SUMMARY

ON INTELLIGENCE

JEFF HAWKINS WITH SANDRA BLAKESLEE

QuickRead

FREE BOOK SUMMARIES

Summary of “On Intelligence” by Jeff Hawkins with Sandra Blakeslee

Written by Lea Schullery

Learn How a New Understanding of the Brain Will Lead to the Creation of Truly Intelligent Machines

Introduction

Jeff Hawkins has always been passionate about mobile computing. He is best known for starting two companies in Silicon Valley, Palm Computing and Handspring. He is also the architect of handheld computers and cellphones like the PalmPilot and the Treo. However, his interest in mobile computing is superseded by his interest in the most complex machine known to mankind: the brain. He aims to understand how the brain works, not just from a philosophical perspective, but in a detailed nuts and bolts engineering way. But his interest doesn't stop there. His desire is not only to understand what intelligence is and how the brain works but also how to build machines that work the same way. The result? A *truly intelligent machine*. Sure, we have plenty of computer programmers who have created some incredible machines, but they have failed to make them intelligent. Therefore, we can no longer ignore the differences between computers and brains.

For instance, why can a six-year-old hop gracefully across rocks while advanced robots are simply nothing more than a lumbering zombie? How are three-year-olds well on their way to mastering language while computers can't? How can you immediately tell the difference between a dog and a cat while a supercomputer cannot make the distinction at all? These are the questions that Hawkins aims to answer. Throughout *On Intelligence,* Hawkins describes a comprehensive theory of how the brain works. He refers to this theory as "real intelligence" instead of "artificial intelligence," since AI fails to address the most important part of building intelligent machines, the intelligence! So if you want to learn exactly what intelligence is and how your brain creates it, then let's get started.

Powerful Computers Doesn't Make Intelligent Computers

Most people today believe that artificial intelligence, or AI, is alive and well. With all the advancements in technology, we now have computers in our pockets. Not only are computers small enough to fit in the palm of our hands, but they are also more powerful and faster than ever before. As a result, researchers have become inspired to create a computer powerful enough to think like a human being. And in some ways, they’ve succeeded. The problem, however, is that no amount of processing power will ever compare to the intelligence of the human brain.

You see, computers are designed to do certain tasks. They don’t learn anything new, they simply store information and are unable to build upon that information in the way human brains do. The human brain, on the other hand, is capable of being creative and interpreting the world in multiple ways. It’s not limited to pre-programming and can continuously learn new things. For instance, when we speak or listen, we manipulate mental symbols called words, using well-defined rules of grammar. Similarly, when we play chess, we use mental symbols that represent the properties and locations of the various pieces. This is what makes the human brain more intelligent.

The famous computer, IBM’s Deep Blue, eventually beat world chess champion, Gary Kasparov, at his own game. But how? The human brain is more intelligent, right? Well, the success of Deep Blue was a hollow one, it didn’t beat Gary because it was smarter than a human; it won by being a million times faster than a human. An expert human player like Gary looks at the board and immediately sees what areas of play are most likely to be successful or dangerous. A computer, on the other hand, has no innate sense of what is important and explores every possible option much faster than a human. Deep Blue didn’t know anything about his opponent, the history of the game, and played chess without *understanding* chess.

Essentially, Deep Blue played chess in the same way a calculator performs arithmetic without understanding mathematics.

Successful AI programs are only good at the one particular thing they are specifically designed for. While computer programmers focus on making computers more powerful and adding more memory capacity, Hawkins understands that these practices won't make computers more *intelligent*. They still won't be able to understand the world and build upon prior knowledge the same way humans do. So Hawkins set out to understand exactly how the brain works.

Neural Networks Seem Like a Promising Solution But Are Missing One Key Component

In January of 1986, Hawkins began at UC Berkeley to begin his journey on understanding the human brain. Around this time, a new and promising approach to thinking about intelligent machines came onto the scene: neural networks. While neural networks have been around since as early as the 1960s, researchers had grown frustrated over the continuing failure of AI and were actively seeking an alternative. They found one in artificial neural networks.

Instead of programming computers, neural network researchers were more interested in learning what kinds of behaviors could be performed by hooking a bunch of neurons together, much like the human brain. Brains are made of neurons; in other words, the brain is a neural network. If neural network researchers, also known as *connectionists,* could successfully create a neural network in a computer, then surely a computer could become as intelligent as a human. Unfortunately, Hawkins quickly recognized how neural networks just simply wouldn't work.

Neural networks still wouldn't allow computers to perceive the world in the way humans do. That's because the human brain consists of one key element in intelligence: the neocortex. The neocortex is the part of the brain responsible for sensory perception and conscious thought. Almost everything we think of as intelligence - perception, language, imagination, mathematics, art, music, and planning - occurs here. The neocortex is about 2 millimeters thick and has six layers, imagine six playing cards stacked upon one another. As information from the senses travels through each layer, each one adds detailed prior knowledge to the raw sensory information.

