyes – they fire when they recognize patterns in the firing of neurons connected to them.

A good question to ask at this point is how neurons pick which other neurons they should connect to. To some extent, this is determined genetically. As mentioned, there are clear rules about where certain types of neurons send their apical dendrites and axons. These rules are based on which layer the neuron is located in, as well as where in the cortex it lives. Moreover, neurons that share the same progenitor cell are likely to be connected - that allows nearby neurons to create local microstructures.

However, the most interesting part of neuronal connectivity is what scientists call "experience-dependent plasticity." It means that as organisms encounter new information in the world, neurons form new synapses. There is a general method that most neuroscientists accept to be true: a neuron forms synapses with neurons that fire at the same time as it does. This rule, frequently described as "fire together, wire together," is known as Hebbian learning. If we pause to think about the cortex as a whole, this rule explains a lot of our thinking processes. Remembering that two things happened at the same time is something humans are very good at. For example, when you commit a certain scene that is important to you to memory – a marriage proposal or a college graduation—you are likely to remember details that are not strictly relevant to the experience, but that happened at the same time, like what you had for lunch that day.

Emotional saliency is often the factor that makes you commit scenes to memory. This is responsible for a few "bugs" in human thinking. One example is the belief that when you think of somebody and then they instantly call you, some higher powers are at work. The reality is far more boring; you think of dozens of people per day, but you don't remember every instance of thinking of somebody. However, when it coincides with them reaching out to you, you react emotionally and store this moment in your memory. Thus, in your long-term memory, there are a disproportionately high number of such instances.

A phrase that is frequently used to describe the operation of the neocortex is "associative memory." It is enabled by the Hebbian learning rule of "fire together, wire together." Recall that the neocortex evolved on top of other parts of the brain – ones that are responsible for hunger, fear, and happiness, as well as complex sensors like eyes and ears. The cortex learns relationships between the various sensory inputs and the states of subcortical structures. It remembers that chocolate is delicious, that barbed wire hurts, and that certain music makes you sad. All of these mechanics are enabled by Hebbian learning.

The action potential mechanic was a terribly exciting discovery. In 1943, McCulloch and Pitts came up with a theoretical model of a neuron that is still the basis (with some changes) of most machine learning technologies today[15]. A general form of it looks like this:

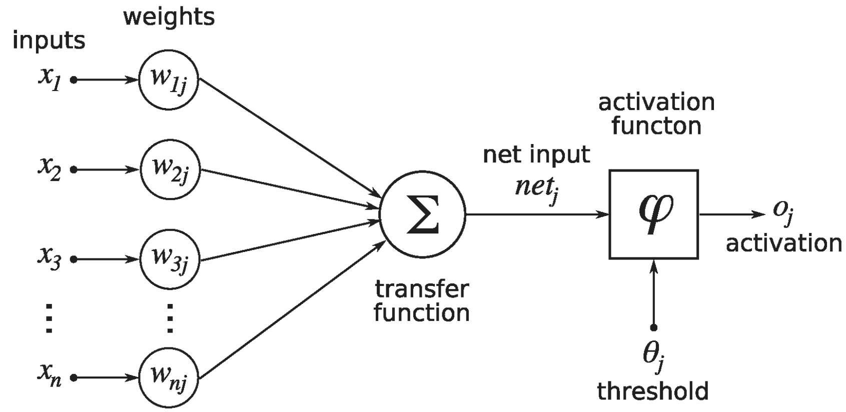


Figure 7 – Model of an Artificial Neuron

(Artificial Model Neuron by Chrislb CC-BY-3.0 https://creativecommons.org/licenses/by-sa/3.0/ Original - https://commons.wikimedia.org/wiki/File:ArtificialNeuronModel_english.png *)*

This model is remarkably simple. Every neuron has a bunch of inputs. Each input has some weight ascribed to it[xvi]. At every time step, all of the inputs that are active (which is meant to represent neurons firing) sum up their weights. If the sum exceeds some activation threshold, the neuron fires. If it doesn't, the neuron doesn't fire. Hebbian learning is used - if two linked neurons fire together, the synaptic weight is increased.

Sixty years later, this is still essentially the model that most AI projects use[xvii]. One important computational feature of this model is that as the number of neurons increases, the number of links between them increases exponentially. It is not computationally feasible (nor will it be anytime soon) to create models with bil-

---

[xvi] Weights can be negative, allowing modeling of inhibitory neurons.

[xvii] Bruno Olshausen does a great job at explaining the evolution of computer science models of neurons here: https://youtu.be/WSzDPfCpcYM

lions of artificial neurons all of which can connect to one another. Programmers deal with this by dividing the neurons into "layers." A layer might have thousands of artificial neurons, and the layers are arranged hierarchically. The terminology here is a bit confusing. Deep learning layers are not modelled after layers of the neocortex, but the same word is used to describe different concepts. In deep learning, "layer" simply means a discrete chunk of neurons, while in the neocortex each of the six layers has unique features.

Neuroscientists, however, did not stay content with the McCulloch-Pitts model for long. It lacks certain features of neurons. One such feature is that neurons, unlike computers, don't have "timesteps." They operate in continuous time. Hence, neuroscientists came up with models that incorporate time. Hodgkin and Huxley won a Nobel Prize in 1963 for coming up with an action potential model in continuous time[16]. A more recent version of it, which is widely used today, is called the "leaky integrate-and fire" model[16]. Its premise is simple: inputs to a neuron boost its voltage, which otherwise slowly dissipates over time. However, if a threshold is reached at a given point in time, the neuron fires.

This model, also, is too simplistic when compared to actual neurons. Even the most casual observer of neurons immediately notices that neurons have branches, which none of these models have. The adherents of these models respond that branches are simply a biological necessity and possess no computational capabilities. Hence, branches can be described simply as cables (the theory is called "passive cable theory"). The further away the synapse is from the soma, the further the electrical impulse has to travel. Since membranes have some resistance, the further away a synapse is from the soma, the more the input has to be “discounted” for the action potential calculation. Therefore, the theory states, the branches can be abstracted away, as the concept of “synaptic weights” can capture the computational meaning of the location of each synapse.

This view is also wrong. Recent discoveries have shown that dendritic branches do in fact possess computational capabilities, which requires of us a new model of how neurons process information.

4

NEW INSIGHTS INTO NEURONS

John Von Neumann had one of the most annoying qualities a human can have: he was generally right about things. As he lay dying in 1957, he was writing a book about how the brain processes information – *The Computer and the Brain*. He had very little neural data. He had no powerful computers or the Internet. But, in his book, he wrote the following phrase:

"It may well be that certain nerve pulse combinations will stimulate a given neuron not simply by virtue of their number but also by virtue of the spatial relations of the synapse to which they arrive. That is, one may have to face situations in which there are, say, hundreds of synapses on a single nerve cell, and the combinations of stimulations on these that are effective (that generate a response pulse in the last-mentioned neuron) are characterized not only by their number but also by their coverage of certain special regions on that neuron (on its body or on its dendrite system), by the spatial relations of such regions to each other, and by even more complicated quantitative and geometrical relationships that might be relevant."[17]

He was completely right. And yet, it took decades to confirm his idea, which has been completely ignored in both neuroscience and AI for a long time. In 2000, a team lead by Jackie Schiller published a paper in *Nature* called “NMDA spikes in basal dendrites of cortical pyramidal neurons”[18]. They used a combination of new (at the time) experimental techniques to see what happens when neurons receive varying levels of input at a single point on a basal dendrite. The results were nothing short of astonishing[xviii]. It turns out that basal dendrites have their own spikes. Up to a point, as you increase the number of inputs on a specific point of a dendritic branch, the inputs are summed linearly, just as the passive cable theory would suggest. However, at some point, once a certain threshold is reached, there is a spike in the local voltage.

That spike in dendrites resembles a neuronal spike – inputs are summed until they reach a threshold, which triggers a spike. Additional inputs do not significantly increase the magnitude of the response beyond the spike amplitude.

Why is that astonishing? First of all, it directly disproves the passive cable theory. If dendrites were but passive cables, they would sum inputs linearly. It is even more curious that this spiking mechanism is almost exactly what McCulloch and Pitts, and the generations of neuroscientists after them, thought to be the function of the whole neuron. It turns out that a single dendritic branch of a biological neuron can perform a computation equivalent to that of an artificial neuron. That raises the question - what are entire neurons in the brain actually capable of? This chapter strives to answer that question, but first some more features of NMDA spikes need mentioning[xix].

[xviii] I hope you forgive me the use of this cliché. Many things about the brain have been called astonishing, fascinating and so on, but this particular paper was largely responsible for my writing this book.

[xix] For reference, NMDA means N-methyl-D-aspartate. That is a class of receptors

The number of synapses that need to be activated to trigger an NMDA spike varies between cell types, the proximity of the branch to the soma, etc. Most importantly, though, NMDA spike thresholds are not discrete – they stretch out in space and time. 10 inputs to the same point on a dendritic branch trigger an NMDA spike, but 20 inputs are required if they are distributed along the length of the dendrite[19]. Similarly, if the inputs are clustered in time, fewer are required to trigger an NMDA spike than if they are spaced apart in time – not unlike the leaky integrate and fire model conceived for a whole neuron.

Moreover, recent nearby NMDA spikes help trigger new ones. That means that if normally 10 synaptic inputs are needed to trigger an NMDA spike, a recent nearby NMDA spike can lower that threshold to 8 synaptic inputs. This feature, referred to as cooperativity[19], is extremely powerful. While a single NMDA spike is not enough to trigger an action potential, multiple NMDA spikes can. Thus, recent NMDA spikes helping to trigger new NMDA spikes on the same dendritic branch is a feature that helps clustered inputs trigger action potentials. It is important to note that cooperativity only works within a dendritic branch. An NMDA spike in a branch does not lower the threshold for NMDA spikes in a nearby but unconnected (except through the soma, of course) branch of the same neuron. That means that a dendrite, not a neuron, is a computational unit.

We have since learned a lot more about the NMDA spikes – that they are present in most excitatory neurons in the cortex, that they occur both in basal and tuft dendrites, etc. But the main conse-

---

in excitatory cortical neurons. "NMDA spikes" can simply be thought of as a name for dendritic spikes for the purpose of this book.

quence of NMDA spikes is enormous - dendrites recognize sequences!

In a paper called “Dendritic Discrimination of Temporal Input Sequences in Cortical Neurons,” Tiago Branco, Beverley A. Clark and Michael Hausser used the same experimental technique to show that if you activate the same synapses in a fixed period of time but *in a different order* you can drastically change the probability of a neuron firing based on those inputs[20].

The reason for this is NMDA spikes. When you spread out in time two inputs to synapses that are close to each other on a dendritic branch, you may dip below the threshold for an NMDA spike. Moreover, even if you get one NMDA spike but further inputs are on a different branch, cooperativity doesn’t come into play, so further NMDA spikes are less likely, and the neuron is less likely to fire.

The Branco/Clark/Hausser study showed neurons are more likely to fire when inputs arrive in a sequence approaching the soma, rather than when they arrive in a sequence moving further away from the soma. That makes sense – impulses arriving far away from the soma take more time to travel to the soma, so if impulses arrive on a dendritic branch in an order approaching the soma, they will arrive in a synchronized fashion, thus making peak voltage higher and increasing the probability of the neuron firing.

Another study found an interesting feature of NMDA spikes on different parts of dendritic branches. The threshold for initiating a spike increases 5-fold from the tips of dendritic branches to parts of them close to the soma. But the effect of the spike on the electrical charge of the soma increases 7-fold in the same direction[19]. In other word, spikes that are close to the soma are a lot harder to trigger, but have a stronger effect on the cell body. Moreover, temporal summation increases as you go along the dendrite away from the soma – inputs that are close to the soma have a stronger effect when they are synchronized, but inputs on the tips of dendrites can be summed up without loss over relatively long periods of time[21].

To illustrate how all of this works, here is a hypothetical example. “Johnny” is a neuron, whose job is to fire when you finish your drink (and there is such a neuron in your brain!) That neuron would integrate information coming in from many sensory modalities.

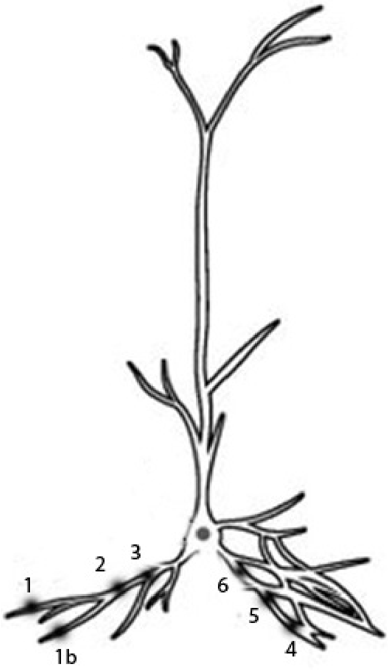


Figure 8 – Johnny The Neuron
(Adapted from OpenStax CC-BY 4.0 https://creativecommons.org/licenses/by/4.0/ Original -https://cnx.org/contents/FPtK1zmh@8.25:fEI3C8Ot@10/Preface*)*

At point 1 Johnny receives inputs from other neurons whose job is to fire when there is something cold in your right hand (a glass of iced water in this case). At point 2 it receives inputs from neurons that fire when your lips are feeling a straw. These inputs arrive constantly, as your lips keep continuously feeling the straw, and your hand is constantly feeling the glass. However, that input is not enough to make Johnny fire. But when it receives input at point 3 – from neurons firing that detect the sound you hear when you start sucking air through a straw— Johnny fires. It is important to note that if Johnny receives input at point 3 first, and then inputs at points 2 and 1, it won't fire. That's because if you hear the sound of air coming through a straw but don't hold a drink in your hand, it's likely that it's not you but somebody else who has finished theirs. Recall the thresholds from the previous paragraph – if there were no NMDA spikes at points 1 and 2, cooperativity would not come into play, and the threshold for an NMDA spike at point 3 would not be reached.

