Quote from the book: „If this is true, we could say the more you know, the less you remember“.

This sentence probably had the most influence on me from the entire book. The first time I read it I literally dropped the book and stopped to think about it. It just forced me to think about this hypothesis and reason whether this implication is valid or not. I find the interaction between remembering specific facts and just understanding the structure just fascinating. In everyday life people often refer to a smart person as to someone who “knows a lot” (remembers a lot of facts). But the odd thing is that with computers the “smartness” is defined entirely differently. With an intelligent or smart computer system we do not care about the factual data, we just want it to understand and make sense out of it. Computer systems can access a lot of data but they fail to see correlation between it. IBMs supercomputer called Watson managed to beat the world’s best Jeopardy players yet the system cannot be called intelligent. Watson uses several algorithms to process the questions and find a good match from its database, but in the end of the day it does not understand the question nor is able to predict anything outside of its factual knowledge.

From the one hand people keep telling that learning is good and we should learn our entire life. There are even some special study programs, like lifelong learning, to provoke old people to learn something new in school. But on the other hand factual memory is considered as the metric to measure our intelligence. So in a sense this quote form the books raises the question if maybe learning could have drawback in the long term.

To answer this question I started to find counter and supporting arguments for the implications followed from the initial quote.

First off, the initial argument seems a bit unrealistic if we only consider the architecture of the brain. If the brain actually works like described by On Intelligence then in order for the new patterns not to reach top of cortical pyramid – the hippocampus, some lower regions of the neocortex have to recognize the pattern as „already known“. In this case the data is actually already stored and if we cannot remember it later the only explanation would be a “missing label”. The data is not accessible from higher levels of neocortex because there is no pointer assigned to this specific pattern. So what could cause this failure to label lower level patterns?

One explanation for the failure to assign new labels could be that it is caused by the overall slowing of the learning process. As you grow older less and less patterns are interpreted as new and therefore reach the hippocampus. So maybe the ability for hippocampus to propagate down and label the new patterns in neocortex just degrades over time due inactivity. In younger age brain meets constantly new patterns and they are propagated high into neocortex as errors. So the brain is used to learning new patterns even on high levels.
Another explanation for the reduced ability to memorize things in old age could also be explained by another phenomenon. It could be the case that after a long learning period brain starts to filter the important from unimportant more efficiently. So brain could just discard the new pattern once it decides that the data at hand is irrelevant or probably useless in the future. Or maybe the pattern itself is stored but not pointer is assigned to it on higher levels of neocortex.

On the other hand some people remember similar patterns (dates, plot of movies, routines of basic actions etc.) very well, even in old age. It is possible that it has something to do with other aspects of the brain. Maybe some genome determines the production rate of hormone X which in turn could control the synapse formation rate (similar to learning rate used in optimization algorithms). So the ability to remember specific patterns even if it is already stored in the invariant representation could be explained by 2 variables. First, how constant has been the learning. If the number patterns reaching hippocampus has been decreasing over time then the production of this specific hormone X could slow down. Second, the initial synapse creation hormone activity could be determined by the individual’s genome.

Even if learning would degrade our ability to remember specific occurrences of already known patterns I think learning would still be a beneficial thing to do. Aiding human brain with remembering facts is probably a lot easier thing to do than to help brain recognizing even more complicated structures in the data (making the human smarter). As no one has managed to create a truly intelligent artificial system we do not have even conceptual models how to increase ones intelligence. With saving/remembering data we have a bit better situation. First off we could just write our memories down. If we think a bit more fictionally we could imagine a system to allow humans to communicate with computers and access its memory for saving and reading data.

It is estimated that for example vision channel has around 2million axons. If we take into account that neurons reaction time is limited to around 5ms we can calculate that human vision is able to transfer $2000000\text{bits}*(1s/0.005s) = 400000000 \text{ bits of information per second (400MBits per second)}$. With that speed you can probably transfer more information than human brain is able to process. Clearly we would not want to replace our vision with the ability to communicate with computers but this gives us the intuition of how much information is flowing into the brain.

I would guess that actually around 16 connections would be enough for us to communicate with computers (8bits to transfer data into the brain and 8bits to transfer the queries out of the brain). $1s/0.005s*8\text{bit} = 1600 \text{ bits of data per second}$. That is, depending on the encoding, around 100-150 characters per second - faster than most people can perceive data by reading. The only difficult part would be to teach the brain how to use this new channel to communicate information and access the memory. For communication we could imagine for example SQL like language. So in general the implementations to aid humans with factual data are a lot more realistic than aiding humans in finding more complex structures in data.

So in the end I concluded that the link between bad memory and learning is pretty saddle and the reduction in ability to label familiar patterns is probably worth less than the ability to recognize higher abstraction level relations in the data.
When designing an intelligent system I would consider the following aspects:

- **Invariant representation of data**

  When the domain where our system needs to operate is fixed and the data variation is not very big we could just save entire data examples. But that is mostly not the case with intelligent systems - they need to be able to adapt and recognize a lot of variation of the same data. In this case it makes sense to save a lot of descriptions of low level objects, so they can be used as building blocks to construct more complicated objects. We can build the entire power set of the low level objects with very small additional memory cost. And those subsets can be labeled and used again like a distinct object. This tree like structure requires a lot of memory space in the beginning because we need to save all the possible lowest level objects, but once we have the basic components we can build many new complex structures with small saving costs. By doing this we would allow a lot more flexible memory and can avoid storing the same data multiple times.

- **Different hardware**

  I would consider using different hardware to build the intelligent system. Our currently well-known silicon based processing chips are not very good for modeling the activities in brain. While our processors work in fixed binary mode, on or off, the signals sent between neurons are more fluid in nature - analog signals would reflect the situation a lot more accurately. Even the artificial neural network algorithm, used for machine learning purposes, uses continues values to predict the activation of nodes representing neurons. So it takes a lot of computational power to calculate the activations.

  Actually few weeks ago MIT (Massachusetts Institute of Technology) announced that they have created a computer chip that models how neurons communicate with each other at synapses. This hardware looks promising and could help neuroscientists to figure out more about our brains.

- **Modelling memory loss**

  In a sense forgetting things allows us the discard the unimportant facts and focus only on the important data. As memory size limits will be a lot weaker on the artifical intelligence systems fully discarding the gained data is useless. But we need a system to somehow classify the data from important to important. So in the normal working process the intelligent system would only consider the data what has been marked as important. If it failes to come up with a good prediction it could start looking into all the available data it has. This implementation idea could speed up the normal working process and make sure usually the predictions are based only on the important data. If finding a good solution seems to be impossible we can still access the irrelevant information in a hope to notice something new and come up with a better estimate.
"On Intelligence" – review and ideas
by Artem Kaliuk

The brain is wider than the sky.

Emily Dickinson

Since the beginning of the XX century many sci-fiction authors have used the idea of intelligence prevailing outside of the human body. Awareness in technical systems was put onto the pages of books and journals in both encouraging and frightening colors. This trend in literature could not be omitted by the research and industry. In the late 1940s Alan Turing has offered a simple approach to define the existence proof of intelligence – if the computer is able to trick the human observer in such way, that the latter thinks he interrogates with another human, then the computer is thought to be intelligent. The Universal Turing Machine, which was thought to be a black box with a set of instructions being able to pass the test, was taken by the Turing’s successors as a model for developing smart computers. As many years of expertise have shown, neither this approach, nor the methods offered by the researches of neural networks, have made us closer to creation of intelligent systems. The main issue here is that behavior is just a sign of intelligence, but not the definition of it.

The most impressive idea – hierarchy, flexibility and prediction

Jeff Hawkins in his book "On Intelligence" offers a new way of thinking about the smart systems' design. Giving the Chinese Room thought experiment as an example, he shows that the trend of behaviorism (adopted by the earlier AI researchers) was yet misleading as the AI approach based on this philosophy simply could not provide flexibility and, which is even more important, understanding when solving the tasks. Even though machine performs elementary operations much faster than the brain, it takes a human just fractions of a second to recognize a face of a friend, while the computer may easily fail the task. The problem lies in how we see the design of the system's core. The most amazing thing about our brain (particularly cortex) is its complex hierarchical structure. It is built of several layers and each of these layers represents some level of abstraction. Hierarchy of abstraction layers is nothing new for the world of science and technology - just think of building a code written in a high-level programming language – first it is compiled into a code of a lower abstraction layer and this “downflow” compilation follows unless we obtain the machine code. In the same way we can represent the flow of information representation from the highest level to the lowest one. An important thing is that each layer of a brain stores the information in the form of patterns or sequences of patterns. The patterns are generalized when going up the hierarchy. This is how we are able to make conclusions about the objects which we see, touch, hear or smell based on the simple sensory input and the experience stored in memory. The long-term studies of the cortex have shown that it is speckled with interconnections between the layers. The numerous feedback connections between the neurons in each level create the phenomena called auto-associative memories – this term refers to all memories being able to retrieve the pattern from only a tiny sample of itself. Even though auto-associative memory was implemented in some neural networks, such systems were still vulnerable to variations in
the input. On the other side, the human cortex is able to fire the same cells in a higher regions even when the input patterns below change over the time. This is the case when we are still able to recognize a friend’s face every time regardless of its position, brightness, etc. The invariant representations in higher layers of cortex is what allows us to overcome the machines in numerous tasks.