Additionally, your neocortex is filled with nerve cells or neurons. But they are so tightly packed that no one quite knows precisely how many cells they contain. Some anatomists have estimated that a typical human neocortex contains around 30 billion neurons, but that figure could be significantly higher or lower. Those thirty billion cells, however, are you. They contain all your memories, knowledge, skills, and accumulated life experience. That thin sheet of cells sees, feels, and creates our worldview. For example, when you look at a familiar face, your brain begins to receive this information. The lower layers of the neocortex process this visual information and focus on things like basic color and contrast information. This area then feeds the information up to other areas until it reaches the higher echelons of the visual cortex that then recognizes the face you see as that of your spouse, child, or boss.

The neocortex is full of patterns. Brains are pattern machines. No matter if you see, hear, or touch something familiar, your brain goes through the same algorithm. This means that you don't need any one of your senses or a particular combination of senses to be intelligent. Helen Keller, for example, had no sight or hearing, yet she successfully learned language and became a skillful writer - better than most sighted and hearing people. She only possessed two of our main senses, yet the flexibility of the brain allowed her to perceive and understand the world the same way individuals with all five senses do.

The neocortex is incredibly fast and efficient and can connect prior knowledge that you aren't even consciously aware of. The brain knows the world through a set of senses, these senses create patterns that are sent to the neocortex and processed by the same algorithm to create a model of the world. This is what allows humans to experience the world in such a fluid and seamless way. Even more remarkably, the brain then holds those patterns in memory to access for future reference. In other words, our brains are just a growing database of memories and knowledge that we continuously build upon. Something computers simply cannot do yet.

Our Brains Use Memory to Predict Future Events

Why is that when a ball is thrown at you, you know to catch it? As someone throws a ball to you, you see it traveling toward you, and in less than a second you snatch it out of the air. This doesn't seem too hard, that is, until you try to program a robot arm to do the same. To get a robot to catch a ball, engineers and computer scientists must first try and calculate the flight of the ball to determine where it will be when it reaches the arm. This calculation requires solving several physics equations. Next, the joints of the robotic arms must be adjusted properly to move the hand into the proper position, requiring even more mathematical equations. Finally, the entire operation must be repeated multiple times for the robot to get better at processing the information necessary to catch the ball. In other words, a computer requires millions of steps to solve numerous mathematical equations to do something as simple as catching a ball. Humans, on the other hand, catch a ball simply by using memory.

When you catch a ball using memory, three things happen. First, the brain automatically recalls a memory by the sight of the ball. Second, the memory recalls a temporal sequence of muscle commands. And third, the retrieved memory is adjusted as it is recalled to accommodate the details of the moment, such as the ball's path and position of your body. The memory of how to catch a ball was not programmed into your brain; it was learned over years of repetitive practice. The memory is stored, not calculated, in your neurons. The neocortex, unlike a computer, uses stored memories to solve problems and produce behavior. In other words, it uses a sequence of patterns to predict future behavior.

For example, your memory of songs is a great example of temporal sequences in memory. Think of a song you know. How do you recall the tune? You simply cannot imagine the entire song at once; instead, you think of it in sequence. Whether you start at the beginning, middle, or end, you continue to play through the song by filling in the notes one after another.

You cannot recall the song backward, you can only recall it in the same way you learned it. The same goes for imagining your childhood home. You cannot imagine each detail at once; instead, you move through the different rooms of the house in sequence, recalling it in the way you experienced it.

These patterns allow us to predict future events. Whenever we experience something, our brain automatically begins to search for a similar prior experience. Nerve cells that became activated by a previous experience suddenly trigger again, allowing the brain to determine the appropriate reaction this time around. For example, when we put our key into the ignition, we predict the engine will begin roaring as it starts up. We can’t ever be 100 percent sure the same will happen each time, but based on experience, our brain can predict that it will. While humans can’t necessarily *see* into the future, our brains can make accurate predictions. This process is ongoing as the brain continues to encounter new experiences. But what about these patterns, sequences, and predictions make a truly intelligent computer?

The Technology for Building Intelligent Machines Already Exists

As we've seen, brains made predictions by analogy to the past. Is there such a way to make computers do the same? Could we create a new tool to do something similar, only faster, more efficiently, or more cheaply? Take a look at the telephone, for example. The telephone has evolved into a wireless voice and data communications network, allowing people from all over the planet to communicate with one another. This technology evolved from a simple telegraph that was only used to communicate important news or emergencies. This simply shows that technology can surpass our expectations. So can we build intelligent machines?

The simple answer is *yes*. But what does an intelligent machine look like? Science fiction books and movies have long introduced us to such machines, training us to believe that the intelligent machine looks like R2D2 and C3PO from *Star Wars* or Lieutenant Commander Data from *Star Trek*. We've even been introduced to limited-application robots, like smart cars and self-guided vacuum cleaners or lawnmowers. So while those vacuum cleaners are certainly becoming more common in today's households, robots like Commander Data and C3PO will remain fictional for quite some time. Here's why.