Neurons' branches provide a built-in mechanism for handling the variability of sequences in the real world. For example, if you are holding the drink in your left hand – Johnny will receive input at point 1b, which is similar to 1 (and, thus, located on an adjacent branch) but corresponds to your left hand.

Moreover, an entirely different basal branch can detect you finishing your drink in a different way, such as finishing a hot cup of tea that you hold in both hands. In that scenario, point 4 would correspond to input from your holding something hot in your hands, point 5 would correspond to input from your lips feeling the hot liquid, and point 6 would correspond to input that you receive from feeling the change in temperature in your mouth when tea is no longer coming in.

So, individual neurons can discriminate between sequences. That is a powerful and beautiful idea, mainly because it suggests that a single discrete element of the cortex, the excitatory neuron, performs the same function that the cortex does as a whole – sequence memory. There is a certain elegance in that. Moreover, we now know how memories are physically stored in the brain - it turns out they are stored in the three-dimensional branching patterns of neurons. Time in the neural storage of sequence memory is physically represented by the distance of a synapse on a dendritic branch from the soma. Different variants of similar sequences are represented by branching patterns. Each neuron's shape is a physical manifestation of a unique memory. Here is a generalized form of what a neuron's dendritic branching pattern represents:

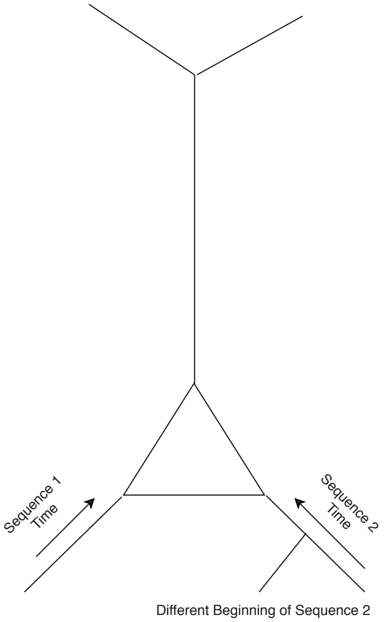


Figure 9 – Role of the Basal Dendrites in Excitatory Neurons

Note that sequence memory is the domain of basal dendrites – dendrites that are connected to the soma. Tuft dendrites, dendrites that are connected to the apical dendrite, do something interesting as well. Inputs to tuft dendrites can also produce NMDA spikes. A single NMDA spike and a little bit of additional input (or two or more NMDA spikes) sent to the tuft dendrites of a neuron result in a calcium spike[23]. Calcium spikes do not trigger action potentials on their own, but when combined with sufficient input to the basal part of the neuron, they cause the neuron to burst – that is, to spike 2-4 times at a frequency of around 200 Hz[xx,24].

		Calcium Spike	
		Yes	No
Basal Input	Yes	Neuron Bursts	Neuron Spikes
	No	No Output	No Output

Figure 10 – Neuron Responses to Different Combinations of Input

In simpler terms, if a neuron receives input only to the bottom part, it fires once. If it receives input only to the top part, it doesn't fire. If it receives input to both parts, it fires 2-4 times rapidly (called BAC firing or bursting). Matt Larkum, a prominent researcher of the neocortex, suggested in a 2013 opinion piece that the calcium spikes provide a mechanism for one of the most important cortical

---

[xx] Layer 2/3 neurons do not appear to produce bursts longer than 2 spikes.

functions[25]. Recall that information in the neocortex flows two ways: feedforward[xxi] and feedback (up and down the hierarchy). Calcium spikes and BAC firing enable individual neurons to associate feedforward and feedback information. In other words, predictions from higher cortical areas come in to the top of the neuron. Sensory data comes in to the bottom.

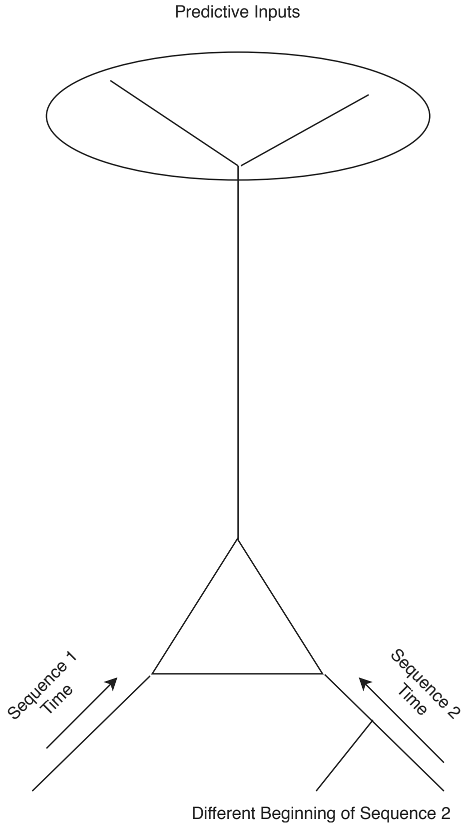


Figure 11 - Inputs to Different Parts of an Excitatory Neuron

If the sensory input matches what is predicted, the neuron bursts. If the inputs to the bottom part of the neuron aren't predicted,

---

[xxi] Frequently "feed-forward" is used in literature.

the neuron spikes just once. Conceptually, bursts are a lot more likely than individual action potentials to trigger NMDA spikes in the targets of the bursting neuron as you get a lot of quick input in succession to the same dendrites (and this has been shown experimentally)[26]. Thus, bursts are more likely to propagate through the cortex than individual action potentials. That makes sense in that predicted inputs are likely to be more reliable in information processing. To sum it up, individual neurons are able not only to learn sequences, but also to test if the sequences they detect have been predicted. Predicted inputs propagate through the cortex using bursting.

Since we know that the location of input on a neuron's dendritic tree matters, we are also interested in how neurons pick where exactly to connect to other neurons. Xu et al. published a study in 2009 that conclusively showed that exposure to new stimuli leads to rapid formation of dendritic spines[xxii,5]. That means that when exposed to new information, neurons create new synapses. Over time, new dendritic spines disappear, unless the stimuli that caused them to appear are repeated, or if the stimuli coincide with neuromodulation (emotional responses). Sometimes new dendritic spines cause a pair of neurons to connect in more than one location[27]. Pruning, however, leaves only the most efficient synapse in terms of timing, thus providing a mechanism by which people improve at activities that require precise timing with training.

A question now appears – how does a dendrite choose which synapses remain and which are "deleted"? Most likely, it is through a mechanism, which is similar to Hebbian learning as described earlier, but with a dendritic spike instead of an action potential being

---

[xxii] Spines can be thought of as physical manifestations of synapses for the purposes of this book.

required at the target neuron to form/strengthen a synapse. Dendritic spikes provide the location for the synapse. It has been shown that dendritic spikes participate in a form of strengthening synapses that does not require action potentials[28]. Moreover, a recent study by Kwon and Sabatini showed that nearby glutamate release (recall that glutamate is an excitatory neurotransmitter, which causes dendritic spikes) triggers spine formation on dendrites[29]. Basically, that means that if a dendritic branch "feels" a nearby axon firing at the same time as the branch is experiencing a dendritic spike, it will form a synapse with that axon.

Inhibitory neurons are far less studied than excitatory neurons, which makes sense given that they are fewer in number and far more diverse. It is not yet entirely clear how inhibitory neurons play into information processing on the dendritic level. However, we do know that inhibitory neurons can target (and effectively "turn off") specific dendritic branches[30] as well as calcium spike initiation zones[31]. This means that inhibitory neurons provide mechanisms for temporarily disregarding parts of certain sequences, as well as certain predictions. A general rule of thumb is that if an inhibitory neuron targets a dendrite far from the soma, it increases the threshold necessary for a dendritic spike. If an inhibitory neuron targets the soma, it reduces the amplitude of the spike[32].

One function of inhibitory neurons is that they support regular spiking[3]. The mechanisms of dendritic spikes and BAC firing work on the timescale of milliseconds. They are useful for rapid responses to novel stimuli. When you first see a zebra, BAC firing will be used to help you recognize that it is a zebra (more on how later). Brains, however, frequently have to work on longer timescales – most zebras (and other objects) persist in sensory domains for seconds or longer. For that, neurons utilize regular spiking. A neuron can spike continually at a certain frequency (30 Hz, for example) for a long time. If you continue looking at a zebra for a while, the neurons that encode the concept of zebra are going to continue spiking regularly until you decide to look at something else. Excitatory neurons on their own don't have any mechanisms to maintain regular spiking, but inhibitory neurons do[34]. Some inhibitory neurons fire regularly without inputs. That can help excitatory neurons maintain frequency by only allowing them to spike during certain time windows and inhibiting them during others. That way excitatory neurons get synchronized.

The main criticism of the idea of neurons being able to discriminate between precisely timed sequences is that synapses are stochastic. Indeed, synapses are notoriously unreliable. They may work, or they may not, in any given case. They can also spontaneously discharge. We don't know why; maybe they are just too small[35,xxiii]. Traditionally, scientists have thought that since synapses are unreliable, spikes are unreliable. If spikes are unreliable, neurons are imprecise. Hence, we can only look at firing rates – a neuron's output is its average firing rate over time.

An important distinction must be made here. The firing rate of neurons is frequently very important, like in the zebra example above. However, only taking the average firing rate of a neuron is not enough to understand what it is doing. Neurons burst when they identify new information and fire repeatedly when the information persists. Hence, we must look at both bursts of neurons and their

---

[xxiii] See Schrodinger's *What is Life* for a detailed examination of how size affects reliability in biology.

regular spiking.

What of the unreliable synapses, then? How are cortical neurons able to learn precisely timed sequences when their inputs are unreliable? It turns out that brains have evolved a way to deal with this fact of life that allows them to process information precisely.

This way is spikes – both NMDA spikes and full-neuron spikes. Spikes have the same peak voltage, even if you apply extra input after you reach the initiation threshold. In live brains, most dendritic branches and neurons are likely to be set up to receive more input than is the absolute minimum to trigger spikes. Thus, if some of the synapses don't work every time, the spikes still happen and the output of a neuron is largely unaffected by these failures. Naturally, that requires some degree of redundancy. There are neurons in the brain that do very similar things to each other. It allows for some neat things – neurons can further specialize when they need to, and having redundancies allows brains to handle neuronal death without losing much information—but mostly redundancies are important in dealing with stochastic synapses.

5

FUNCTIONS OF CORTICAL LAYERS

The power of the cortex lies in three things: the sheer number of neurons, the computational power of each individual neuron, and what is known as the "canonical structure" or the "common algorithm" – the way neurons in different layers are linked in a complex pattern.

There is a saying often attributed to Albert Einstein: "Everything should be made as simple as possible, but not simpler." Unfortunately, and as may be expected, in the case of the cortex, things can only be made so simple. There are six layers, each with different cell types, dozens of types of connections between them and a huge number of cortical areas, which may or may not be discrete. There is no soundbite that can explain how the cortex works. It is a common problem with biology – there are frequently huge numbers of elements in a system. Luckily, we can infer general principles, which can make any particular element easier to understand.

In the case of the cortex, any general principle has to explain its laminar structure and the connections between various layers. While different parts of the cortex can have certain layers that are thicker or thinner than usual, the one thing that remains the same everywhere in the cortex is the connectivity pattern between the layers. Layer 4 always sends[xxiv] outputs to layers 2/3, for example. Hence, we now have network diagrams of the layers of the cortex. They look like this[xxv]:

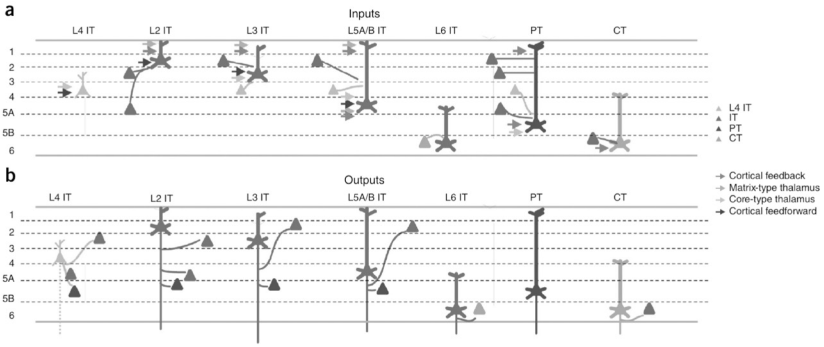


Figure 12 – Connectivity between Cortical Layers
Reprinted by permission from Macmillan Publishers Ltd: Nature Neuroscience[37], © 2015

While such diagrams are woefully incomplete without inhibitory cells, they do offer a glimpse of the general structure of the cortex. However, there is very little theory about how this structure translates into function. This chapter introduces some theories, as well as ways to test them. They are not necessarily right, but they are definitely "truthifiable"[xxvi].