Another important observation is that even though traditionally each type of information (visual, auditory, sensory) is processed in a separate region of a cortex, all the regions of a cortex are performing the same operations and the difference between them is only in the way the are interconnected to each other and other parts of the brain. Cortex is extremely “plastic” - it can rewire itself and learn to process auditory information in a visual region and vice versa. The flexibility of the brain was proved by the author in several examples. I can not keep myself from mentioning the experiment with a blind person, who learned to obtain the visual information taken from the camera by perceiving the inputs from the tongue-mounted device. So it really does not matter where the inputs come from – what matters is how these input patterns correlate over the time.

The third most impressive part (basically, the essence of intelligence and the main idea around which “On Intelligence” is built) is the ability of our brain to predict. This ability is enabled by a specific hierarchy of the cortex and complex interconnections of neurons in different layers of the cortex. Prediction allows us to retrieve the patterns from the higher regions based on the frequently observed temporal patterns in lower regions. It is done by propagation of a “name” of the sequence - this eliminates the details in higher levels and allows to learn sequences of higher level of abstraction. Every second we use prediction – even if we do not realize that. Prediction is not an entirely cortical process – it also involves the parts of an “old” brain such as thalamus and hippocampus. We use prediction when solving a mathematical problem, we use prediction when humming a melody; right now I am using prediction in order to type a text which will make sense to everybody. Prediction is also what our creativity is based on.

When I started reading about prediction, I thought again that the idea of prediction has already been used. As a rough example I will give a modern wireless network. The data is transmitted in the form of a modulated electrical signal. At the receiver this signal is corrupted by the noise – imagine this as a correspondence to a varying input of a lower cortex layer. The signal is then demodulated using a method based on some probabilistic approach (ML, MAP). All those approaches make use of the likelihood function in order to retrieve the data. The likelihood function can be thought of as a probability of getting a certain value of a parameter given a set of observed outcomes, which we can treat as some representation of experience or memory. Probability function of some variable based on a set of past observations is what the so-called Bayesian networks (which in recent years already became a core of modern probabilistic robotics) are built upon. Even though the Bayesian networks’ approach is alone a very powerful tool, the memory-prediction framework created by Mr. Hawkins allows us to understand how exactly our brain works. And only the full understanding of the mechanisms can help us to build indeed intelligent machines.

Some of my colleagues might argue that prediction should be the first choice to state as the main idea of the book. I would agree on that – according to author, prediction is the cornerstone of intelligence. However, if we recall the three properties of cortical memory which enable prediction (and these are: 1) storing sequences; 2) auto-associative recall;
3) invariant representations), we would see that all those properties come into play only due to the specific structure of cortex and interconnection of neurons. Thus, I would still insist on the peculiarity of the brain's architecture to be something that enables us to make predictions, hence, to make us intelligent.

*How does this book inspire us to build intelligent systems?*

This question addresses to the last section of the book. As it was mentioned, there are still many disputes on how soon we will be able to create something in silicon which will learn, infer and, most importantly, predict in the same way as humans do. Although the research is yet far from conclusion, we can already say that it is not about resources which prevent us from creation of a first smart system. There are still many questions on how exactly the neurons are connected within the cortex and externally, what is the mechanism of invariant memory creation and so on. But once we get the framework completed, we can hopefully launch a new era of technical progress.

The future intelligence will most probably have nothing to do with a popularized image of androids. The new systems could be box- or rack-like, mostly hidden from our field of view, but performing complex tasks using both semiconductor-driven computational power and human-essential ability to learn, analyze and predict. After adopting the structure of a brain, the systems will inherit its flexibility, so enabling us to use any type of sensors to push the data first to the lower parts of cyber-cortex with the following propagation, generalization and association of patterns on the way to upper parts. The freedom in choice of sensors may lead us to explore different environments and conditions with the same ease as we daily explore our rooms and streets of our cities. In the same way as humans do, computers will need to spend time on learning. But this time will be significantly smaller as compared to the one needed by people due to the higher processing speed, better reliability of the physical memory and intentionally reduced functionality (as we will most probably not need to make the weather monitoring system learn Shakespeare's sonnets or cook Chinese food). By reducing the functionality I simply mean decrease in the area of the cyber-cortex.

In recent years the ethical part of the progress has become a real issue. Jeff Hawkins provides us with examples of technology being a scare factor for people in XIX and XX centuries. The situation did not change much – today the privacy aspect is often raised by the world community with respect to technical systems. Also, in the same way as a computer became a necessary and ubiquitous device, smart computers may become available on the mass market in some years after being introduced. According to the author, the system is unlikely to be hostile and harmful as long as it: a) will not have an emotional component (typical for people and highly dependent on the environment in which one exists); b) will be trained for ad-hoc tasks only. Though, to my opinion, the safety of is highly dependent on how humans would like to use the technical intelligence. The affordability of such a computer can cause many unforeseen problems (imagine an intelligent spam bot, the potential harm it may cause and possible precaution measures which might confuse the normal users).

The evolution has been bringing changes into all the creatures on our planet. To step up on the next stage of evolution, the mankind has to wait another couple millions of years. With intelligent systems we will be able to speed up the progress in science and also develop our own abilities by supplementing our environment with real cognitive systems. Even despite many drawbacks, the end justifies the means.
1 Introduction

*On Intelligence* by Jeff Hawkins of Palm Computing served as the opening book for the seminar Brain, Mind, and Cognition, which seeks to examine concepts of intelligence their implications on the future of computing through the reading of several recent popular science books. Through his work in computing, Hawkins claims that current computer science has been going astray when trying to create artificial intelligence, relying on complex designs which conform to existing fundamentals of computation but ignore nature’s example, the human neocortex, which it is attempting to replicate. As an attempt to unify the many assorted sub-theories of neurobiology, Hawkins spends the first half of the book outlying the basic principles he believes summarize what we call *intelligence* and providing basic observations about the brain. After laying out these foundations, the book climaxes with a basic but unified hypothesis explaining much of the physical workings of the brain itself, followed by a denouement explaining how he believes we can apply these concepts in the development of machines which can learn and think similarly to how humans do.

2 Main Point of *On Intelligence*

Hawkins spends a good deal of the book going into details and clarifications which somewhat detract from his primary thesis, such as his emphasis on the six layers in the brain. One does, however, notice a few concepts coming up over and over again throughout the book: Hierarchy, Invariance, and Prediction.

2.1 Hierarchy

A key point that is often ignored, according to Hawkins, is the hierarchy present throughout the brain. While AI largely ignores this, and artificial neural networks largely relies on a simple three-layer input-processing-output hierarchy, the brain uses a deeply-nested structure of simple pattern matchers whose emergent behavior is what we call intelligence. Basically, these individual components fire when they detect a pattern, and then groups of firings trigger higher-level patterns. Most importantly, recognized patterns provide feedback to associated parts of the brain, acting as a secondary input to assist in recognition of an ambiguous pattern, similar to belief propagation used...
to decode LDPC codes. In short, Hawkins asserts that distributed pattern matching over a large domain serves as the foundation upon which intelligent thought is built on. How this is stored, however, is the second key point:

### 2.2 Invariance

The brain does not care about details in its pattern matching, but instead focuses on general concepts which may appear in a variety of forms. In fact, Plato’s concept of Forms succinctly summarizes this, in which many different dogs/tables/etc. are all recognized as members of their respective class, even though there may be a huge difference between individual elements of these groups. For example, you have probably noticed the change in this section to a sans serif font, but despite the fact that the letters are not the same as in the rest of the document, you will have no trouble reading the text. This is because an A in Times and an A in Helvetica are both easily recognized as the same thing by the reader, and the specifics of the instantiation are ignored as unimportant details. Likewise, I could describe the letter as an upside-down V with a horizontal line through it, and you would understand I meant an A. Hawkins goes through a few further examples, citing songs in different keys and written vs. spoken text, but the concept is always the same: grouping information by extracting the important details is the key to both how the brain can hold such a large amount of information and also how patterns can easily be exploited for prediction.

### 2.3 Prediction

Pattern matching and grouping is all well and good, but not useful on their own. Rather, they must be coupled with a prediction mechanism. Intelligence, from Hawkins’ view, can be summarized as the ability to recognize patterns and predict what else would reasonably fit that pattern. Prediction is then useful for two things: detection and consequences. The first, detection, has to do with the sheer volume of information the brain needs to sort through. If I can guess that I will see Y after doing X and it does happen, it’s probably not very relevant for me to make note of Y. However, when these predictions fail, one will take notice, like the font example in the previous section. As Shannon eloquently described it, information is simply uncertainty, so by discarding things I expect, I don’t have to spend all my time analyzing unimportant details. The prediction framework serves as the arbiter of input processing.