The human mind is more than just the neocortex; it is also created by the emotional systems of the old brain and by the complexity of the human body. So to be human, you would need biological machinery, not just a neocortex. Additionally, for a robot to converse like a human, it would need to live a humanlike life and have the experiences and emotions of a real human. In other words, humanlike robots are simply born out of fiction and should not be looked at for inspiration in developing genuinely intelligent machines.

Perhaps the biggest challenge faced when building intelligent machines is creating the memory. To build intelligent machines, we would need to build them with the same memory capacity as the human brain. To do this would require 8 trillion bytes of memory. A hard drive on a personal computer today has 100 million bytes, so we would need about eighty of today's hard drives to have the same amount of memory as the human cortex. The exact numbers may be a bit off, but the point is, building a computer with this amount of memory is totally doable thanks to silicon chips. Silicon chips are small, low power, and rugged, and soon they will have enough capacity to build intelligent machines.

The second problem we will have to overcome is connectivity. Real brains have a large amount of white matter that is made up of millions of axons streaming in all directions just beneath the cortical sheet. An individual cell in the cortex might connect to five or ten thousand other cells. This kind of wiring is near impossible to implement using traditional silicon chips. One possible solution, however, is the use of single fiber optic cables. A single cable can quickly transmit large amounts of data over a million conversations at once. So while real brains have axons between cells that talk to each other, we can build machines that share connections similar to the telephone system.

Once these technological challenges are solved, there is nothing preventing scientists and engineers from building genuinely intelligent systems. There are certainly the challenges of making these systems small, low cost, and low power, but nothing is standing in the way. While it took fifty years to go from room-sized computers to ones that fit in your pocket, today, we are starting with the advantage of having more technological knowledge; therefore, the transition for intelligent machines should go much faster.

Intelligent Machines Should Be Embraced Instead of Feared

So now that we know we *can* build intelligent machines, the question becomes, *should* we build intelligent machines? Think about the realm of science fiction movies that have posed various ethical issues and the possible dangers of creating highly intelligent machines. Some people still fear that we may develop intelligent robots that could escape our control and swarm the Earth, taking back what they believe is theirs. Ray Kurzweil even has a theory that nanorobots could crawl within our brains and record every synapse and connection, then report that information to a supercomputer. The supercomputer could then reconfigure itself into you! Now you'll simply be a practically immortal "software" version of yourself.

The theories and worry over intelligent machines revolting in an attempt to exterminate humankind is simply science fiction. Intelligence in machines can only be based on the brain's neocortex, machines will not suddenly become capable of having emotions and feelings. You see, human emotions like fear, desire, love, and hate aren't generated in the brain's cortex but in the old, more primitive part of the brain. Intelligence machines will not have these faculties. They will not have personal ambition and will not desire wealth, social recognition, or sensual gratification. They cannot have appetites, addictions, or mood disorders. Instead, intelligent machines cannot be humanlike unless we go out of our way to make them like that.

So why should we build them in the first place? Robots will not set out to hurt humanity but to benefit it. If a computer can have a memory that exceeds that of the human brain *and* never dies, it can eventually accumulate a vast amount of knowledge, more than any individual human. Having such machines can essentially change the way we operate and live each day. Today, we have technology that exists that heightens our senses, like night-vision goggles, radar, or the Hubble telescope. These instruments convert information that we cannot sense into visual or auditory displays that we can easily interpret. But what if intelligent machines could perceive the world through any sense found in nature as well as new senses of human design? For example, sonar, radar, and infrared vision are examples of nonhuman senses that we may want our intelligent machines to possess.

With sensory technology, intelligent machines could essentially change the way weather is predicted. Imagine weather sensors spaced every fifty or so miles across a continent. Using these sensors, an intelligent machine could collect global data and gain an even more detailed understanding of weather patterns. By attaching the sensor to a cortical-like memory, the system would learn to predict the weather in the way that humans learn to recognize objects and predict how they move over time. The intelligent weather brain could think about and understand global weather systems the way humans understand people and discover weather patterns that have long been hidden to humans.

Intelligent machines will one day surpass the human ability to think and process knowledge. However, intelligent robots will never possess human emotion so they will never try to “break the chains” of slavery in an attempt to destroy mankind. We mustn’t be afraid of what these intelligent machines might do; instead, we must embrace them as they will provide an inconceivable value to mankind.

Final Summary

Astronomer Carl Sagan once said that understanding something does not diminish its wonder and mystery. You see, many people fear that scientific understanding means a trade-off with wonder, they believe knowledge simply sucks out the flavor and color of life. But with understanding, we create a universe that is even more colorful and mysterious. So understanding how our brains work doesn’t diminish the wonder and mystery of the universe, our lives, or our future. Instead, our amazement for the world around us deepens as we apply this knowledge to understanding ourselves and building intelligent machines. Therefore, learning how the brain works and how to build intelligent machines only seems like the next logical step for humanity.