Layer 4 is the input layer of the neocortex. In primary sensory regions it is the first cortical layer to receive input from the thalamus, which transmits sensory information from sensory organs to the cortex[38]. In higher regions of the hierarchy layer 4 receives

---

[xxiv] When I say a layer does something, I mean the excitatory neurons in that layer. Unless otherwise specified, "layer X neurons" refers to excitatory neurons in that layer.

[xxv] This is *a* diagram, not *the* diagram. Many of the distinctions made here are a product of human judgment.

[xxvi] See chapter 19 of *The Lightness of Being* by Frank Wilczek for a discussion of truthification.

feedforward input from lower regions[39]. Neurons in all layers connect to other neurons in the same layer; however layer 4 neurons do that a lot more than neurons in other layers[40]. In this way, they find patterns in broad slices of data. Consider, for example the device on which you are reading this book – whether digital or paper. Its top edge extends across a large portion of your visual field. Because of that, neurons that identify the left side of it and neurons that identify the right side of it can be pretty far apart. But because layer 4 neurons are so interconnected, they can recognize that it's the same edge.

It has long been known that thalamic axons, which carry sensory information to the cortex, branch out in layers 6, and, to an extent, 5, as they enter the cortex. Recent experiments show that even if layer 4 is disabled, information still flows into the cortex from the thalamus through these branches going to layers 5 and 6[41]. The reason for having two feedforward pathways is attention. Attention is a process of selecting useful information from the environment. It is by nature an inhibitory process – it works by filtering out irrelevant information[42]. Consider this paragraph. Attention causes the higher regions in the cortical hierarchy to receive more visual information from the part of your visual field where this paragraph is located than from other parts of your visual field. Although there are no "pixels" in the brain, high definition (HD) vs standard definition (SD) are useful concepts when thinking about attention. The part of your sensory input that you are paying attention to is transmitted in HD, while the rest of it is transmitted in SD. Hence, more neurons are firing to figure out the HD part of the sensory input than the SD part of it. Two separate inhibitory processes are required for such processing. The first one downsamples the HD information from the eyes to SD. The second one picks out the part of input in HD that you are paying attention to and disinhibits that part of the input so you get it in the original HD.

In the cortex layer 4 receives sensory input, and the inhibitory neurons in the layer reduce that input to SD. Basal dendrites of layer 6 neurons receive cortical feedback[43] from higher-order cortical areas[44], which determines where to direct attention. Then the layer 6 neurons disinhibit (inhibit the inhibitory cells of) the parts of the sensory information that you want in HD by sending information to inhibitory cells in the thalamus[45]. Layer 6 neurons extend their apical dendrites to layer 4 to match the HD and the SD streams[46]. Their tuft dendrites in layer 4 receive feedforward information[xxvii]. They match the information about where the object that attention should be paid to is located in the sensory field, and they send information to the thalamus. The thalamus is a central organ in attention[47], and the information input from Layer 6 allows it to direct attention to where the higher cortical regions want it.

It is known that there are different types of corticothalamic cells (aka cortical neurons that send information to the thalamus) in layer 6[46]. They send projections to different parts of the thalamus. Many parts of the thalamus are poorly studied and the structure is outside the scope of this book. It suffices to say that there are two types of attention – frequently referred to as "top-down" and "bottom-up"[48], although both of them involve feedback mechanisms. Top-down attention allows mammals to focus on specific objects. The previous paragraphs talked about this kind of attention. Bot-

---

[xxvii] In this, layer 6 is an exception. Other layers receive feedforward input in their basal dendrites and feedback inputs in their tuft dendrites. Such an arrangement is necessary for attention to work.

tom-up attention makes sure that new and unexpected stimuli catch the animal's attention. A classic example of this type of attention is the classroom door – nobody pays much attention to it, until it does something unpredicted and opens. This type of attention works slightly differently – as long as inputs are perfectly predicted, they are inhibited. Thus, if a sensory input suddenly becomes unpredicted, it is disinhibited compared to the predicted inputs. The precise mechanism of this is not yet clear, but it is likely that both attention mechanisms rely heavily on layer 6. The different types of layer 6 cells that project to the thalamus probably correspond to the different types of attention.

In many scientific papers, layers 2 and 3 are bundled together. The reason for that is not theoretical but experimental – many experiments are performed in ways that do not allow for distinguishing between them. In this theory, however, these layers have separate, if highly intertwined, functions.

Layer 3 neurons receive feedforward input to their basal dendrites from layer 4[49]. Hence, layer 3 searches for sequences in the spiking of layer 4 neurons. This allows it to track information that persists in your sensory input, even if that input is affecting different layer 4 neurons over time. An example of this in the auditory cortex might be that while layer 4 neurons respond to individual frequencies, like musical notes, layer 3 neurons respond to sequences of these notes. Apart from the obvious benefits of sequence learning, layer 3 neurons help increase sparsity as you go up the hierarchy. If you are looking at a tree, there are many more neurons firing to describe the visual features of the tree in the visual cortex than there are neurons in the higher cortical regions firing to describe the concept of a tree. Such an increase in sparsity is necessary, as it allows the top of the hierarchy to focus on the most important information, which is conceptual.

Layer 3 neurons' apical dendrites extend into layer 1. There they receive feedback input from higher levels of the hierarchy[50]. This allows them to differentiate between predicted and unpredicted patterns, through the BAC-firing, or bursting, mechanism. As described earlier, a neuron only bursts when its sensory input is both new and predicted. Then, if the input persists for a while (as most sensory inputs do in the real world), the neuron transitions to spiking regularly[51].

Some inhibitory neurons, called Martinotti cells, respond to bursts of excitatory neurons[52]. They then inhibit nearby excitatory neurons. This is a winner-take-all mechanism that in layer 3 is responsible for discrete object identification. This deserves an example – let's say that you are at home and you know that your wife is also at home. The layer 3 neurons responsible for identifying her will be getting predictive input to their tuft dendrites because of that knowledge. When you actually see her, many layer 3 neurons that describe women will receive input to their basal dendrites and spike. The neurons that are also receiving the predictive tuft input will burst. This will trigger nearby Martinotti cells to fire and inhibit neurons adjacent to the ones bursting. Martinotti cells make sure that nearby neurons to those receiving prediction are silenced and that your wife is easily identified as herself, and not as a woman who looks like her. After she is identified, the neurons describing her will keep spiking regularly. I predict that layer 3 neurons stop firing when they stop receiving feedforward sensory input.

What happens if a layer 3 neuron receives input that is not predicted? This is the case of bottom-up attention. It means that there is an object out there in the world that needs identification. The neuron spikes, sending information up the hierarchy. As soon as a prediction (which can be also thought of as an identification) is generated, it is sent by feedback to the tuft dendrites of the layer 3 neuron. Then it bursts, and, if the input persists, transitions to regular spiking. Such an event is useful for the cortex to remember, for it is then making a correct prediction or identification. From this stems another prediction of this theory: there exists a separate plasticity mechanism in the cortex based on bursts.

Layer 3 neurons send information to a few different places. The first one is feedforward – layer 3 neurons project to layer 4 neurons in a higher level of the hierarchy[53]. The second one is layer 5 in the same region – more on that later[54]. The third place they project to is layer 2 in the same region[37].

Layer 2 neurons receive local input from layer 3 and layer 4 neurons[40], spike regularly, and send feedback projections down the cortical hierarchy. I predict that they do not burst because of their role, which I contend to be short-term memory. Layer 2 neurons keep information in mind that is not directly evoked by the sensory input at each exact moment. Consider the following scenario – an office worker just looked up from his screen to see a motivational poster. He is now looking back at his screen. Layer 4 and layer 3 neurons that describe the poster have stopped firing. However, it is likely that at some point in the near future he will see the same poster again. Hence, it makes sense for some of the neurons that describe the poster to keep firing That way, if the worker looks at the poster again, the cortex will be able to identify the object quickly without involving the top of the hierarchy. I suggest that those neurons are located in layer 2. Every time layer 3 neurons identify an object, they start regular spiking in both layer 3 and layer 2 neurons. While layer 3 neurons stop firing when the object leaves the sensory field, layer 2 neurons keep firing. This allows the object to be quickly identified if it is encountered again.

There is likely a decay function – the layer 2 neurons' firing rate decreases over time. Otherwise, you would predict all of the objects you have ever seen, which is unsustainable. If layer 2 neurons receive input, however, their firing rate is likely to increase. It is likely that both feedforward and feedback input can bump up the firing rate. In less complex terms – if you hear your dog barking, the layer 2 neurons that describe "dog" increase their firing rate. Also, thinking about your dog, or even just being aware that it's in the same room, will boost the layer 2 neurons' firing rate. How does it work on the level of neurons? Layer 2 neurons' tuft dendrites are close to the cell bodies (because their tuft dendrites are in layer 1, which is adjacent to layer 2); as such, layer 2 neurons do not have pronounced apical dendrites[55]. I predict that NMDA spikes from their tuft dendrites have the same effect on their cell bodies as NMDA spikes from basal dendrites, without the involvement of calcium spikes. Finally, there probably exists a feedback mechanism to stop layer 2 neurons from firing. A part of the hippocampus, which is at the top of cortical hierarchy, stores maps of the environment. When you move from one distinct place to another (e.g. when you go outside), a new map has to "load". Such a change is likely to send waves of information down the cortical hierarchy to stop layer 2 neurons, which were describing information about the inside space that you left, from firing. Other layer 2 neurons will start firing as you receive the sensory input from the place you just moved to.

Memories of recently encountered objects are not the only type of information that layer 2 neurons are responsible for. They are responsible for predictions in general – it's just that memories of recently encountered objects are really good predictions. Predictions that are not recent memories originate high up in the hierarchy and are transmitted down the hierarchy via layer 2 neurons projecting to layer 2 neurons in lower hierarchical regions[56]. Recall how descriptions of objects become sparser as you ascend the hierarchy. Similarly, as feedback flows down the hierarchy in layer 2, the neurons' sparsity decreases. A higher cortical region only needs to predict "dog," while a visual region in the middle of the hierarchy needs to predict visual information about a dog from all angles of view, which involves considerably larger numbers of neurons.

Layer 2 neurons send output to tuft dendrites of layer 3 neurons. That is intuitive – that's where feedback predictions come in to layer 3. They also send output to layer 5 neurons[54].

Layer 5 is the output layer of the neocortex[57]. Neurons in that layer extend axons to various subcortical structures, which is how they influence behavior. For example, layer 5 neurons in the visual cortex send information to the superior colliculus, which is the part of the brain responsible for controlling gaze by coordinating eye and head movements[58]. Layer 5 neurons in the somatosensory and motor parts of the cortex send information to the brainstem to control movement[59]. Higher up in the hierarchy, layer 5 cells can extend axons to all sorts of subcortical structures. For example, some layer 5 neurons in the prefrontal cortex send information to the Raphe Nuclei, which is a serotonin center in the brain heavily involved in regulation of emotional states[60]. Most such projections are outside the scope of this book, but it is worth remarking that from the perspective of the cortex there is no big difference between altering mood and moving muscles. The cortex is an associative learning structure, and its job is to make predictions about information inputs by taking into account its previous actions, no matter whether those are physical or emotional.

As the output layer, layer 5 participates in action selection. For that reason, it receives sensory information, information about prior movements, goals, and predictions. Layer 5 neurons receive inputs from all layers of the cortex[40]. Layer 3 sends information to the basal dendrites of neurons in layer 5. That information is about identified objects in the sensory field. Layer 2 neurons from a higher region in the hierarchy connect to the tuft dendrites of layer 5. Why is that? Why should predictions about sensory objects activate tuft dendrites of output neurons? It turns out that layer 2 neurons encode something else in addition to information about recently encountered objects, and that something else is goals. If you think about how goals need to be treated by the neocortex, you'll realize they need to persist in the absence of sensory input. That is what layer 2 neurons are perfectly suited for. For example, let's say that a sprinter is waiting for the gun to fire to start running. Layer 5 neurons responsible for beginning to run would receive continuous input to their tuft dendrites from layer 2 neurons that encode the goal of sprinting. Then, as soon as the auditory cortex layer 3 neurons send input to the same layer 5 neurons, they burst and the athlete begins running. Interestingly enough, because those layer 5 neurons are not at the top of the hierarchy but somewhere in the middle of it, professional sprinters begin running before they are conscious of hearing the sound[61]. Beginner runners, however, don't have that connection from auditory cortex neurons to motor cortex layer 5 neurons. For them to identify what to do based on the sound of the gun firing, the involvement of higher level cortical areas is necessary. As such, more time is required for them to begin running.