The second use, consequences, is the use of prediction to determine reasonable courses of action. The brain must use its current model of its surrounds along with current inputs to figure out what will happen soon, even if it has never been experienced before. For example, when we saw Jeff Goldbloom watch his water start shaking in Jurassic Park, we all knew a T-Rex was about to make the party its lunch; nobody had to tell us that, and even though the T-Rex hasn’t even appeared in the movie yet and the individual bits of information on their own don’t lead to much in the way of prediction, the brain is easily able to combine a few bits of information together to know exactly what will happen.

These three concepts - hierarchy, invariance, and prediction - are the keys to understanding the cognitive process. The hierarchical structure allows fast processing of information through distributed pattern matching, the invariant representations allow simple, error-tolerant matches of
a wide variety of information, and from these matched patterns one can guess what else reasonably fits into the current situation and use this to quickly determine relevance and also predict future actions.

3 How to Build Intelligent Machines

To build intelligent machines, it seems that error-tolerant pattern matching should be the first area of research. Hawkins’ book implies that rather than using a big, beefy CPU, many simple yes/no binary conclusions on matches over a large amount of inputs, both from normal sensory input and from conclusions of other match questions, are the way to go. This kind of structure could be implemented on FPGAs serving as large customizable logic gates, although even better would be FPGAs which could reprogram themselves on the fly. Anything which can do many, many simple pattern matches, easily feed back outputs as inputs, and ideally ‘learn’ how to make better matches based on past experience would be the main idea.

Of course, this technology is at least somewhat available already. Perhaps the holdup, then, is the sheer magnitude of connectedness required. As anyone working in VLSI knows, connections inside a chip are a huge obstacle to implementation, often taking up much more space than the silicon itself. Although the brain uses inefficient point-to-point connections rather than shared buses, the neurons have the advantage that they are essentially electrically isolated from each other, so crossings are not a problem. Perhaps examination into applying these point-to-point connections could be fruitful.

Connections will be difficult, but also data storage is a bit tricky. Hawkins points out the error-tolerance of the memory provided through invariant representations. I’ve heard that similar work is going on in image compression fields, such as not storing information for grass in a photo, but rather noting a few basic properties and recreating a fitting representation of the grass during decompression. After all, most of the textures in the photo are only as relevant as saying “There’s grass here and concrete there and a partially cloudy sky above”, and the actual positions of individual blades of grass, cracks in the concrete, and clouds in the sky are usually quite irrelevant. It would seem that applying these same concepts to data storage outside of the fields of image compression would be a sensible path to follow.

So, in short, it appears to me that Hawkins’ book tells us we should focus on simple, highly-parallel processing units with high degrees of connectivity, and also on lossy compression schemes.

4 Conclusion

Ignoring irrelevant details and making general assumptions and predictions from basic associations is the general framework for intelligence we can derive from the book. Hierarchical associations of objects are simple to implement and allow general predictions, and these predictions are, according to Hawkins, the core of intelligence.
First of all I would like to point out that reading this book was very exciting to me and that it is very hard to pick out the most interesting thought of the book. Therefore I will state those ideas which were impressive and outstanding to me and give a brief description why they were of personal interest.

The first time I was astonished was when I read about the proof that computers were not intelligent. In 1980 John Searle conducted a thought experience, called the Chinese Room. Basically it was a closed room with an English, non-Chinese, speaking person inside. This person was sitting at a table and he was featured with a book about how to manipulate, sort and compare Chinese characters. Someone outside then committed a story, written on a piece of paper and with Chinese characters, to the person inside the room. The story was concluded with some text related questions, written in Chinese as well. The man then started to read and apply the instructions in the book. After a while he was finished and returned his answers. A Chinese outside read them and was wondering how exact they were and supposed the person inside to be intelligent.

The conclusion, however, was that although the person inside actually was not able to speak or write Chinese at all; with the set of instructions in the book he could solve the given problem. Now recollect how computers work. Inside they have a CPU which is simply executing defined instructions. Shifting million bits and bytes per second without really knowing what they are doing. Nevertheless, after finishing the calculations the results are correct, exactly as those from the person inside the Chinese Room. I personally think that this is a very nice experiment in order to disprove the common belief that computers are smart devices. They might seem to be, but in fact it is an intelligent person who tells the computer what to do.

The next thrilling topic was about the method for displaying visual patterns on a human tongue. I had to read the relevant page a couple of times until I could believe what it exactly was about. In brief, a subject was equipped with a small camera sensor, placed on his forehead, and a chip attached to his tongue. The visual image then was converted to visual patterns and finally processed to tiny pressure points on his tongue. After a while the proband was able to grab a can, later he could distinguish whether a door is opened and even noted signs on the door. “Images initially experienced as sensations on the tongue were soon experienced as images in space” (Hawkins 2004: 61) in my opinion reveals perfectly what the main approach of the idea was.

Later that day I was out with my friends and I told them what I have recently read in this book and how impressed I was about the mentioned method. No one really could or was willing to believe my story. Despite I illustrated the thought as far is could they were refusing it. Some of them claimed that a tongue only can handle the sense of taste, but is not
able to perceive patterns or even images. This claim though led the discussion to a very radical idea about how our brain processes exterior sensations.

The idea was introduced by Mountcastle in the paper “An Organizing Cerebral Function”. It suggested that the cortex handles signals in a universal way, which means that the way how information is processed is independent from the type of sensory. They sensory only differ in the manner of collecting patterns. At this point I had to introduce the theory that all their brain knows are patterns, from perceptions to knowledge. This may sound absurd at the beginning but after a while the different concepts perfectly fit together. Explaining the distinction between spatial patterns for vision and temporal pattern for hear and touch revealed many light bulb moments. Anyway, I stopped the discussion, offered to lend out my book and suggested to continue by the time they read it ;)

The thing I liked most about reading about the pattern issue was that I was presented like a big puzzle. Everything seemed to be so obvious and logical, complemented with the idea that patterns are the fundamental currency of intelligence. In particular the thought that our brain doesn’t compute results; rather it uses stored memories, sequences of patterns in a hierarchical invariant form, to solve problems and produce behaviour is amazing. This makes it difficult, for example, to say the alphabet backward as you normally have only stored the forward sequence. On the other hand, for instance, you constantly complete patterns by filling in missing words when you are listening to someone.

Finally, this brings us to the most important idea of the book: What we perceive is a combination of what we sense and of our brains memory-derived prediction. I loved the following example: “Have you ever missed a step on a flight of stairs? Then you know how fast you react when a stored prediction is not met with the reality. Even though the foot doesn’t feel anything but as it lowers and passes through the anticipated stair tread you know you are in trouble” (Hawkins 2004: 92).

This allows the more sophisticated point of view: “Intelligence is measured by the capacity to remember and predict patterns in the world, including language, mathematics, physical properties of objects and social situations.” (Hawkins 2004: 96).

The second part of this essay is on the gained inspiration for building intelligent systems. I would like to start with the definition of intelligence. In the early days the famous Turing test was introduced. It says that “… if a computer can fool a human interrogator into thinking that it too is a person, then by definition the person must be intelligent” (Hawkins 2004: 14). Later on in the book it is redefined as “prediction and not behaviour is the proof of intelligence”.

What I personally think is that the general acceptance that an intelligent machine must have human behaviour and/or human looking is very old fashioned. Thus I support the following statement: “Robots are a concept born of the industrial revolution refined by fiction. We should not look to them for inspiration in developing genuinely intelligent machines.” (Hawkins 2004: 208).
Further, don’t limit yourself when it comes to the shape of the devices. It is not essential for an intelligent machine, that the sensory is directly attached. Thus the physical embodiment can be adapted to the desired field of application. In my opinion this is nicely emphasised in: “The strongest application of intelligent machines will be where hum intellect has difficulty are in which our sense are inadequate, or in activities we find boring.” (Hawkins 2004: 216).

Throughout the book a general approach how to build intelligent machines was formed. First you need a set of sensory to extract patterns from the world. Then you need a memory system in a hierarchical fashion which interacts with the sensory and acts cortex-like. The most challenging thing here is the hierarchical memory system. A reverse engineering of the full cortex capacity would require roughly 8 trillion bytes of memory. Although this would be feasible under laboratory conditions, for many applications much less is actually needed. Furthermore, an individual cell within the cortex is connected with up to ten thousand other cells. This parallel architecture provides a vast connectivity between them. Something that is almost impossible to realize with present silicon manufacturing techniques. Nevertheless, since electrical wires transmit data much faster than the axons of neurons, we can create an adequate framework by reusing existing wires and hence increasing the throughput.

Additionally, I think that you should forget about the idea of creating intelligence with a straightforward sensor-motor approach. If you take into account that our brain is steadily checking his nested world model against the actual perception it is essential to add an auto-associative function to the memory block. This provides the machine with the required feedback and the possibility of predicting things.