Layer 5 neurons can fire regularly, as well as burst[62]. That makes sense as behaviors like walking need to be maintained for a long time. There are also Martinotti cells in layer 5[52] that are responsible for discrete behavior selection, much like those in layer 3 responsible for discrete object recognition. For example, you don't want to say half of "Hello" and half of "Good day"[xxviii], even if both are applicable; you want to say one specific phrase.

Some layer 5 neurons send feedback information to lower regions in the hierarchy. Behaviors are arranged just as hierarchically as objects in our brain. A high-level behavior like opening the door after you hear a doorbell involves many separate behaviors of lesser complexity like walking and reaching and grabbing, and a myriad of variations in the movements of actual muscles. Much like the feedback flow in layer 2, feedback flow in layer 5 decreases sparsity as you go down the hierarchy. An interesting feature of both layer 2 and layer 5 feedback connections is that they arrive to layer 1 of the lower cortical region. David Hubel, one of the most influential neuroscientists ever, once described layer 1 of the neocortex as its "crowning mystery"[63].

[xxviii] Unless you do, in which case, Heday to you too!

Now that we know about BAC firing and how it enables matching feedback and feedforward information streams, the purpose of layer 1 is clear – that's where layer 2, 3 and 5 neurons receive feedback information to their tuft dendrites from farther up the hierarchy.

The information that layer 5 neurons send to tuft dendrites of layer 3 helps the cortex make sense of sensorimotor transitions. Knowing how you move your head helps predict what you're going to see. The connection from layer 5 to layer 2 is a bit trickier. Layer 5 sends information about actions to layer 2 where it persists for a bit. This means that recent actions are stored in the cortex, allowing you to learn consequences which do not immediately follow actions – on the cortical timescale, there is quite a gap between you saying something and hearing the response.

There are also layer 5 neurons that send information feedforward to the basal dendrites of layer 4 in the higher cortical region. While lower regions in the hierarchy primarily learn only patterns in sensory data, higher regions learn mostly sensorimotor transitions. That makes sense, given that most of the changes in our sensory input are caused by our movements. A reasonable question is – isn't this what the layer 5 feedback connection to the tuft dendrites of layer 3 neurons is for? Almost, but not quite. The feedback connection to layer 3 neurons makes predictions about lower level features – like that moving your eyes will shift a table that you currently see to a different part of your visual field. The layer 5 feedforward connection makes predictions about higher level features – like moving your eyes will cause you to see a piano, which you currently aren't able to see. Since maps of our environments are stored very high in the cortical hierarchy, information about movements has to go up the hierarchy to predict which objects will be encountered. From there, information about these objects propagates down the hierarchy through the layer 2 feedback pathway to make predictions about the distinct sensory features of these objects.

It also should be noted that some layer 6 apical dendrites terminate in layer 5. Moreover, layer 5 neurons sometimes project to layer 6[46]. It is likely that these links have to do with coordinating the movement of sensors and the internal mechanics of attention in the brain. If you want to pay attention to somebody's face, you need to both turn towards them and utilize the attentional mechanics explained earlier. Since your turning changes the other person's position in your visual field, layers 5 and 6 need to coordinate their activities to make sure that attentional mechanics apply to the part of the visual field where the person is going to be after you turn, not before you do.

In summary, layer 4 neurons identify patterns across a wide range of the sensory field. Layer 3 neurons identify discrete objects in spatiotemporal patterns of layer 4 neurons' firing. Layer 2 neurons make sure that information not immediately present in the sensory field persists in the cortex. Layer 5 neurons select actions. Layer 6 neurons work in tandem with layer 4 neurons and the thalamus to implement attention.

This view is undoubtedly somewhat wrong. However, at present, neuroscience offers no compelling theories about what the different cortical layers do. As such, I think that this view is less wrong than other theoretical descriptions of the neocortex, which is why I thought this book was worth writing. This chapter introduced theories about what the different pieces of the neocortex do. The next chapter puts them together in an example to show how they cooperate on a task in a living brain.

6

CORTEX IN ACTION

The capabilities of the cortex come not only from those of individual neurons or the laminar structure, but also from its hierarchy and its sheer computational power. The following example is set to illustrate how the hierarchy works together to accomplish a task that, while trivial for humans, is nearly impossible for even the most advanced robots today – getting food from the fridge. Note that this example by no means describes all of the neural circuitry involved in that process – just a slice of it that shows how cortical regions use the connections between different layers in the hierarchy.

In this example, the behavior starts with a goal: hunger. The feeling of hunger in the brain originates in a subcortical structure called the hypothalamus, which gets information directly from your stomach[67]. The hypothalamus then relates the information about hunger to a region of the cortex called the insular cortex[68]. This is not the only place in the brain the hypothalamus projects to, however. It also projects to the thalamus, the amygdala, secondary motor cortex, and other places[68]. The insular cortex is important here because it's the part of the cortex known to be implicated in hunger[69].

The insular cortex has reciprocal connections[xxix] with both the sensory parts of the cortex and the prefrontal cortex that sits near the top of the cortical hierarchy[68]. In fact, this is a common theme of the cortex – information about goals from subcortical structures is first sent to the middle parts of the cortical hierarchy. It is then integrated into the feedforward information stream in the cortex – layer 3 of the insular cortex transmits information about hunger to layer 4 in the prefrontal cortex (PFC). This allows the PFC to make predictions about goals (e.g. I'm going to be hungry later). It also allows the PFC to see all of the current goals and make decisions based on them[xxx].

The inputs about goals to the PFC are not discrete, but scalar. In your brain, there is a scale of how hungry you are. In this case, the scale is likely represented by how many neurons that describe hunger are firing in the insular cortex. Thus, the method of choosing an action is something like neurons voting – the goal that sends the strongest input (the most spikes) to the PFC wins, and the PFC decides to embark on a course of action to achieve that goal. Other goals are temporarily not acted upon. In terms of layers, the PFC is a bit different than most of the cortex. While most cortical regions only have information about goals encoded in layer 2, the PFC deals with goals in every layer. It receives information about all the different motivations in the feedforward stream that enters layer 4. Layer 3 is where the goals compete and the ones to act on now are chosen. These are sent to layers 2 and 5, from where they propagate down the hierarchy.

[xxix] Most feedforward/feedback connections in the cortex are reciprocal.

[xxx] Subcortical structures, notably the basal ganglia, are heavily involved in decision-making.

The PFC also sends information from layer 3, which in this example is "I should go get some food" to the hippocampus[xxxi,70]. The hippocampus sits at the very top of the cortical hierarchy. The hippocampus stores short-term memories, maps of places, information about recently encountered objects, events and so on[71,72]. Note that the hippocampus only stores very basic representations of objects; their full descriptions are stored in the sensory parts of the cortex. In this example, the hippocampus has stored information that there is an apple in the fridge, which would satisfy the goal of hunger. The hippocampus has two parts called ventral and dorsal. Not coincidentally, there are two parallel ascending pathways in the visual cortex, also called ventral and dorsal. People also frequently refer to them as the "What" (ventral) and the "Where" (dorsal) streams[73,xxxii]. The What stream performs object recognition. Hence, in the ventral part of the hippocampus, there is a memory of an apple. The Where stream creates a 3D map of the environment. Thus, in the dorsal part of the hippocampus, there is a map of how to get to the fridge[xxxiii]. In our perception, objects are inextricably linked to their location in space. It is not known at what point the brain integrates information about their location – it may be in the hippocampus, somewhere lower in the hierarchy, in the claustrum, or all of the above.

Regardless, in this case the hippocampus sends information about the apple and the location of the fridge back to the PFC. They are now concrete goals propagating down the cortical hierarchy through layers 2 and 5. The first part of getting the apple involves moving to the fridge. To do this a human utilizes their motor cortex

---

[xxxi] Through the midline thalamus.

[xxxii] There is likely a similar division in the auditory cortex.

[xxxiii] The What stream from the PFC to the hippocampus goes through the perirhinal cortex, and the Where stream goes through the parahippocampal cortex.

and visual cortex extensively. The motor cortex receives input from the PFC[74] in its feedback pathway (PFC's layer 2 projecting to the motor cortex's layer 1). That information is largely about the goal of getting to the fridge. The motor cortex also receives information about the current body position from the somatosensory[75] cortex. Interestingly enough, that information does not come via a feedforward pathway. The primary motor cortex does not have a traditional layer 4 because the information there is dependent not on sensors but on movements, which are caused by the cortex[76]. Hence, the somatosensory cortex projects to layers 2/3, 5 and 6 in the motor cortex. Then the motor cortex integrates information about body position with the feedback information of where it should get to and plans a route. Layer 5 neurons of the primary motor cortex fire when a behavior is selected and project to subcortical structures that handle movement[77].

For navigation, the visual cortex mostly utilizes the "Where" pathway. Incidentally, an influential paper once proposed to rename it to the "How" pathway[77]. Recall that the cortex evolved on top of a brain that already handled action selection. Thus, the analysis of visual information became heavily integrated into this mechanism via the Where/How pathway. Information analysis for the purposes of identification evolved a separate pathway – the What pathway[xxxiv].

There are a few things happening in the Where pathway as part of our quest for the apple. A region of the cortex called the Frontal Eye Field (FEF) directs targeted eye movement[78]. It receives information via the layer 2 feedback pathway that the human is currently interested in the fridge. It also receives feedforward information

---

[xxxiv] It is curious that in reptiles and insects action selection works without the What pathway.

from visual regions MST and VIP that encode the environment in head-centered space[79]. By matching the desired object and the position of the head, FEF makes the decision on where to direct eye movement. It sends out the motor command via layer 5, as well as the attentional command via layer 6 to respectively look at and focus on the fridge. The information about the eye movement is also sent out both feedforward and feedback – feedforward from layer 5 to layer 4 in the next higher hierarchical regions and feedback to layer 2 of lower hierarchical regions. The higher regions predict what new objects are going to be seen, and the lower regions use the feedback to adjust the position of objects due to eye movement.

The What stream of the neocortex receives feedback predictions about both eye movements and the objects that it expects to encounter. The higher regions of the cortical hierarchy have a pretty good idea about what objects are located between the desk and the fridge and they send feedback down the layer 2 pathway. As mentioned in the previous chapter, these predictions decrease in sparsity as they go down the hierarchy. In V4, which is a cortical region in the middle of the What stream, high level features of objects such as colors and large shapes are identified. In V1, which is the region of the visual cortex most closely connected to the retina, low level visual features like edges are predicted – and there are a lot more of those than there are large blobs of color. In general, feedback in the visual cortex (and in other parts of the cortex) is at least as important as feedforward information. Because of this, some scientists call the cortex's method of information encoding "predictive coding"[10]. As a side note – even though at this point your back is turned to the desk, layer 2 neurons describing the visual features of objects on your desk are still spiking regularly. They are still good predictions, in case you turn around.

Note that the regions of the visual cortex mentioned above are only examples of what the visual cortex does. The structure has many regions – here are two examples of maps of the visual cortex[53]. There are many other competing maps. In the cortical hierarchy, many regions connect to regions of the same level. Such connections have not been mentioned before, but they go through the regular feedforward and feedback pathways.

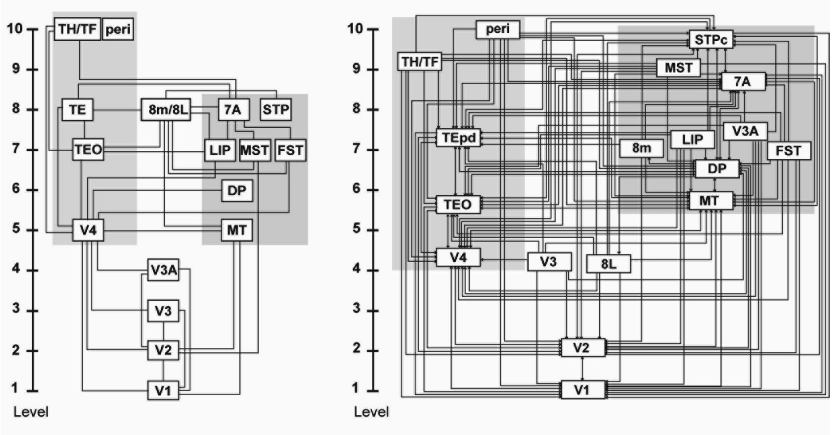


Figure 13 – Two Maps of the Visual Cortex

(Adapted from Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex under CC-BY-3.0 https://creativecommons.org/licenses/by/3.0/ Original - https://www.ncbi.nlm.nih.gov/pmc/?term=23983048%5BPMID%5D&report=imagesdocsum*)*

Eventually the motor cortex and the Where stream in the visual cortex, together with the help of the maps in the hippocampus, are going to get you to the fridge. When the PFC discovers that you are within reaching distance of the fridge, a new set of behaviors will be required. Reaching and grabbing are the domain of the posterior parietal cortex – a structure most prominent in primates, since they are the best at reaching and grabbing things among mammals.

The posterior parietal cortex (PPC) is near the top of the Where stream, and thus receives feedforward visual information about the map of the environment. It also receives sensorimotor and vestibular information[80] so that it knows the location of the hand. All of this information is feedforward, so it enters layer 4 of the PPC. It also receives information about the reaching target from the PFC – in this case, the fridge door handle. The PPC matches the feedback and the feedforward information in its layer 5 cells, which then burst to start the reaching behavior.