Let’s continue with our example. With this physical framework the intelligent machine, then has to learn, such as children do. The system will build up its world model as it is perceived with its sensory system. Important is that, for doing so, no set of rules, databases or any coded facts, as they are widely known from the AI, are required.

From my point of view it is fundamental to consider that our brain doesn’t differ between perception of what we see or hear. The processing unit is universal and the algorithm can handle everything it was taught before. So what I want to point out is that if you are creating a hand writing recognition machine, for instance, you should not only focus on the particular recognition algorithm but also try to think how our brain solves this issue in a more general way.

To sum up, there is much more I could write about how this book helps us to create intelligent machines. Although I can’t fit everything on 3 pages, the description of creativity: “Creativity is mixing and matching patterns of everything you’ve ever experience or come to known in your lifetime” (Hawkins 2004: 187) proves that I obtained many new ideas and that I created reams of new patterns while reading this book 😊

ON INTELLIGENCE

Eric Lambers

With no personal psychology or intelligent systems background myself this was a very interesting read covering broad aspects of how the brain functions. By far and beyond the most interesting concept that he raised was the concept of prediction by the brain. To comprehend that every single second of daily life your brain is constantly making predictions of what will happen next, who will appear next, what you will hear next – it is quite amazing to think of the mind in such a way. To be honest the book makes it sound quite simple to build a robot/mechanical brain, yet it clearly hasn’t been done before or at least presented to the public! In some ways I find it hard to accept Hawkins’ view given he says it so surely at times and yet I find it hard to believe that there is no significant part of science following in his footsteps and building a brain!

An easy way that I can think of when I subconsciously (but obviously) predict is listening to music. I also have very little background in music (except for singing in the shower and listening to the radio) to understand how chords and transitions work. Yet when I hear a popular song on the radio, I can predict what the next chord or movement will be, when the lead singer will sing and even after just hearing one chorus and verse, I am able to predict most of the rest of the song up to a couple of bars ahead. My brain was predicting the next beat based on a preconceived notion somewhere in my head that the song will follow certain rules and patterns, like many other songs I have heard before. I have never thought of it that way. I have noticed that I can predict chord movements and tones [I guess you would call it something similar] even though I have not studied music and the architecture of music and what sounds good to the human ear.

It is hard to get my head around what he proposes, but for the brain to predict every single thing that occurs to me highlights the sophisticated, and efficient make up of the brain throughout evolution. The sequential layering system of the cortex enable this prediction, with the segmented hierarchical system (as in there are more layers in the layer below) appears to be a reasonable explanation to me. Clearly Hawkins has a high level of knowledge on the matter and has done experiments and made many connections as to how the cortex is made up and how it may potentially function. I am still very unclear how the brain remembers (as in physically, where that information is stored and how it is recovered). I do not know if it is possible to model something else that works in a similar fashion to the brain, but the day that it happens, if ever will be a revolutionary day in history. Brains, I believe, have an amazing memory. People always say “oh its up there I’m just trying to remember...” I agree that memory is never lost, just archived as it is not used frequently, until it is needed again.

I have never believed that it would ever be possible to build truly intelligent machines/systems, as it is simply far too complex. I have never even for one second that that any one would ever get to the bottom of how it works, let alone reconstruct it on a small scale! Hawkins mentions that people say it will
be possible once computers are fast enough, but he goes on to quash that argument by stating that computer transistors are millions of times faster than the brain's neurons. But I fail to see how hard it would be for a computer to 'see' and recognize objects. Obviously a baby does not recognize what anything is when it is born, but it rapidly learns and creates associations and learns patterns, well before it can walk or talk. So why can robots not remember, or predict? I think that the prediction model is quite important and decisive for building intelligent systems, and furthermore to recognize patterns in order to predict. Thinking about what I have read has made me think about how I behave and how I see things everyday, and I agree that the brain must make predictions. How else do I know how far to put my foot in front of me before it hits the ground without looking or even consciously thinking? The ability to predict is far more powerful than the ability to recognize, which is the current train of thought in building intelligent systems. For a robot to recognize a stair it must then compute distances and speeds in order to step with the next foot in the correct place. It is not quite as simple as the "less than 100 steps" that Hawkins refers to. Even for an intelligent building system, to recognize that someone is cold, or to recognize that something is wrong, it cannot simply predict what should be there and notice the differences. I agree with Hawkins that it is far easier for me to say what is wrong than what is right if my door at home is altered. But consider how long it takes a human to recognize these differences. As I mentioned before a child would not recognize these differences until (I assume) being at least 3, let alone would a child think something is wrong, only something is different. However an adult would acknowledge that something as arbitrary as the location of a doorknob had been changed, when the prediction in his brain is not met. So it takes many years for a human to learn these things; and what I think is that we are being very tough on robotic and system engineers as I think that if a robot with memory and recognition skills was created, it would still take many years for it to be anywhere near as smart as a human.

Considering it would be possible for a robot to learn, to remember, to predict and to recognize. It would only take a few robots to "grow up" before that memory could be transferred to newly made robots. It is like software that is developed not by keystrokes, but over years of a robot observing and learning. We all know that computers can remember when instructed to save a word document, or to make a reminder alarm. But it is still only simple automated commands like auto save, or when a user inputs commands. It begs the question – when will it, if ever, be possible for a computer to learn and predict concurrently? Imagine a day when a robot would be able to watch how humans behave, recognize if spilling a glass of milk on the floor is a bad thing; and then the next time a similar incident happens (invariant representations as well), if a glass of tea is spilt, a robot instantly realizes (before seeing a human move to clean the mess) that it should help clean the mess. Just to simply be able to learn and be intelligent as humans are. Not just behavior but as Hawkins explains, prediction and learning – and not just exact objects, but invariant representations. I do think that one day it will be possible and each household who is affluent enough will own a robot that is a new companion as a man's best friend – but somehow I think I will still prefer a dog!
With no programming experience, I am quite blind. However the layering system or rather approach, presented by Hawkins, I believe could be possible. It would have slow beginnings, but given the speed of computer processors, it may be possible to firstly classify that an object is NOT something. For example, a robot sees an object with 4 wheels. It is instantly not a motorbike, and all motorbike associations can be ignored. Thinking of all programmed options with 4 wheels, simple yes/no commands might be able to be resolved to find out what it truly is. There is this game called “20 questions”. You think of an object [tell your friend next to you], and then start the game. It is all done by a small handheld device much the same size and shape as an apple. After 20 “intuitive” questions, the small device usually has a very good chance of guessing the object. So after 20 Y/N questions, I found that when it was unspecialized objects, the game guessed correctly the object that I had in mind. Ranging from a book to things such as bottles, many thousand different objects must be there, with a checklist of yes/no answers for a variety of different questions. Obviously this is limited as it still takes sometime between questions, but this was 5 years ago. If there was any way to learn the yes/no answers, like I proposed earlier, it would in theory lead to smart robots.

The question still remains: would it ever be truly intelligent? How would it predict? There can be so many different objects in any given space at any given time, how would it be able to tell that a cat for example, is an object that can move? I personally think that it would only be possible if it learnt, and remembered patterns.

Just a quick comment on the ethics of developing a truly intelligent robot: I think that they would be extremely helpful and even beneficial to society. Intelligent machines would possibly replace mundane jobs that few people like. However it is not long before one realizes the detrimental effects of such a replacement of uneducated cleaners to ‘intelligent’ robots. It would surely result in losses of jobs, for people in society who struggle to find any other jobs. I think that it is a conflict of wanting scientific advancement, however at the potential cost of the livelihoods of uneducated people.
Thoughts on “On Intelligence” by Jeff Hawkins
by Marius Loch

In his book “On Intelligence” Jeff Hawkins sets out to explain how the human brain works. He claims that all previous attempts to develop artificial intelligence are condemned to fail, because they don’t consider how the brain works. That’s why he started his research and came up with a theory describing the brain (or rather the part responsible for intelligence) as a memory-based prediction-machine.

On idea that came up while reading the book was the possibility to add new senses to the human experience. In chapter three “The Human Brain” the author states, that intelligence is located in a part of the brain called the neocortex. The neocortex is a very thin layer (about 2 mm), covering the brain. It consists of several different units (functional areas) that process all the sensory information arriving through our nervous system.

There are units for visual input, tactile input, auditory input and so forth. Although all these units process different inputs and work in a complex hierarchy, Hawkins claims (based on observations by Vernon Mountcastle) that all units essentially work the same way. This is possible, because in the end all sensory perceptions are the same: electrical patterns with spatial and temporal dimensions. Therefore all the functional areas only need a kind of pattern-processing ability with only minor adaptions to the respective input.

One indication for this theory is observed on people that congenitally lack senses. People born deaf for instance use their auditory units to process visual information.

What it comes down to: every unit in the neocortex can adapt to process any kind of sensory information. This seems a very interesting idea, since it implicates the possibility to add new senses to the human body. From my understanding of this issue it should be possible to take some artificial sensors and link them to our
nervous system – as long as they provide an electrical pattern understandable by the neocortex (it probably needs some research on how sensory information are encoded). One functional area should then be able to “learn” to handle the new kind of input like it would have to learn to handle hearing or vision as a child. This would open up huge capabilities and enable all kinds of new experiences to mankind. We could see light beyond visible wavelength, maybe even through walls, we could hear frequencies out of reach, we could sense radiation or magnetic fields; it may even be possible to use sophisticated sensory devices such as microscopes or spectrosopes; maybe even wireless communication would be possible. Imagination is the limiting factor here.

Really linking the idea, I still see two technical problems with it. When a blind man uses his functional area that is usually dedicated to vision for hearing, he trades off one sense for another. The question would be: is there enough brain for additional senses? Or would the approach only allow replacing (broken) senses? Would people even be willing to sacrifice their natural senses for improved, artificial versions?

The second technical question is: will the brain still be able to adapt to new senses later on during life or only as a child when it learns to use senses anyway. In the book the author only refers to congenital defects. This would make the whole endeavor really difficult.

Finally there always is an ethical question about adding technical devices to the human body. We don’t want to become machines. Mankind has always done it though, with crutches, glasses and contact lenses. Nowadays we even have sophisticated prostheses replacing whole limbs. But usually we only use these devices to repair our natural capabilities not to upgrade them. But the desire exists: some people already try to experience new senses by implanting tiny magnets into their fingertips.

Another interesting topic is addressed in the last chapter of the book, “The Future of Intelligence”. In a short section the author discusses, whether
intelligent machines could become a threat to mankind. This is a very important question to ask as an engineer. The fear of new technologies has always been present in the human history and found expression in art, literature and movies. Frankenstein’s Monster is an example from the times the discovery of the battery introduced electronic devices. The thought of losing control over a powerful “technology” already prevailed in earlier times e.g. Goethe’s *The Sorcerer's Apprentice* or the Jewish Golem.

The relatively new idea of artificial intelligence evoked the horror scenario of intelligent machines which would start to reproduce and - due to their intellectual supremacy – enslave or extinguish the human race. This is topic of recent movies like *The Matrix* or *I, Robot*.

Hawkins explains we don’t have to fear his intelligent machines. The justification lies in a statement from the very beginning of the book: “[…] I am not interested in building humans. […] Being human and being intelligent are separate matters.” (p. 41).

The explanation is very simple, but yet enlightening: intelligent machines are not like humans (which have proven to be inclined towards world domination and enslavement). Since Hawkins claims intelligence stems from the neocortex and is therefore somewhat separated from the “old brain” which is responsible for our emotions and urges, it should be possible to build machines which are intelligent, but not emotional. In a way they only get our “good parts”.

Humans are biological beings with the urge to reproduce to preserve their race. A memory-based prediction-machine as described by Hawkins couldn’t even get the idea. Why should it reproduce? It’s simply processing sensory input and checking it against its predictions based on previous input. Therefore engineers can – and should! – safely work on Hawkins intelligent machines, which might soon help us in all areas of our lifes.
On Intelligence

When thinking of the future of computers, I truly believe researchers should dare to aspire as grandly as possible. Jeff Hawkins is such a person; he argues that Artificial Intelligence scientists “program computers to act like humans without first answering what intelligence is and what it means to understand”. He doesn’t only propose a daring and provocative approach on how the brain actually works but also convinces the reader to share its beliefs.

Saying that, I must mention recently seeing the talk of Simon Sinek on “How great leaders inspire action”. This may seem at first sight off the subject, but I would like though to structure this essay based on the notion of the “golden circle” he introduces (see figure below). As it turns out, says Sinek, all the great and inspiring leaders and organizations in the world think, act and communicate the exact same way, by answering the following three questions in the exact same order as listed below and not the other way around like the others do:

- Why? – purpose and beliefs;
- How? – method;
- What? – product/ findings.

Sinek argues that “people don't buy what you do; they buy why you do it. And if you talk about what you believe, you will attract those who believe what you believe”. I red the book of Jeff Hawkins and surprisingly found myself believing in his theory. I am confident that this is the “secret ingredient” that makes his work so important for the future development of truly intelligent machines: he starts by giving an answer to the “why” question; he makes people believe.

So why don’t we have yet intelligent computers? Why did Artificial Intelligence (AI) scientists and neural networks researchers fail in delivering such computers? Why is understanding the nature of intelligence important? Why do we first have to understand how the human brain works? Why do we need this new approach? The answer to these questions is what most impressed me in this book: contrary to the actual approach, intelligence and understanding cannot be measured by external behavior; designing an intelligent system doesn’t mean programming it in a way that it only produces a humanlike behavior. Intelligence is not a matter of acting or behaving intelligently. Of course, the behavior is a manifestation of intelligence, but not it’s central characteristic or the primary definition of being intelligent.
Alexandra MARINESCU

You might argue that there is no element of surprise in the answer given by Hawkins to the “why” question and I agree. But let me explain why I find this so important. When we ask ourselves what an intelligent system does, we intuitively think in terms of behavior. People normally consider the other fellow humans intelligent on account of the behavioral observations they make and mostly by assuming the others have the same internal structure. If we now think of a computer or, for a better understanding, of a robot, we know that the latter has a mimic behavioral output, but no similar internal constituency to humans. In order to convince human observers that they are truly intelligent, they only behave as though they were. All together, researchers concentrate on programming a system in such a manner as to only behave intelligent and define the product of their work as an intelligent machine. As you now see, Hawkins creates an added value by bringing true intelligence again to light: he returns to the very beginning, to the “big picture” if I may call it that way and restates the goal of computer science by pursuing “real” intelligence.

Again: why? One “can be intelligent just lying in the dark, thinking and understanding. Ignoring what goes on in our head and focusing instead on behavior has been a large impediment to understanding intelligence and building intelligent machines. [...] To succeed we will need to crib heavily from nature’s engine of intelligence, the neocortex. We have to extract intelligence from within the brain”. And there we already have our answer to the “how” question: how do we get to build intelligent machines? By understanding the human intelligence; and how do we understand human intelligence? By understanding how the brain works.

This brings us already to the third question: what does this new approach consist of? What is the basis of intelligence? What are the most important findings this theory relies on? The senses, writes Hawkins, create patterns that are sent to the cortex, and processed by a single powerful cortical algorithm to create a model of the world. The latter is then remarkably being held in memory. Our brain isn’t at all a computer; it doesn’t “compute” any answers to problems; it retrieves all answers from its memory. The entire cortex is actually a memory system. Based on the stored sequences of patterns, it is constantly predicting what we will see, hear, feel, mostly in ways we are unconscious of. These predictions are out thoughts, and, when combined with the sensory input, they become our perceptions.

According to Hawkins, prediction and not behavior is the proof of intelligence. So how can we use this to ultimately build intelligent machines?

The brain model introduced by Jeff Hawkins is daring and provocative yet very convincing, mostly because of its simplicity. It uses a few ground rules to explain how the brain actually works and gets therefore rid of the complexity leading to confusion among researchers. The made predictions weren’t though proven to be true, so I will try to stick by the main idea of the “memory-prediction framework” when shaping a statement on how this book inspires me to build intelligent systems.

In order to do that, please allow me to bring the robotics research field into discussion. As a whole, the latter departs from an engineering perspective, gains insight from biology and neuroscience and ends in a social perspective, by aiming towards a human-like behavior.
While reading the book of reference I found myself mostly thinking about the possible implications of this theory for the robotic research area, so I must disagree at this point with the author, who doesn't see human-like robots as a possible future development.

It is true that such robots are not nearly human-level in their abilities today. Yes, they fail. “They lose the topic in conversation, misunderstand us, and they disappoint as much as they exhilarate us. At times these failures frustrate the public and robotics researchers alike” (David Hanson). But being able to implement the mentioned cortical algorithm will most certainly solve all this problems. The robot will be able to understand and make predictions of the future, even make decisions on its own based on the already learned sequences of patterns. If such a unitary cortical algorithm is to be discovered, we will achieve total autonomy of robots.

I do agree robots will not be able to have any feelings, but I also believe the aim of robotic research isn’t for robots to actually be able to experience emotions. Most of the people consider affects as part of the human intelligence and therefore expect intelligent machines to be able to understand and communicate their feelings. But that isn’t the case, nor will it ever be. It isn’t wrong to aspire to make robots beautiful, compassionate, loveable, and capable of love, on the contrary, but my personal belief is that the social skills of a robot will never be of first priority. “You might think intelligent machines would need emotions to foresee patterns involving human behavior, but I don’t think so. [...] Intelligent machines can comprehend human motivations and emotions, even if the machine doesn’t have those emotions itself” writes Hawkins.

The author also argues that given the necessary cost and effort to build and maintain humanoid robots, it is difficult to see how they could be practical. “A robot butler would be more expensive and less helpful than a human assistant”, he says. As far as I am concerned though, humanoid robots are of great importance when considering Robotic Re-Embodiment (physical embodiment of a person onto a robot avatar). Also, studies show that autistic children respond favorably to such robots, promising treatments and social training uses.