The pattern of matching feedforward and feedback information continues to repeat in the cortex. It is the primary operation of each cortical region. It is an incredibly powerful operation that is the center of the "common cortical algorithm". It is versatile and allows for pattern recognition and action selection in any part of the cortex.

The PPC is similarly responsible for opening the door. Once you open the fridge, there's a surprise – your wife has bought groceries! This is where the What stream of the visual cortex shines. Visual information about unpredicted items is matched with the feedback of items that are usually found in fridges to identify objects like yogurt and beer. However, if there is an unpredicted object in the fridge, like a very unhappy cat, information about it flows all the way up the hierarchy so it can be identified. That's why surprising information captures a disproportionate amount of our attention.

The apple is also identified. Attention in all parts of the visual cortex is directed to it (assuming that there is no cat in the fridge) through the top-down mechanism originating in the PFC. The PPC plans a reaching and grabbing motion. Here is a good example of single-neuron mechanics coming into play. Consider closing your fingers on an apple. You need to change the behavior from closing your fist to carrying the apple as soon as you grab it. There are neurons in your somatosensory cortex that respond to your touching the apple exactly as Johnny from chapter 4 responded to your finishing your drink. However, Johnny was not localized to a specific layer. In this example, neurons in layer 3 of the somatosensory cortex are the ones to respond, identifying that you are now touching the apple. Layer 5 neurons, meanwhile, are receiving input to their tuft dendrites telling them that as soon as you touch the apple you should bring it to your mouth. When they receive input from the aforementioned layer 3 neurons, they burst to begin that motion.

You bite the apple and eventually it enters your stomach. The hypothalamus receives information that there is food in your stomach and you are no longer as hungry[xxxv]. The PFC and subcortical structures pick other things to do.

Decision making almost always includes neuromodulators (generated in the aforementioned subcortical structures). They are the "value function" of the cortex. Thus, the predictions of the apple also trigger reward circuits in the brain, to make sure they win the competition for attention the next time.

Another mechanic omitted so far is learning. In this example, you learned that there is yogurt and beer in the fridge. That learning started with identifying the objects in the What pathway of the visual cortex. This information is carried via the feedforward pathway to

---

[xxxv] Actually, it is likely that even before you know that you are no longer hungry through the stomach, the cortex makes a prediction that soon you will not be hungry anymore because you've eaten and, consequently, it inhibits hunger.

the PFC, the perirhinal cortex and the ventral hippocampus where it is stored. Neuromodulators are involved here also – they make sure that you remember what food there is from only seeing it once, since it is associated with reward. Most of the other learning in the cortex takes repetition.

The thalamus has also been conspicuously absent from this chapter. Its method of operation remains unclear, although we know it relays sensory information to the cortex and is important in attention. Consciousness has been largely ignored as well. The one thing we know is that you are not conscious of the information in the primary sensory parts of the cortex, like V1. Consciousness appears somewhere higher up in the hierarchy and involves the claustrum.

Finally, an attentive reader might have noticed an inconsistency. Earlier this book mentioned that cortical regions aren't necessarily discrete entities, but this chapter treated them as such. Unfortunately, there is no way around this yet. The types of studies conducted so far mention specific cortical regions, and our understanding of the cortex can only be as good as the studies we have. When we have cell-level connectomes of the cortex (and algorithms to analyze them, for it is impossible to manually analyze trillions of connections) we can make better subdivisions.

Hopefully, this example was helpful in understanding the cortex. The next two chapters discuss how we can use what we learned in building better AI. If you are only interested in biology, feel free to skip them.

7

USING INSIGHTS INTO THE NEOCORTEX IN ALGORITHMS

Modern neural networks can do remarkable things. They can recognize spoken words, drive cars, beat humans in many games, and predict which news stories will interest which people. In some respects, however, they are far behind living animals. They cannot move well through forests, play volleyball, or make movies[81]. The biggest gap between existing neural networks and brains is that neural networks cannot learn new skills – each network can only do the thing it's designed to do. A neural network that recognizes cats in pictures cannot play Go, for example. Mammals, on the other hand, can acquire a large variety of new skills. While many improvements can be made with current technology, to truly bridge this gap, we will need to create completely new types of neural networks, inspired by the neurons and the architecture of the neocortex. This chapter discusses how the new ideas about neurons and the neocortex that have been presented in this book can be used to create better neural networks.

Many improvements are available across the board – better artificial neurons, cortex-inspired network structures, and proxies for other parts of the brain. However, the first element that warrants a discussion is architecture. Certain features of computers that we take for granted are actually results of architectural choices made decades ago:

1. Computers are built to work extremely reliably. If you perform a calculation on your phone 1000 times, you are going to get exactly the same answer every time. Computers extensively use error correction codes to achieve that.
2. Computers consume a lot of power. A desktop computer can consume 1000 watts of power. An adult human brain consumes only around 100 watts.
3. Computers utilize a low number of very fast cores. A typical CPU may have 2-8 cores. In recent years, computers have begun making extensive use of parallelization – using hundreds or thousands of computing cores at once, especially in processing graphics. But the way this parallelization is set up, all of the cores are highly dependent on each other. Frequently, until all cores finish a certain job, none of them can start a new job.

In sum, computers are highly reliable, consume a lot of power, and their parallelization is highly interdependent. Brains are not like that. As mentioned above, synapses are unreliable. Brains use large numbers of neurons and spike thresholds to overcome this. They consume little power because they utilize spiking and sparse representations. Brains take advantage of massively parallel processing units (neurons) that operate in an asynchronous fashion. That means that while neurons depend on other neurons firing, and sometimes are synchronized, no neuron "waits" on another neuron to finish its processing. Each neuron operates on its own clock, so to speak (although, as mentioned before, inhibitory neurons can help synchronize excitatory neurons). Ideally, new hardware would be needed to create next-gen neural networks, and the next chapter talks about some efforts in that direction. But to know what hardware would run artificial neurons the best, we would first need to know what those artificial neurons are. This chapter describes Artificial Branching Neurons (ABNs) – a new type of artificial neurons. The word branching is used in their name because branches serve three very important purposes:

1. They increase the computational power of neurons.
2. They provide a mechanism for handling variable sequences.
3. They introduce a limit on the number of synapses a brain can have.

The third reason deserves some attention. As mentioned in Chapter 3, as you increase the number of artificial neurons in a model, the number of potential synapses increases exponentially. That's not scalable. However, if neurons can only form synapses with their dendrites, or branches, they can only form synapses with neurons whose axons are close to their branches. That introduces a limit on scaling and allows us to create networks the size of the human brain[xxxvi].

There are three ways to model branches. The first one is purely logical. In this way, the dendritic tree is modelled only as a set of bifurcations, without such attributes as length or direction. The second one is spatial. In this way branches exist in a modelled space, and every branch can be thought of as a line segment. The third way is "real 3d spatial". In this model, branches have width and two branches cannot exist in the same space, much like in the real world. Video games use a type of software called colliders to make sure that such overlaps don't happen – they are fairly computationally intensive. While it is possible that branch width has a computational role in biological neurons through electrodynamics, that may be just as accurately modelled by synapse weight. Colliders offer no obvious computational benefit. Length, however, is important – as we know, the length of a dendrite helps neurons deal with time. Hence, ABNs follow the second approach – spatial – and use line segments as their branches.

The laminar structure of the neocortex is a huge component of its computational power, and, more importantly, its versatility. It would be wise to incorporate a similar structure into networks of ABNs. As the last few chapters showed, neurons in different layers of the cortex have slightly different mechanics. Therefore, ABNs will be different depending on the layer they are located in. Inhibitory cells do not need to be modeled – in code they can exist as mechanics of excitatory cells[xxxvii].

Regardless of the layer, all of the ABNs have similar den-

[xxxvi] A neuron typically has thousands of synapses. If a neuron could be connected to any other neuron, it could have tens of billions of synapses.

[xxxvii] It is likely that as we understand them better, inhibitory cells will reveal functionality that will require us to model them – but we're not there yet.

drites. All dendrites can generate dendritic spikes, modeled after NMDA spikes. Here is an algorithmic description of what a dendritic segment would compute in a timestep[xxxviii]:

```
for s in synapses:
if s.active:
        dendrite.potential=dendrite.potential+s.weight
        for d in nearby_dendrites:
                d.potential=d.potential+s.weight*power(dendrite,d)
if dendrite.potential>dendrite.threshold:
        dendrite.trigger_NMDA_spike()
        dendrite.potential=0
if dendrite.potential>=0:
        dendrite.potential=dendrite.potential-DECAY
```

This looks a lot like existing artificial neurons. The only exception is that an active synapse increases the potential not only of the dendritic segment where its located, but also that of the nearby and connected dendritic segments. The *power* function computes how much effect a synapse on one dendrite has on another by multiplying the distance between the segments by some negative coefficient constant. The interesting bit comes in the trigger_NMDA_spike() method:

```
a    neuron.potential= neuron.potential+SPIKE_STRENGTH/dendrite.distance
b     for d in dendrite.nearby_dendrites:
c          d.threshold=nearby_dendrites.threshold-coop_power(dendrite,d)
```

[xxxviii] This book uses discrete timesteps and simple thresholds to make it easier to read. Differential equations can also be used in implementations of ABNs.

```
d   for axon in nearby_axons:
        if axon.fired_recently:
            if synapse(axon,dendrite) in synapses:
            synapse(axon,dendrite).weight+=W
        else:
            synapses.add(synapse(axon,dendrite))
```

Line *a* shows the effect of an NMDA spike on the cell body. It is dependent on how far away the dendritic segment in question is from the cell body. Lines *b* and *c* implement NMDA spike cooperativity. An NMDA spike reduces the threshold for further NMDA spikes in adjacent dendritic segments. The *coop_power* function is a function similar to *power* that computes the effect of the cooperativity of one dendritic segment on another based on their distance. The constants have to be chosen in such a way that distant dendritic spikes have a large proportion of their effect on the soma through cooperativity (and therefore through facilitating other dendritic spikes), while dendritic spikes that occur close to the soma directly increase the neuron's electric potential on their own. On a separate thread, for every dendrite there is a function that allows for thresholds to return to normal (it can be thought of as the decay of the cooperativity effect). It looks something like this:

```
while (dendrite.threshold<DEFAULT_THRESHOLD):
    dendrite.threshold=dendrite.threshold+X
    sleep(Y)
```

The learning algorithm starts on line *d* above. It states that if an NMDA spike was preceded by a nearby axon firing, it should either strengthen the existing synapse with said axon or create a new one, if one doesn't. Naturally, synapses need decay – otherwise every dendrite would be connected with every nearby axon eventually. The decay can be flat, like this:

```
for synapse in synapses:
    synapse.weight=synapse.weight-SYNAPTIC_DECAY
    if synapse.weight<0:
        synapse.remove()
```

Or it can be contingent upon the strength of the synapse:

```
for synapse in synapses:
    if synapse.weight<permanence_threshold:
        synapse.weight=synapse.weight-SYNAPTIC_DECAY
    if synapse.weight<0:
        synapse.remove()
```

This choice is dictated by the application. In humans and animals the second mechanism is in play – some things are never forgotten, like that the sky is blue. But in some machine learning applications this need not be the case.

Neurons in most layers also have tuft dendrites, which can initiate calcium spikes that lead to bursting. Their algorithmic implementation is straightforward – an NMDA spike in one of the tuft dendrites sets a flag for the neuron that there has been a recent calcium spike:

```
if dendrite.zone = “tuft”:
    neuron.recent_calcium_spike=true
```

This recent calcium spike is utilized in the code of the whole neuron. The neuron code goes like this:

```
if neuron.regularly_spiking==false:
    if neuron.potential>neuron.threshold:
        if neuron.recent_calcium_spike==true:
            neuron.initiate_burst()
        else:
            neuron.initiate_spike(0)
else:
    frequency=frequency+neuron.potential*M – FREQUENCY_DECAY
    neuron.initiate_spike(frequency)
neuron.potential=neuron.potential-NEURONAL_DECAY
if frequency<FREQUENCY_THRESHOLD:
    frequency=0
    neuron.regularly_spiking=false
```

There are a few things happening here. First, there is a check whether the neuron is already spiking regularly. In real brains, it is implemented by inhibitory neurons. If the neuron isn’t regularly spiking, it checks if its potential is over the spike threshold. If there has been a recent calcium spike, the neuron bursts. If there hasn’t, it spikes.

If the neuron is firing regularly, however, the inputs are summed up to increase the frequency of its firing. In real brains, that mechanism is handled by the regular spiking inhibitory cells.

Note that this algorithm is the general form of an ABN neuron. It is applicable to neurons in layers 3, 5 and 6. Layer 4 and layer 2 neurons do not burst. There are no neurons in layer 1.

There are also some differences in the constants for neurons of different layers. For example, FREQUENCY_DECAY is much higher for layer 3 neurons than for layer 2 neurons. This difference allows for layer 2 neurons to persistently fire in the absence of sensory input, while layer 3 neurons will stop firing when the input stops.

Chapter 5 posited that bursts initiate regular firing. In this algorithmic representation, this happens in the initate_burst() method shown below:

```
initiate_spike()
sleep(Z)
initiate_spike()
sleep(Z)
initiate_spike()
neuron.regularly_spiking=true
frequency=starting_frequency
for neuron in nearby_neurons:
    neuron.threshold=neuron.threshold+I
```

Note that only layer 3 and layer 5 have the competitive inhibition part in the last two lines. Layer 6 is responsible for attention. Unlike action selection and object identification, where we want to select just one option, attention is somewhat elastic. We can pay attention to multiple input sources at once, even though that may produce diminishing returns at times.

Furthermore, since the thalamus is not yet properly understood, it cannot be implemented in code. Luckily, we know that layer 6 is involved in attention, which is a disinhibitory mechanism in the feedforward stream. Hence, for the purpose of this implementation, layer 6 neurons can project to layer 4 neurons.

Since ABN neurons exist in three dimensions, their learning mechanisms have to incorporate growing branches. The mechanism for that is homeostatic, meaning that a neuron must maintain a certain number of synapses. If the actual number drops below that threshold, the neuron forms new synapses with axons that fire close in time to the neuron's dendritic spikes. This is similar to how biological neurons operate – they protract and retract dendritic spines all the time.

The first thing needed for this mechanism to work in code is the ability to form new synapses. The code for that is executed when an NMDA spike occurs in a dendrite:

```
if neuron.number_of_synapses<MAX_NUMBER_OF_SYNAPSES:
        for axon in nearby_axons:
                if axon.fired_recently:
                        form_synapse(axon, dendrite)
```

This code checks for axons that are near a dendritic segment after an NMDA spike occurs in that segment. If a nearby axon had fired before the NMDA spike, a synapse is formed. New branches are only formed if the number of synapses on a neuron falls significantly below the maximum number of synapses – to, say, 80% or so. This check occurs independently of the neuron firing (on a separate thread), with some regular frequency:

```
if neuron.number_of_synapses< MAX_NUMBER_OF_SYNAPSES*.8:
        dendrite_to_extend=pick_random_dendrite ()
        form_new_dendrite(dendrite_to_extend)
```

If the number of synapses falls below a threshold, a new dendritic branch is grown. The pick_random_dendrite method simply picks any branch with an endpoint that is not at maximum length yet. The form_new_dendrite (dendrite_to_extend) method is shown below:

```
new_dendrite=dendrite()
new_dendrite.starting_point=dendrite_to_extend.ending_point
new_dendrite.zone=dendrite_to_extend.zone
new_dendrite.direction=dendrite_to_extend.direction * dendrite_variation()
new_dendrite.ending_point= new_dendrite.starting_point + DEFAULT_DEN-
DRITE_LENGTH * new_dendrite.direction
new_dendrite.nearby_dendrites=[]
new_dendrite.nearby_dendrites.add(dendrite_to_extend)
dendrite_to_extend.nearby_dendrites.add(new_dendrite)
dendrite_to_extend.neuron.dendrites.add(new_dendrite)
```

This code extends an existing dendritic branch. The direction of the extension is the direction of the existing branch with some variability added through the dendrite_variation() function. If the new branch does not form synapses after some period of time, it is retracted.

Axons have a similar homeostatic mechanism, except they branch out more widely. To achieve that, they would use an axon_variation() function, which produces larger variations than dendrite_variation(). Speaking of axons – there is no need to model the entire 3D structure of an axon for ABNs. Real axons can be myelinated – covered by an insulating layer of glia cells – along most their length, only allowing axons to form synapses close to the end of the axon. This allows for long-distance transmission that does not activate every neuron in between the neuron and its target. ABNs can only model the axon as its unmyelinated segment at its target (which is arborized, that is to say it branches out at the end). Note that some axons are unmyelinated – those are mainly useful for local transmission. Those would need to be modelled in their entirety.

In brains, about 80% of all connections within the cortex are local (within the same cortical region), while 20% are long-distance connections between different cortical regions[82]. This is a useful heuristic for those building networks of ABNs.

The specific hierarchy of regions in an ABN network would depend on the application, and its computational requirements, sensors and behavioral range. Nonetheless, a few guidelines are likely to be universally applicable. There should be single sense regions for each of the sensors. Multisensory regions should be created based on the correlations between the data that comes in through different sensors. Reward values and other modulators should come in the middle of the hierarchy. The complexity of the sensor determines the size of the sensory region; the complexity of the patterns in the data determines the height of the hierarchy. In humans, the top of the hierarchy is the hippocampus, which helps store episodic memory and maps. A lot of applications would need that functionality, but since the hippocampus is not yet completely understood, a simpler solution can be utilized that would work for many machine learning applications. Layer 3 cells at the very top of the hierarchy can simply complete the loop by projecting only to layers 5 and 2 in the same region and nowhere else. It should be noted that there was a famous case of a person losing their hippocampus and still being able to learn new skills[83]. That state would still be a great improvement on current generation AI. Episodic memory is important but not crucial at this point.

One important feature of the cortex that was described earlier is the dual feedforward streams – the What and the Where streams. It is curious how the separation between the two streams occurs – why do cortical regions with the same structure at the same level of the hierarchy, both of which receive information from the primary visual cortex, perform such different functions? How does one of them identify objects and one of them create maps? The difference is purely in the connectivity of these regions. The Where stream receives extensive information about the movement of the eyes and the head. Its job then is integrating visual information with information about movement, which is basically putting objects on a map. The map itself comes via feedback from the hippocampus. The What stream, on the other hand, does not receive information about the movement of the sensors. It only receives two types of information: what the eyes currently see, and what objects have been seen earlier, which comes from feedback. Such an arrangement allows for object identification. For computing purposes, this differentiation may be useful in many applications.

As far as neuromodulators go, there currently is no need to create separate artificial neurons for them. They can just be assigned as parameters for existing ABNs. Say that a predicted outcome has a negative reward value – the motor neurons leading to that outcome would be inhibited. If a predicted outcome has a positive value associated with it, associated actions are boosted. The way these parameters are assigned these values is simple: when something good happens, the active neurons get assigned a positive value. When something bad happens, the opposite occurs. Humans have a variety of neuromodulators that are responsible for different outcomes: hunger, sex, fear, etc. Initially, ABNs can just use positive and negative reward values. However, more specific parameters, such as "Critical failure – do anything to avoid," can also be assigned.

One feature of biological intelligence that helps it achieve impressive outcomes is its ability to experiment. ABNs can do that too by having motor neurons spontaneously fire with some probability. This mechanic introduces variability to actions, and makes ABN networks learn what actions yield the best outcomes.

One issue that has not been addressed so far in this chapter is threading. Unlike traditional neural networks, networks of ABNs do not need to operate in a synchronized fashion. Separate threads run for dendrites, neurons, and functions that are not dependent on immediate activity like growing new branches. It is not feasible in such a configuration for threads to be synchronized and dependent on each other. As a result, there is no "master clock" and no "network time step" – real neurons run on their own clocks. An obvious problem with such an approach is memory – wouldn’t memory locks create massive desynchronizations causing neurons that represent the same sensory input to fire at different times? The solution to this issue is simple, if unorthodox from the perspective of computer science – it’s a laissez-faire approach to error handling. Let’s say that a dendrite thread is trying to see if a synapse is active, but the synaptic weight is currently being updated by a separate thread. The dendrite thread will not wait for the update to end, but will instead just report the synapse as not being active. This is the beauty of dendritic spikes – if the threshold for a spike is, say 7 active synapses, it does not matter if there are 11 or 12 synapses active at any given point – the spike still goes through. Hence, occasional “faulty synapses” would have no strong effect on the network. This is a feature of the neocortex, and, with “laissez-faire error handling”, it is a feature of networks of ABNs.

8

HARDWARE FOR NEXT-GEN NEURAL NETWORKS

Recently, a few articles appeared in the news about IBM creating a computer as powerful as a brain of the mouse[84]. IBM developed a state-of-the-art computer called TrueNorth that has 48 million artificial neurons. Unfortunately, their neurons are closer to McCullogh and Pitts neurons than real biological neurons. They do not model dendritic spikes, calcium spikes, bursting and other complex neuronal behaviors. This means that in terms of sheer computational power, the state of the art technology today is still orders of magnitude away from the simplest mammalian brains. But, as computing power increases exponentially, we may yet see computers as powerful as brains within our lifetime.

IBM was right about something big, though: computers as powerful as the brain are not going to utilize the Von Neumann architecture that powers most modern computers. Von Neumann computers simply use too much power to be scaled to the size of a brain. An entirely new class of hardware will be developed for the next generation of brain-like computers. While it is not yet certain what that class will be, this chapter reviews the technologies that look promising today.

The IBM TrueNorth computer is a chip, which runs pre-trained neural networks (optimized for deep learning), while requiring very little power. Such a chip would be incredibly useful for many commercial applications – low power speech recognition, for example – but gets us no closer to the brain. The brain constantly learns, which the IBM TrueNorth chip does not[xxxix].

The Human Brain Project's original goal was to build a computer that simulates the entire human brain. They have not achieved that, but the project now works with several novel computing architectures. The first one is BrainScaleS[85]. BrainScaleS is a hybrid computing system. Instead of representing information only as 1s and 0s, BrainScaleS also represents information as physical quantities – voltage, current and charge. The analog computing elements are extremely efficient at modelling certain aspects of a neuron, such as membrane potential, while using very little power. When the spiking threshold is reached, the spike signal is transmitted digitally, which makes sense due to the binary nature of spikes. BrainScaleS neurons support such features of real neurons as bursting, regular spiking and spiking frequency changes. They implement dendritic spikes and plan to implement calcium spikes. The good thing about the BrainScaleS analog implementation is that dendritic spikes are easily implemented without greatly increasing the power budget. Another helpful feature of the BrainScaleS system is that it runs at 10000 times the speed of actual neurons. This is somewhat counteracted by the fact that the system requires millions of parameters to run and, consequently, configuring the system for an experiment can take a

---

[xxxix] Even though the chip is optimized for "deep learning", it only allows running pre-trained models. These models have to be trained on other computers.

very long time[86].

About 400 million BrainScaleS synapses fit on a silicon wafer, so 2.5 million wafers are needed to achieve the level of complexity of the human brain. That does not seem feasible at the moment. That being said, the system most accurately simulates biological neurons as we understand them today[xl]. This allows running experiments that expand our understanding of networks of neurons in the brain.

Another architecture supported by the Human Brain Project is the SpiNNaker architecture. It involves taking many ARM cores and linking them together in a supercomputer. The advantage of the SpiNNaker system over traditional supercomputers is that it consumes less power, as ARM cores are primarily developed for smartphones where power usage is an extremely important consideration. The advantage of the SpiNNaker system over BrainScaleS is that it is easily scalable. It does not, however, aim to model biologically realistic neurons, at least not to the extent that BrainScaleS does[xli].

SpiNNaker is similar to many projects that have been launched in the commercial space that use GPUs or TPUs. They generally have many more cores than CPUs and consume less power. None of these projects, however, are aiming to simulate real biological neurons or the architecture of the cortex.

Some approaches eschew traditional digital computation altogether. They are in their infancy, yet they seem promising. The first one is optical computing. Optical computing uses light instead of electricity. It can perform such operations as the Fourier transform and linear algebra operations quickly and in a massively parallel

---

[xl] The limits of our understanding are discussed in the next chapter.

[xli] Both BrainScaleS and SpiNNaker allow researchers to use their systems for experiments. You can find more information about how the system works and request access at neuromorphic.eu

fashion[xlii]. It is not yet obvious how optical computing can be used for creating brain-like systems.

Quantum computing is at its infancy. It can loosely be divided into two groups – quantum annealing systems and a full digital quantum computer. Quantum annealing systems exist today and are not unlike the BrainScaleS system in that they utilize an analog computational process. However, in the case of quantum annealing, the analog component computes the global minimum of a given objective function. This computation may be very important in brain function. One of the leading theories about overall brain function describes it as a system the goal of which is to minimize its free energy (this is called the free-energy principle)[87]. Quantum annealing could be extremely helpful in this computation. That being said, there is still an extremely wide gap of understanding between the free-energy principle as a principle, and the biological reality of neurons with their active dendrites. Moreover, quantum annealing is only a little bit beyond the proof of concept stage. To put it simply, it is possible that quantum annealing systems will help model brain-like computers, but there is no immediate work to be done in this area – we are years, if not decades, away from beginning to build such systems.

A full digital quantum computer is a fully programmable computer that uses qubits to reliably perform digital computation. No such computer exists today. It is unlikely that one will be built anytime soon. However, building such a computer would constitute, well, a quantum leap in computing. It would allow for full scale digital brains, but as of today, it remains only a speculation.

[xlii] See Optalysys for a more detailed explanation.

Finally, biological computing is an interesting area. The term "biological computing" can refer to many things from programmable genomes to linking brains together to form a supercomputer. Only the ones that are relevant to brain-like structures are mentioned here.