In conclusion, I strongly believe that Hawkins’ theory, if it will be proven to be true, will give robotics research a new grounding. It will probably be the breakthrough we have all been waiting for. We will have robots as truly intelligent machines embedded in our everyday life. I find it therefore appropriate to conclude with the words of David Hanson: “the field has so much room left for innovation and diversification in design. Let a thousand robot flowers bloom".
The book *On Intelligence* explains Jeff Hawkins’s theory about what intelligence is and how it works. He first introduces the subject by explaining that for him an important expression of intelligence is the ability of the brain forming representations of the world. He makes the reader aware of how the brain is able to reconstitute consistent images / representations from ‘basic’ patterns. During this part of the book, he explains how the brain fills in lacking or damaged patterns with what he expects in order to constitute the representations we are used to.

I read these chapters during my first weeks in Munich and so I could experiment much of what he explains since I have a new home, a new university, a new language… For example when I speak German with someone I am aware of this fill-in process because it simply doesn’t systematically happen and I have to concentrate much more about what the other person says to really understand. A lot of things that were assured like basketball specific vocabulary are to be learned again and in some ways I feel like a child having to learn everything that surrounds me. It is in fact very exhausting and I now understand better why changes are so difficult. Taking Hawkins’ theory about how the brain works, it likes to make predictions about the world. So when you change almost everything in your life, it can’t predict anymore because its model of the world is not accurate. For example I was yesterday attending a lecture (in English) and a boy from another class came to me and asked me about the lecture I attended on Tuesday, the question was very simple but as long as I was very surprised and he was speaking very softly I could not understand a word of what he asked me. In fact I never knew him before and he came asking me in German while I was listening in English. The theory of *On Intelligence* explains this little ‘moment de solitude’: my neocortex was not able to predict what sentence or word he will say and with the noise of the class more the fact that he was speaking so softly I did not get anything. This is just an example I have in mind because it happened yesterday but I find amazing that a theory of what is intelligence and how does the brain work concerning intelligence can be applied in my everyday life so easily. Moreover, the ‘technical’ explanation of the neocortex and this model of memory-prediction is vulgarized at a comprehensible level for a lot of people even not scientists.

Getting over this first quite pleasant surprise (I am an intelligent being) I sticked blocked to the conclusion that intelligence is “just” the combination of memory and prediction. When I think about the qualities of an intelligent person I don’t think she has good memory. On the contrary people who learn everything by heart have a sort of lack of intelligence to me because they do have to memorize reasoning and are not able to “create” it. Memory can be trained. The more someone learns thing the more he will be able to learn. For example I made my previous studies in France and as much of the future engineers I had to make two intensive years of “Classes préparatoires” to prepare the entry exam for engineers’ schools. When I look back at this time it is hardly conceivable how many information we learned within two years… But our brain was trained and so could always learn more patterns and as Hawkins explains it, some at first very abstract ideas become always easier to tackle. Nevertheless, every student should, and with training could, memorize mathematics and physics but some of them passed the entry exam much better than others. Considering *On Intelligence* theory this difference between students is due to their capacity of making
analogies. It makes sense after going through the book but considering all information we now have about brain without any framework, isn’t easy to fit any theory? I mean, when convinced that a general framework is true, couldn’t we pick just the research results that fit our framework to prove it? The intelligence of human being is, as Hawkins says something that interest scientists since years and an amount of experiment have been made in order to understand how it works. Regarding the book, Hawkins answers the questions about creativity or animal intelligence but what about people very well educated and with culture unable to construct any ‘new’ reasoning? Marcel Proust used to say « Il vaut mieux une tête bien faite qu’une tête bien pleine » which would be translated as a well-made brain is better than a very full brain. Hawkins chooses to explain intelligence with memory, I do not agree: memory is a part of intelligence since without remembering anything we couldn’t be intelligent but the Einstein example is no exception since known inventors were not always raised in memory stimulating environments.

Now let’s tackle the second part of this exercise. What would I do with all the material of this book if I was part of a team trying to build intelligent machines?

Even if I do not agree with the all-puissant-memory theory, I think memory is an important part of intelligence and actually if a machine could have a good hierarchical memory it would probably allow very interesting uses. Within the examples Hawkins gives about intelligent machines, I think supervision and security is a great field because it is very boring to humans. Poor guardian who must watch six hours long supervisions screen in the middle of the night! If a machine could learn to detect an intrusion or any other “suspect” move it could watch with a higher efficiency. Moreover as Hawkins mentioned the machine sensors can be much more adapted to the watching of building that human’s senses (supervision cameras are most of the time infrared based to capture motion during the night which is not very human vision friendly). Watching the images is tiring for a human eye and after a couple of hours the attention of the security officer decreases.

So first of all our “super watching machine”, let’s call it Watchnight must have adapted sensors such as infrared camera and a sonar. The infrared camera allows seeing in the night. I thing a “normal” camera is not necessary. Then the sonar can be useful to detect movement and the exact position of objects in the room. Maybe a microphone could be an advantage but let’s first try with just infrared and sonar as core sensors. Watchnight has now sensors adapted to building supervision. The sensors would be in the building but the “intelligent” machine can be anywhere else.

After the description of the sensors, the most important part of our machine should become clear: how are we going to build the hierarchical memory in order to allow learning? Machine learning already exists but is based on algorithms in computers. Considering Hawkins theory, this way of learning cannot lead to human learning. But let’s focus on these methods of machine learning. The algorithms used to make the machine learn are categorized within different types based on “the desired outcome”. The three most used are:

- Supervised learning;
- Unsupervised learning;
- Semi-supervised learning.
The supervised learning consists in classifying inputs regarding a known model that an oracle (expert) enters in the machine. This method is near to our bucket classification made by the human brain as explained in *On Intelligence* but still algorithms and determined classes are used in this learning: after learning, the machine cannot adapt buckets to the inputs it must classify.

The unsupervised learning consists in classifying inputs too but there is neither model nor predetermined classes. The machine discovers the “nature” of data by its own. This method is also called clustering.

The semi-supervised learning is a combination of both methods.

These learning methods are based on mathematical algorithms that can sort out inputs. The unsupervised learning should be useful considering that the machine just takes the inputs and try to create the categorization. It could be interesting to see if this learning allows adaptability or not since it is a very important aspect of the brain-like hierarchical memory. So our *Watchnight* could be trained based on a method inspired by the unsupervised learning.

*Watchnight* has adapted sensors and can learn. Yet, what kind of representation will it have from its sensors and how do we will train it?

*Watchnight*’s two sensors are sonar and infrared camera. The basic patterns form those are respectively a wavelength and emissions in the infrared wavelength. Then the different level of abstraction of the sensors representations could be:

- Received wavelength characteristics (frequency, amplitude...) for the sonar and wavelength emission zones (red, orange, yellow, blue that are more or less warm) for the camera.
- Differences to the sent signal and reconstitution of objects by grouping the different zones into forms.
- Positioning and size of the objects in the room and recognition of the object nature (human, animal, piece of furniture...).

The machine could, at the end, have a common level of high abstraction with a precise representation of the room in terms of distance and nature of objects. It should then learn to detect what an intrusion is. To do this, training sessions with different scenarios can be made. Just as for a human security officer, some realistic intrusion case can be simulated and then we tell *Watchnight* this was an intrusion. It is difficult to realize how many times would be necessary to really train the machine because it must see a lot of realistic situations to be reliable.

When operational and reliable, *Watchnight* can be used to supervise buildings such as banks or stores. It should naturally be connected to a security agency and send an intrusion signal as soon as it detects one. The staff of the agency could do some more interesting work even if they should be able to intervene rapidly if *Watchnight* detects an intrusion.

This machine is an example of what kind of application I would like to build inspired from *On Intelligence* and using the memory-predication framework.
What’s your opinion of the most interesting thought?
Where did the book give you inspiration for building intelligent systems and what is your inspiration?

It is an established theory, that the world and everything it consists of, can be separated into smaller pieces until finally there are some kind of atomic elements left behind. So it seems logical that not only the universe consists of several solar systems, a solar system consists of planets and there are many different things on a planet that can also be divided into smaller constituent parts. I think it is very interesting that referring to Mountcastle’s theory (chapter 3) the brain is made up of a homogeneous structure of neurons because this fits the view of the world that we have. It eases the approach to figuring out how the brain actually works very much. There is only one particular algorithm that has to be understood comparing to an inhomogeneous brain structure for different senses. This is what makes the brain as flexible as it is. Sometimes it happens that one of our senses fails for some reason. Imagining that the brain consisted of an inhomogeneous structure where different regions would have a fix job, there was no possibility of using that area of the brain to make up for the lack of one particular sense. If there were fixed jobs for each region of the brain, there was no flexibility as we find it proven by real-life examples. One of them, mentioned by Hawkins, is Weihenmeyer who could use a special camera on his tongue to see.