There are attempts to build small chips out of neurons[88]. It is an interesting concept but is unlikely to help us build a computer that has the power of a human brain. Another approach is to link together many animal brains into a network to perform computations[89]. Such an approach holds promise, but we are currently limited by brain-computer bandwidth. Modern brain-computer interfaces connect to hundreds of neurons at once. That is far too few for high-performance computing applications. DARPA's NESD program aims to record from a million neurons and stimulate one hundred thousand of them – should it succeed, some version of using animal brains for computing purposes would become possible. The last biological computing idea worth mentioning utilizes synthetic biology, and would genetically program organisms to build nervous systems useful for specific applications. It also holds promise but we are nowhere near the understanding of genetics we would need to implement something like this.

There is another approach to building a brain-like computer that is worth mentioning: the one used by DARPA in its Cortical Processor program. It states that the first step to building a digital brain is understanding the algorithms of the cortex – only then can we build the hardware required for it. It makes sense that to design hardware to run an algorithm efficiently, it would be helpful to know what the algorithm is!

In conclusion, there are many efforts aimed at building a computer with the power of a human brain today. The most likely course of development for ABNs and other algorithms that aim to have more biologically-plausible neurons is to first test them on regular GPUs/TPUs and then implement them on a larger scale in a BrainScaleS-like digital/analog system.

9

THE LIMITS OF OUR UNDERSTANDING

If the last chapter occasionally bordered on science fiction, this chapter deals with the harsh reality of the limits of our understanding of the brain. Despite some advances in figuring out the cortical algorithm, we could not model the brain today even if we had unlimited computing power. This chapter outlines the things we would need to work on understanding next.

The first one is inhibitory cells in the cortex. There is a huge diversity of them[90]. We have not yet figured out how the different types of inhibitory neurons connect to each other and to excitatory neurons. Such a connectome is crucial for completing our understanding of the common cortical algorithm. This book presented a few conjectures as to what some of the inhibitory neurons might be necessary for, but such theories need to be confirmed. Neither this book, nor any other publication, offers a compelling theory of how inhibitory neurons work on the dendritic level. Learning algorithms for inhibitory neurons are also lacking.

The second thing that we need to learn more about is the deep layers of the cortex. Many papers divide layer 5 into layers A/B, others divide layer 5 neurons into regularly spiking and intrinsically bursting, and still others classify layer 5 cells by size. It is entirely possible that information about goals is expressed in some of those cells. Layer 6 has an even greater number of cell types, many of which are connected to parts of the brain, like the thalamus, that we don't understand very well. The thalamus is not a single sheet of neurons like a cortex, but a collection of distinctly separated parts called nuclei. Some nuclei are understood pretty well. For example, many nuclei receive sensory information and then send it to the primary sensory cortices. LGN and pulvinar do that for vision, ventral posterior and posterior medial nuclei do that for somatosensation, and the medial geniculate nucleus does that for hearing[91]. These are known as first order nuclei and are relatively well studied. However, the nuclei that receive inputs from cortical regions and send outputs to other cortical regions are known as higher order nuclei, and they are not understood well today. Higher order nuclei provide second pathways for cortico-cortical connections, but their role is unclear at this point. Given how tightly interwoven these nuclei are with the cortex, they are likely to play an important role in cortical functioning. Layer 6, in particular, cannot be understood without understanding the thalamus as well.

The hippocampus is actually understood pretty well (or so the people studying the hippocampus say). We know what the function of the organ is. We know all the different parts of it, and we have figured out the roles of the different cell types in it. That being said, we do not yet have a good algorithmic representation of the hippocampus. One of the reasons for that is that we do not yet have a good set of learning rules for the hippocampus. To obtain those, we need to further study the hippocampal neurons at the level of dendrites and synapses.

The claustrum is, perhaps, the most mysterious part of the brain (the word claustrum means "hidden away"). It is small - its volume is just 0.25% of the volume of the neocortex - yet it is highly interconnected with many parts of the neocortex. Furthermore, the claustrum seems to be strongly involved in consciousness, one of the most alluring concepts known to man. One bit of evidence for this is that if you implant an electrode into the claustrum, you can have an "on-off" switch for consciousness[92]. Another relevant fact is that salvia divinorum, a psychoactive plant, affects k-opioid receptors that are most concentrated in the claustrum[93]. This is relevant because users of the drug describe the experience as being conscious, but not having full access to the sensory stream and memories (aka the neocortex and the hippocampus). Francis Crick and Christof Koch have described the claustrum as the "conductor of consciousness"[94]. Their theory is that the claustrum synchronizes the cortex so that various conscious precepts occur on the same timeline. The salvia divinorum finding suggests that the claustrum may be even more involved in consciousness than that. As the structure is poorly studied, it may be too early to theorize further. Experimental findings describing the types of excitatory and inhibitory neurons in the claustrum, their dendritic mechanics and their connectome are necessary to advance our understanding of both the claustrum and consciousness. A better understanding of the thalamus and layer 6 would also be helpful for understanding the claustrum, as the structures are highly interconnected.

Electrical synapses received only a brief mention in this book.

Our knowledge of their function is fairly limited. We know that they are important during cortical development in early childhood – they help the newly born neurons assemble into circuits[95]. However, they may also play an important role in the adult brain. A recent study suggested that electrical synapses help neurons synchronize in the cortex[96]. Beyond that, electrical synapses in excitatory neurons of the cortex remain very poorly studied.

Another potential game changer is active information processing in axons. In this book axons are treated like passive cables – which is reminiscent of how people used to treat dendrites. Axons have traditionally been very hard to study, but new methods are allowing us to examine their inner workings. Early evidence suggests that axons possess active capabilities that help them precisely regulate the timing of spikes during high frequency firing[97]. Whether this is an auxiliary mechanism or a crucial information processing feature we do not know, nor do we know what other active mechanisms axons might possess. It is possible that they are just cables, but it is also possible that, like dendrites, they process information.

Finally, the inner workings of chemical synapses are perhaps our greatest unknown. In this book synapses are described as relatively straightforward links between neurons, which are sometimes unreliable. But once again, things are rarely straightforward in biology. A synapse has a lot of machinery that we don’t understand[98]. It is possible that this machinery introduces yet another layer of computational complexity. To those attempting to figure it out, I wish the best of luck.

AFTERWORD

HOW DO WE UNDERSTAND THE CORTEX?

At some point, we are going to understand all of the aforementioned things: synapses, axons, inhibitory cells, etc. One problem will still remain – how do we understand the neocortex? When we examine a dictionary, we can see many distinct parts, each of which looks different and has a function that is unique or close to it. When we examine a tree, we find that it has many leaves, all of which look the same and do the same thing.

The neocortex, however, is like neither of those things. It has many excitatory neurons, all of which look the same[xliii] but mean different things. It's a devilishly tricky system to study. First of all, we don't even have the methods to study it. We don't have a way to observe a meaningful part of the cortex in action. That problem is being actively worked on – IARPA, DARPA's younger sister, has a program called MICrONS, which aims to devise techniques to record neural activity at single neuron resolution from large volumes of the neocortex[99]. Let's assume that they are successful beyond our wildest expectations and we gain the ability to record firing of neu-

---

[xliii] Sure, they can be divided into classes, but each class will still have millions of neurons in it.

rons from the entire neocortex. What then? How do you begin to approach recordings from tens of billions of neurons?

Currently, most labs would try to do something like this: make a mouse perform certain activities, like running, jumping, or drinking, and look at which neurons fire during that activity. Then they would label those neurons as “running”, “jumping” or “drinking” neurons. A recent study tried to utilize such an approach to try and understand a microprocessor[100]. Unsurprisingly, the approach failed – while they were able to find “Donkey Kong” transistors and “Space Invader” transistors, such information offered very little insight into how a microprocessor actually works.

We need better questions to ask of neural data. We know that the neocortex operates in a world of sequences, so we need to look for sequences of neurons firing in different modes, including bursting and regular firing. Moreover, to understand the principles of the neocortex we need to look for a very specific thing: how sequences on short-term timescales are combined to form sequences on longer-term timescales. That question is at the center of the quest to understand the cortex.

We need better ways to ask such questions, too. While it is possible (but not very fun!) to manually analyze recordings from 30 or even 300 neurons, manual analysis will become an unrealistic way of dealing with data when we are able to record from millions and billions of neurons. New software will have to be developed to look for sequences and hierarchies of sequences in neural data. It will be machine learning software – hopefully new algorithms based on the principles of the cortex will be very helpful in this regard.

There is an issue with using machine learning to study the brain, though. Consider the Deepmind algorithm that plays Go. It can make moves that are better than moves that humans would make, but it cannot at present communicate why those moves are better. Similarly, machine learning algorithms might find patterns in neural data, but, as of now, there are no good ways for them to explain the patterns to humans. Luckily, theoretical advances will be able to help with this problem. If we think that layer 2 neurons start firing when an object is identified and persist even when the object is out of the sensory field, we can train algorithms to look for such neurons and link them to the object, assuming the algorithms track everything in the sensory field as well. If we know that layer 5 neurons cause physical actions, we can also make algorithms look for such connections. At the end of the day, if you have labels that you as a researcher understand very well, algorithms can find the neurons that match the labels. Further theoretical advances, such as understanding cortical columns will help bridge the human-machine gap in understanding even more. Though the problem of understanding the neocortex seems daunting right now, the combined advancement of better data collection tools and theoretical understanding of the cortex are likely to bring this biological marvel within our grasp sometime this century.

REFERENCES

1. The lac operon. Khan Academy 2017.
2. Menda G, Shamble PS, Nitzany EI, Golden JR, Hoy RR. Visual perception in the brain of a jumping spider. Curr Biol 2014;24(21):2580-5.
3. Petersen CC, Hahn TT, Mehta M, Grinvald A, Sakmann B. Interaction of sensory responses with spontaneous depolarization in layer 2/3 barrel cortex. Proceedings of the National Academy of Sciences 2003;100(23):13638-13643.
4. Orr AG, Hsiao EC, Wang MM, Ho K, Kim DH, Wang X, Guo W, Kang J, Yu G-Q, Adame A. Astrocytic adenosine receptor A2A and Gs-coupled signaling regulate memory. Nature neuroscience 2015;18(3):423-434.
5. Xu T, Yu X, Perlik AJ, Tobin WF, Zweig JA, Tennant K, Jones T, Zuo Y. Rapid formation and selective stabilization of synapses for enduring motor memories. Nature 2009;462(7275):915-919.
6. Mountcastle V. An organizing principle for cerebral function: the unit model and the distributed system. 1978.
7. Von Melchner L, Pallas SL, Sur M. Visual behaviour mediated by retinal projections directed to the auditory pathway. Nature 2000;404(6780):871-876.
8. Sampaio E, Maris S, Bach-y-Rita P. Brain plasticity:'visual'acuity of blind persons via the tongue. Brain research

2001;908(2):204-207.

9. Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cerebral cortex 1991;1(1):1-47.
10. Rao RP, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature neuroscience 1999;2(1):79-87.
11. Matyas F, Sreenivasan V, Marbach F, Wacongne C, Barsy B, Mateo C, Aronoff R, Petersen CC. Motor control by sensory cortex. Science 2010;330(6008):1240-1243.
12. Craddock RC, Jbabdi S, Yan C-G, Vogelstein JT, Castellanos FX, Di Martino A, Kelly C, Heberlein K, Colcombe S, Milham MP. Imaging human connectomes at the macroscale. Nat Meth 2013;10(6):524-539.
13. Spruston N. Pyramidal neurons: dendritic structure and synaptic integration. Nat Rev Neurosci 2008;9(3):206-221.
14. Shepherd GM, Stepanyants A, Bureau I, Chklovskii D, Svoboda K. Geometric and functional organization of cortical circuits. Nature neuroscience 2005;8(6):782-790.
15. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics 1943;5(4):115-133.
16. Burkitt AN. A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. Biol Cybern 2006;95(1):1-19.
17. Neumann Jv. The computer and the brain. Yale University Press; 1958. 98 p.
18. Schiller J, Major G, Koester HJ, Schiller Y. NMDA spikes

in basal dendrites of cortical pyramidal neurons. Nature 2000;404(6775):285-289.
19. Major G, Larkum ME, Schiller J. Active properties of neocortical pyramidal neuron dendrites. Annual review of neuroscience 2013;36:1-24.
20. Branco T, Clark BA, Häusser M. Dendritic discrimination of temporal input sequences in cortical neurons. Science 2010;329(5999):1671-1675.
21. Branco T, Häusser M. Synaptic integration gradients in single cortical pyramidal cell dendrites. Neuron 2011;69(5):885-892.
22. Lashley KS. In search of the engram. 1950.
23. Larkum ME, Nevian T, Sandler M, Polsky A, Schiller J. Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. Science 2009;325(5941):756-760.
24. Larkum ME, Waters J, Sakmann B, Helmchen F. Dendritic spikes in apical dendrites of neocortical layer 2/3 pyramidal neurons. Journal of Neuroscience 2007;27(34):8999-9008.
25. Larkum M. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. Trends in neurosciences 2013;36(3):141-151.
26. Polsky A, Mel B, Schiller J. Encoding and decoding bursts by NMDA spikes in basal dendrites of layer 5 pyramidal neurons. Journal of Neuroscience 2009;29(38):11891-11903.
27. Kasthuri N, Hayworth KJ, Berger DR, Schalek RL, Conchello JA, Knowles-Barley S, Lee D, Vázquez-Reina A, Kaynig V, Jones TR. Saturated reconstruction of a volume of neocor-

tex. Cell 2015;162(3):648-661.