As the brain consists of homogeneous neurons the logical consequence is that all incoming signals are treated the same inside of it. Although it is not natural for a human to think of visual or audio input as a similar kind of signal, again, it is a way to simplify to understand the functionality of the brain. Moreover if somebody is listening to music from a notebook or watching a movie on the same notebook, that person gets input via the ears and via the eyes. Processing into the brain, the structures of the signals become similar. But it is also interesting to think about where the audio and video signals are actually coming from. They are retrieved from a notebook, so they must be stored on a hard drive. Data that is stored on a hard drive usually consists of zeros and ones. So if somebody takes a closer look at this data, this person will not find any differences between the video and the audio data, just as there is no difference processing different kind of signals inside the brain. Even if it is not natural to us to imagine input of different senses as similar kinds of signals, this is exactly what has been done by storing different kind of data on memory. Hawkins says that the brain is not magic and can be figured out by humans (chapter 6). Although the structure of the brain is not similar to the structure of a computer, the principle of encoding the data into some kind of patterns is done by both of them.

Hawkins often uses a song to give an example of patterns. The melody consists of several notes put together to a pattern that a brain can recognize. Not only the way the data is encoded by the brain is a very interesting point, but also the way it is stored in the hierarchy of the neocortex. I consider it as another very interesting thought that the neocortex consists of different layers that are used to recognize patterns and retrieve information. Referring to this thought the neocortex consists of a layer hierarchy (chapter 6). Known pattern are stored in the lower part of the hierarchy and new patterns in the upper part. As a pattern gets more familiar
to the brain, it moves down in the layer hierarchy. So whenever a new sequence of signals enters the brain it is compared to the known patterns to make some predictions. But if there is an incoming pattern that is not recognized by any layer, it is moved up to the top layer, Hawkins refers to as hippocampus. In my opinion there are a few very interesting points concerning this theory about the hippocampus.

The first point is mentioned by Hawkins (chapter 6). He describes how his children are able to remember details of a play but as Hawkins himself gets older he cannot remember them anymore. This means that the more information is stored in your brain, the more patterns match existing patterns and the less patterns get into the hippocampus, where new patterns to remember things are generated. Does this mean that the more a person gets to know in young years the faster this effect will occur? Does it mean that a person is not able to remember things when one gets to a certain point where one knows too many patterns? Does it mean that the brain might also have a certain capacity, where it is not able to adapt to certain kind of input anymore? That the brain starts to confuse known patterns with new but akin patterns? It can be observed that older people often confuse details of stories you tell them. Moreover there are certain people that are geniuses in one special subject, for example mathematics. Often those people are really good in their own subject but have problems with other things. Those things might be as simple as preparing some food or washing their clothes. Are the brains of these people so full of mathematical patterns that there is no more room for simple things? Or is there a certain pattern for preparing food that matches some patterns for solving an equation?

Another point that is interesting about the hippocampus is that a person cannot remember new things without it being there. But this person still does know everything that one had learned before the hippocampus being damaged. This basically means that if someone suffers from hippocampus damage, this person will actually never be able to learn about the own deficiency. Imagining that kind of disease is scarifying to me. The person without the hippocampus will not be able to react to any long term changes in the environment. When time passes by, the people without the hippocampus may be able to recognize other people they know but what if their friends get older and start to change? As this person is not able to learn new things, the brain cannot adapt to the new situation and by some time will not be able to recognize that friends anymore.

There are a lot of different points in Jeff Hawkins “On Intelligent” that are worth mentioning. To me the most important thing is that the neocortex actually consists of a homogeneous structure which fits in our perspective of the world. The processing of different input signals inside the brain using similar pattern for different senses derives from that homogeneous structure. Finally one of the interesting thoughts is the structure within the brain and what actually happens to the brain as it evolves during a lifetime.

Before reading “On Intelligence” I have never wanted to create an intelligent system, because I have thought that the brain has a too complex structure to copy and to actually understand it. So for me there was no option that building really intelligent systems is really possible.
Furthermore I have tried programming neural networks and the algorithms are very inefficient as it gets towards bigger networks. But considering that there is a basic structure underlying the brain, it seems more practicable to me that there might be a way of figuring out how the neocortex works someday. As Hawkins points out, there is no magic behind it. Moreover humans have already found a way to store data on hard drives. This data can be encrypted and recognized by eyes and ears and stored in our brains although in a rather different way than on a hard drive. So my inspiration is that now I do believe that building an intelligent system could actually work.

In my opinion the most important point to understand is that building intelligent systems does not mean creating humanlike behavior but to use the structure of the human brain to do more intelligent things than humans can do (chapter 8). Those things could be done without being limited by the needs of the human body as food or rest. The idea is to use the brain’s very effective structure and use it for things humans cannot do like collecting values from weather stations placed all over the world and predict tomorrow’s weather.

The first step of building an intelligent system is to design a physical structure of the brain. So referring to the status quo of the current development of neuroscience it is still the most important thing to figure out. How shall an artificial brain look like? Similar to a computer program on silicon or will there be any new approaches? In my opinion this a very important question to work on. A brain can only work efficiently if it is given the right architecture.

After having the physical design set up, the brain is supposed to learn. I imagine that an artificial brain has some special sensors and I wonder how to teach that brain about what it is supposed to do. How do I tell a brain that I want it to produce a weather forecast? How do I extract the information I am actually interested in? When a child is born it needs a lot of time to develop its brain so that it can actually survive in a world like ours by itself. How long will humans need to train an intelligent system? As Hawkins pointed out, the advantage of a system which is built by humans is, that it can be copied. A human brain – more the content of a brain – cannot be copied from one hard drive to another. I think this is a shame because it would be very interesting to see how the same intelligent system develops in different surroundings. If there was an intelligent car that was driving on a busy road in town every day, it sure would develop other characteristics than the same car driving on a quiet road in the suburbs. So there are some experiments that could be done with artificial brains that cannot be done with real human brains. But still we could carry the conclusions from those experiments over to human behavior and human brains.

My personal conclusion from Jeff Hawkins’ “On Intelligence” is that the structure of the brain is homogeneous and therefore not as complex as I expected it to be. So my main inspiration from the book is that it actually is possible to create intelligent machines with an artificial brain. To get there we first need to work on a physical structure that makes it possible to hold such a construct. With that been done there are a lot of fields where artificial brains can be used in for example weather stations or intelligent cars. In my personal opinion it is very interesting to use artificial brains for a better understanding of our own human behavior.
Written Discussion on Jeff Hawking's
"On Intelligence"

Martin Reverchon,
November 21, 2011

1 What's in your opinion the most interesting thought/idea?

The most interesting thought is in my opinion the fact, that there is just one universal structure in our brains. More stunningly, this universal structure is capable of handling a vast amount of different tasks. From vision to hearing, from thinking to speaking and moving - one single principle of computation leads to a wide variety of tasks that can be handled. Moreover, the set of tasks one brain is able to cope with is arbitrarily exchangeable and expandable. The vastness of things one learns (and of course forgets) during his lifetime is enormous - especially when considering that all this information is stored in just about 1.3kg of biomass.

This structure is not only versatile but very effective in learning and execution. In fact our efforts to teach robots basic things like walking and talking seem to be inept when compared to the speed a toddler learns how the world works. It handles many routines of daily life in an astonishing speed. It is hard to believe that a biological system with all its disadvantages in terms of speed, reliability and vulnerability can cope with the challenges of daily life. It easily outsmarts everything humans have built and does so with just one universal algorithm.

The beauty of this fact is, that we only need to research this one algorithm. It will give us the answer to many problems researchers all around the world focus on in a great variety of subjects. With the knowledge about this one algorithm many of these subjects would be gifted with an elegant solution. Furthermore, every device and every tool built with this algorithm would be, like the human brain, capable of solving different tasks. In this way a normal smartphone equipped with neuronal nets could
not only write the appointment I just made over the phone into my calendar but it could also remember me that I should need to plan more time because I need to fill up the tank of my car before I let my smartphone drive me to the meeting. There it could fluently translate my order at the Chinese restaurant and inform the kitchen that I usually like my dishes well-spiced.

Recently I heard a lecture that described a cure to certain brain related diseases. It outlined that by analysis of people afflicted with morbus parkinson a certain pattern of neuronal activity could be found. Neurons in brain areas where these dysfunctions are considered to stem from were firing synchronously and thus triggering each other again. If stimulated with a brain pacemaker in "coordinated reset" patterns, the synchronization can be broken and the patient be nearly completely healed. This shows that the search for the brain algorithm has a delightful byproduct. The knowledge about this one algorithm and the insight into how the brain works many brain-related diseases could easily be remedied.

2 Where did the book give you inspiration for building intelligent systems and what is your inspiration?

Even if not directly leading to innovation I take much inspiration from the authors confidence in the soon upcoming progress in this field of research. If one of the leading experts on this subject feels so confident that huge progress in the field of intelligent machines is imminent I look with great curiosity into the future.