28. Golding NL, Staff NP, Spruston N. Dendritic spikes as a mechanism for cooperative long-term potentiation. Nature 2002;418(6895):326-331.

29. Kwon H-B, Sabatini BL. Glutamate induces de novo growth of functional spines in developing cortex. Nature 2011;474(7349):100-104.

30. Chiu CQ, Lur G, Morse TM, Carnevale NT, Ellis-Davies GC, Higley MJ. Compartmentalization of GABAergic inhibition by dendritic spines. Science 2013;340(6133):759-762.

31. Miles R, Tóth K, Gulyás AI, Hájos N, Freund TF. Differences between somatic and dendritic inhibition in the hippocampus. Neuron 1996;16(4):815-823.

32. Jadi M, Polsky A, Schiller J, Mel BW. Location-dependent effects of inhibition on local spiking in pyramidal neuron dendrites. PLoS Comput Biol 2012;8(6):e1002550.

33. Hasenstaub A, Shu Y, Haider B, Kraushaar U, Duque A, McCormick DA. Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. Neuron 2005;47(3):423-435.

34. Buzsáki G, Chrobak JJ. Temporal structure in spatially organized neuronal ensembles: a role for interneuronal networks. Current Opinion in Neurobiology 1995;5(4):504-510.

35. Schrödinger E. What is life?: With mind and matter and autobiographical sketches. Cambridge University Press; 1992.

36. Ziegler JF, Curtis HW, Muhlfeld HP, Montrose CJ, Chin B. IBM experiments in soft fails in computer electronics

(1978;1994). IBM J. Res. Dev. 1996;40(1):3-18.

37. Harris KD, Shepherd GM. The neocortical circuit: themes and variations. Nature neuroscience 2015;18(2):170-181.

38. Miller KD, Pinto DJ, Simons DJ. Processing in layer 4 of the neocortical circuit: new insights from visual and somatosensory cortex. Current opinion in neurobiology 2001;11(4):488-497.

39. Coogan TA, Burkhalter A. Hierarchical organization of areas in rat visual cortex. Journal of Neuroscience 1993;13(9):3749-3772.

40. Lefort S, Tomm C, Sarria J-CF, Petersen CC. The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. Neuron 2009;61(2):301-316.

41. Constantinople CM, Bruno RM. Deep cortical layers are activated directly by thalamus. Science 2013;340(6140):1591-1594.

42. Houghton G, Tipper SP. Inhibitory mechanisms of neural and cognitive control: Applications to selective attention and sequential action. Brain and cognition 1996;30(1):20-43.

43. Hooks BM, Mao T, Gutnisky DA, Yamawaki N, Svoboda K, Shepherd GM. Organization of cortical and thalamic input to pyramidal neurons in mouse motor cortex. Journal of Neuroscience 2013;33(2):748-760.

44. Vélez-Fort M, Rousseau CV, Niedworok CJ, Wickersham IR, Rancz EA, Brown AP, Strom M, Margrie TW. The stimulus selectivity and connectivity of layer six principal cells reveals cortical microcircuits underlying visual processing.

Neuron 2014;83(6):1431-1443.

45. Lam Y-W, Sherman SM. Functional organization of the somatosensory cortical layer 6 feedback to the thalamus. Cerebral cortex 2009:bhp077.

46. Thomson AM. Neocortical layer 6, a review. Frontiers in neuroanatomy 2010;4(13).

47. Rees G. Visual attention: the thalamus at the centre? Current biology 2009;19(5):R213-R214.

48. Buschman TJ, Miller EK. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. science 2007;315(5820):1860-1862.

49. Lübke J, Roth A, Feldmeyer D, Sakmann B. Morphometric analysis of the columnar innervation domain of neurons connecting layer 4 and layer 2/3 of juvenile rat barrel cortex. Cerebral Cortex 2003;13(10):1051-1063.

50. Muckli L, De Martino F, Vizioli L, Petro LS, Smith FW, Ugurbil K, Goebel R, Yacoub E. Contextual feedback to superficial layers of V1. Current Biology 2015;25(20):2690-2695.

51. Fleidervish IA, Friedman A, Gutnick M. Slow inactivation of Na+ current and slow cumulative spike adaptation in mouse and guinea-pig neocortical neurones in slices. The Journal of Physiology 1996;493(1):83-97.

52. Silberberg G, Markram H. Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. Neuron 2007;53(5):735-746.

53. Markov NT, Vezoli J, Chameau P, Falchier A, Quilodran R, Huissoud C, Lamy C, Misery P, Giroud P, Ullman S. Anat-

omy of hierarchy: feedforward and feedback pathways in macaque visual cortex. Journal of Comparative Neurology 2014;522(1):225-259.

54. Gilbert CD, Wiesel TN. Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. Nature 1979;280(5718):120-125.
55. Little JP, Carter AG. Subcellular synaptic connectivity of layer 2 pyramidal neurons in the medial prefrontal cortex. Journal of Neuroscience 2012;32(37):12808-12819.
56. Shipp S, Adams DL, Moutoussis K, Zeki S. Feature binding in the feedback layers of area V2. Cerebral cortex 2009;19(10):2230-2239.
57. Feldmeyer D, Sakmann B. Synaptic efficacy and reliability of excitatory connections between the principal neurones of the input (layer 4) and output layer (layer 5) of the neocortex. The Journal of Physiology 2000;525(1):31-39.
58. Hattox AM, Nelson SB. Layer V neurons in mouse cortex projecting to different targets have distinct physiological properties. Journal of neurophysiology 2007;98(6):3330-3340.
59. Donoghue JP, Kitai S. A collateral pathway to the neostriatum from corticofugal neurons of the rat sensory-motor cortex: An intracellular HRP study. Journal of Comparative Neurology 1981;201(1):1-13.
60. Dorocic IP, Fürth D, Xuan Y, Johansson Y, Pozzi L, Silberberg G, Carlén M, Meletis K. A whole-brain atlas of inputs to serotonergic neurons of the dorsal and median raphe nuclei. Neuron 2014;83(3):663-678.

61. Crick F, Koch C. Consciousness and neuroscience. Essential sources in the scientific study of consciousness 2003:35-53.
62. Chagnac-Amitai Y, Luhmann HJ, Prince DA. Burst generating and regular spiking layer 5 pyramidal neurons of rat neocortex have different morphological features. Journal of Comparative Neurology 1990;296(4):598-613.
63. Hubel DH. Cortical neurobiology: a slanted historical perspective. Annual review of neuroscience 1982;5(1):363-370.
64. Mountcastle VB. The columnar organization of the neocortex. Brain 1997;120(4):701-722.
65. Horton JC, Adams DL. The cortical column: a structure without a function. Philosophical Transactions of the Royal Society of London B: Biological Sciences 2005;360(1456):837-862.
66. Meyer HS, Egger R, Guest JM, Foerster R, Reissl S, Oberlaender M. Cellular organization of cortical barrel columns is whisker-specific. Proceedings of the National Academy of Sciences 2013;110(47):19113-19118.
67. Leibowitz SF. Reciprocal hunger-regulating circuits involving alpha-and beta-adrenergic receptors located, respectively, in the ventromedial and lateral hypothalamus. Proceedings of the National Academy of Sciences 1970;67(2):1063-1070.
68. Risold P, Thompson R, Swanson L. The structural organization of connections between hypothalamus and cerebral cortex. Brain Research Reviews 1997;24(2):197-254.
69. Frank S, Kullmann S, Veit R. Food related processes in the insular cortex. Frontiers in human neuroscience 2013;7:499.

70. Vertes RP, Hoover WB, Szigeti-Buck K, Leranth C. Nucleus reuniens of the midline thalamus: link between the medial prefrontal cortex and the hippocampus. Brain research bulletin 2007;71(6):601-609.
71. Kumaran D, Maguire EA. The human hippocampus: cognitive maps or relational memory? Journal of Neuroscience 2005;25(31):7254-7259.
72. Sullivan Giovanello K, Schnyer DM, Verfaellie M. A critical role for the anterior hippocampus in relational memory: evidence from an fMRI study comparing associative and item recognition. Hippocampus 2004;14(1):5-8.
73. Goodale MA, Milner AD. Separate visual pathways for perception and action. Trends in neurosciences 1992;15(1):20-25.
74. Wilson FA, Scalaidhe SP, Goldman-Rakic PS. Dissociation of object and spatial processing domains in primate prefrontal cortex. SCIENCE-NEW YORK THEN WASHINGTON- 1993;260:1955-1955.
75. Petrof I, Viaene AN, Sherman SM. Properties of the primary somatosensory cortex projection to the primary motor cortex in the mouse. Journal of neurophysiology 2015;113(7):2400-2407.
76. Shipp S. The importance of being agranular: a comparative account of visual and motor cortex. Philosophical Transactions of the Royal Society of London B: Biological Sciences 2005;360(1456):797-814.
77. Oswald MJ, Tantirigama ML, Sonntag I, Hughes SM, Empson RM. Diversity of layer 5 projection neurons in

the mouse motor cortex. Frontiers in cellular neuroscience 2013;7:174.

78. Schall JD. Neural basis of saccade target selection. Reviews in the Neurosciences 1995;6:63-63.

79. Zhang T, Heuer HW, Britten KH. Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. Neuron 2004;42(6):993-1001.

80. Andersen RA. Multimodal integration for the representation of space in the posterior parietal cortex. Philosophical Transactions of the Royal Society of London B: Biological Sciences 1997;352(1360):1421-1428.

81. NEWITZ A. March 30. Movie written by algorithm turns out to be hilarious and intense. <http://arstechnica.com/the-multiverse/2016/06/an-ai-wrote-this-movie-and-its-strangely-moving/>. Accessed 2017 March 30.

82. Markov N, Misery P, Falchier A, Lamy C, Vezoli J, Quilodran R, Gariel M, Giroud P, Ercsey-Ravasz M, Pilaz L. Weight consistency specifies regularities of macaque cortical networks. Cerebral Cortex 2011;21(6):1254-1272.

83. Scoville WB, Milner B. Loss of recent memory after bilateral hippocampal lesions. Journal of Neurology, Neurosurgery & Psychiatry 1957;20(1):11-21.

84. Metz C. 2015 IBM's 'Rodent Brain' Chip Could Make Our Phones Hyper-Smart. Wired <https://www.wired.com/2015/08/ibms-rodent-brain-chip-make-phones-hyper-smart/>.

85. BrainScaleS. <https://brainscales.kip.uni-heidelberg.de/public/index.html>.

86. Accelerated Analog Neuromorphic Hardware Schemmel J2016.

87. Friston K. The free-energy principle: a unified brain theory? Nature Reviews Neuroscience 2010;11(2):127-138.

88. Bien-Kahn J. 2016 The Man Who's Building a Computer Made of Brains. <https://motherboard.vice.com/en_us/article/komiku-neuron-computer-agabi>.

89. Medicine D. Neuroscientists demonstrate operation of the first network of brains

(Brainets) in both primates and rodents In: Halkiotis S, editor2015.

90. Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. Interneurons of the neocortical inhibitory system. Nature reviews. Neuroscience 2004;5(10):793.

91. Sherman SM. Thalamus plays a central role in ongoing cortical functioning. Nature neuroscience 2016;16(4):533-541.

92. Koubeissi MZ, Bartolomei F, Beltagy A, Picard F. Electrical stimulation of a small brain area reversibly disrupts consciousness. Epilepsy & Behavior 2014;37:32-35.

93. Stiefel KM, Merrifield A, Holcombe AO. The claustrum's proposed role in consciousness is supported by the effect and target localization of Salvia divinorum. Frontiers in integrative neuroscience 2014;8:20.

94. Crick FC, Koch C. What is the function of the claustrum? Philosophical Transactions of the Royal Society of London B: Biological Sciences 2005;360(1458):1271-1279.

95. Yu Y-C, He S, Chen S, Fu Y, Brown KN, Yao X-H, Ma J, Gao KP, Sosinsky GE, Huang K. Preferential electrical coupling regulates neocortical lineage-dependent microcircuit as-

sembly. Nature 2012;486(7401):113-117.

96. Wang Y, Barakat A, Zhou H. Electrotonic coupling between pyramidal neurons in the neocortex. PloS one 2010;5(4):e10253.

97. Chéreau R, Saraceno GE, Angibaud J, Cattaert D, Nägerl UV. Superresolution imaging reveals activity-dependent plasticity of axon morphology linked to changes in action potential conduction velocity. Proceedings of the National Academy of Sciences 2017;114(6):1401-1406.

98. Feng W, Zhang M. Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. Nature Reviews Neuroscience 2009;10(2):87-99.

99. IARPA. Machine Intelligence from Cortical Networks (MICrONS). <https://www.iarpa.gov/index.php/research-programs/microns>.

100. Jonas E, Kording KP. Could a neuroscientist understand a microprocessor? PLOS Computational Biology 2017;13(1):e1005268.