Another non-technical more methodological aspect was the fact that the solution to seizing such an enormous task as building intelligent machines lies in generalization. If a solution to a problem is not directly visible one sensible approach is to take a step back and to look at the problem from maybe very different functional point of view. I had the feeling that the author sees more sense in reconstructing the humans brain than to develop mathematics to achieve intelligence in machines. He as an electrical engineer took several steps back to look at the problem from a biologists and neurologists point of view and thereby developed a distinct and novel way to tackle intelligence.

Another very interesting aspect is the idea that brainlike structures may be connected to different sensors. It proofs the versatility of the brain and once again the mighty versatility of one single algorithm. Furthermore, I like the idea that in this way man and machine do complement each other. The machines look at our world with a mind similar to ours but with probes to feel completely different facets of reality. In a world which is thoroughly researched with our eyes and ears this might bear the possibility of new, fast and unforeseen research.

To take a step further, man himself might profit from this fact. Science-fiction tales of
cybernetic enhancement do suddenly not seem to be far fetched. The human brain should be able to adjust itself quickly to a new set of sensors wired to it, if the assumptions the author made are correct. It could cope with infrared sensors and cochlear-implants that are able to perceive ultrasonic sound. Telecommunication could be able without using special handheld devices. Other languages could easily be learned if the right words of a dictionary are simply routed to our invariant representation of the idea of the word. Furthermore we could project our ideas and imaginations directly onto screens, showing others directly what new ideas we have.

Transhumanism might be the subject of current technophobia. But I hope this is only the natural humans fear of change driven by the prospect of alteration of ones own body in ways one does not understand. Yet the possibilities would be incredible. To feel radiation, to see chemical reactions, to augment the reality we perceive with our senses would lead to a completely new view of our world. The technical progress that would stem from our new views is unpredictable. With our new senses we would learn new pattern that, applied to other problems, lead to novel inventions.

While these are all applications of future technology one needs to have a look into contemporary research. Here I find the idea that information is only useful when put into context very interesting. We do that well and automatically with our auto-associative memory. This means for machines that metadata itself is not feasible. To know something means to have an idea in which relation the object stands to its surrounding. This might be one key in contemporary research. Speech recognition might be enormously enhanced if the computer knows what we are talking about. The knowledge about a topic can be used to find words that are likely to come up during the conversation and might thus make less errors in recognition. The search for objects in computer vision might be accelerated if the context of where the object might usually be found is available. Therefore, a new way to link the data stored in the memory of a computer to other data and knowledge would bear much improvement for data processing itself and all its applications.
Report on the book “On Intelligence”

1. What is the main point of the book according to your opinion?

The author, Jeff Hawkins, tries to put together years of neuroscience experimental data and theories into a unified paradigm of how the human brain works and makes it intelligent. He touches on many subjects, such as the brains ability to recognize patterns (irrespectively of where those patterns come from) and build invariant representations of the world from them, but the main idea of the book is that prediction is the key thing that makes the brain intelligent.

Intelligence and prediction are intertwined. He argues that a single cell organism is also intelligent, because it can make predictions about its environment by using its senses and its “programming” embedded in its DNA. Going up the hierarchy, reptiles are more intelligent because they can accomplish more complicated tasks. Mammals though are even more intelligent due to the fact that they can not only make predictions about the world, but because these predictions can also change with experience. They are not static (at least not all of them), a result from evolutionary “programming”, but they are ever changing during the animal's lifetime. The point here is that while prediction is the key to intelligence, it requires a complicated memory system in order for it to work. The author calls it the “memory-prediction framework”. Humans are at the top of the “intelligence pyramid” because of their ability to pass on information from their memories and create such memories in others without the need for the other to have actually experienced the memory directly.

Prediction can not exist without memory, this is clear. The environment is very complicated and could not be predicted by using just the immediate sensory information. Just any memory, however, will not do. The memory in the human brain is totally different than that of a computer first of all because it is decentralized. There is no one area of the brain whose task is to store data, rather there are multiple such places, organized in a hierarchical fashion. Storing and processing is done in the same place and all these “functional units” are identical (there are no specialized cells or structures that are involved in vision or hearing, for example) and the borders between them are not clearly drawn. These areas respond to and store patterns, irrespective of whether these patterns come directly from the senses (down on the hierarchy) or from other areas (upper on the hierarchy). The way they respond to patterns is also different from the traditional computer memory mechanism of address and content. The memory in the human brain is auto-associative – a pattern applied to an area of the brain provokes a response in the places where that pattern is stored (the pattern does not have to be exactly like the one stored, it can be missing parts, slightly mixed up and so on).

In the author’s theory, the brain is always making predictions about what it is going to see, hear, touch, etc. next based on current sensory inputs and its past experiences. Areas that are on the lower levels of the hierarchy work with the sensory input directly. Let's take the example of looking at a face. The first vision areas work only with notions such as edges, colors, basic shapes, etc. They pass on this information to the higher levels, which create more complex models: edges and basic shapes get
translated into more complicated shapes, which form individual objects such as eyes, mouths, noses, which together form an entire face. In this hierarchical model, information is also passed down the hierarchy. If the brain is expecting to see a face, the face is deconstructed down the hierarchy in the same way the sensory input was constructed to form the face and the resulting feedback pattern is compared to the actually received input. When something unexpected happens (predicted and actual stimulus do not match), the brain takes action. In this way, simple tasks are automated, allowing for higher thought processes to occur while the basics are taken care of automatically. If something unexpected happens at this basic level, attention is focused on solving the issue and then, upon solving it, the higher thought processes can be resumed.

The author also talks about how these patterns are stored within the brain. Clearly it cannot store every individual pattern by itself. Instead, they are classified and stored together as a single class of parameters. The author calls these invariant representations. This happens all the way up the hierarchy, with the highest levels having representations of extremely complicated world objects or even abstract notions that may not have direct physical manifestations (they did not come directly from sensory inputs, but are the product of thought processes). The brain learns such new representations every day, when repeatedly exposed to new patterns.

2. How can the knowledge obtained from this book be used to build intelligent machines?

It is because of the fundamentally different ways in which computers and human brains work that the author believes current approaches will never lead to building truly intelligent machines. Fields like AI, while they can be successful in some applications, will never be able to capture human-like intelligence because they focus too much on functionality (programs for speech recognition differ greatly from programs for image recognition) and they lack a certain degree of understanding (speech recognition programs may recognize individual words, but they have no idea about what is actually being said – they lack the context).

It is a bit difficult to come up with original ideas on how the author’s framework fits into the task of building intelligent machines, seeing that the book contains a whole chapter dedicated to the topic. Things like adding new senses like radar, sonar, etc. to machines to complement or even replace the ones humans possess (totally possible since in the author’s view all cortical areas are identical and they are not physically specialized in any way to the task they accomplish) or using a prototype to teach a class of machines and then download the developed “mind” into newly build ones were ideas that popped into my mind while reading the book (and probably not only mine), but the author beat me to it.

One original thought I might have is this: even though computer memory works totally different from human memory, using the ideas outlined in this book it should be possible to implement (on a small scale) something resembling the hierarchical structure of brain areas in software running on a traditional computer.

It is my feeling that the author tends to get caught up a bit too much on topics like the structure and physiology of the brain and his ideas of building an intelligent artificial brain involve more or less copying the biological one (he talks about building silicon chips for memory that can be faulty in some storage locations, but have to be able to be interconnected to many more such chips, because this is the
way the brain is built and wired). Even he realizes that the brain is a product of evolution, so it may contain a lot of “spaghetti structure” (analogous to “spaghetti code” in software engineering – overly complicated code that results from redoing certain things or adding new things later on, because of improper planning before actually starting to write code), but still insists that copying the brain is the way to go.

I am not saying that his idea of building hardware that is more brain-like is a bad one. Implementing the connections and the decentralized memory already in silicon would create a brain that could outperform the human one in terms of speed, but basically just that. I like the view of the brain being the hardware platform and the mind being the software running on it. It should be possible eventually to just study the mind's behavior without necessarily always running back to how the brain implements it physically. My guess is that the author and I have slightly different views on what constitutes behavior: he considers behavior as something that can be observed (moving the arms, speaking, etc.), but I would include thinking and understanding in there, although it does not produce any tangible outputs, it does produce thoughts, notions, ideas that could be shared. This would be why he is against only studying the way the mind works, without necessarily studying the brain in tandem.

I do understand and agree with the author that AI is in no position to build truly intelligent machines, because of its fundamental flaw of approaching different classes of problems in totally different ways and without being backed up by any type of understanding. I also believe that his views on pattern based, auto-associative memories organized in a hierarchy could be tried out and most probably implemented on something that does not require a copy of the human brain to run.
“On Intelligence”

By Jeff Hawkins

Grazvydas Ziemys

21.11.2011
Jeff Hawkins is a computer architecture and entrepreneur in Silicon Valley. In his book “On Intelligence” the author presents his theory of Intelligence, describes how the brain works and discuss why all attempts in the past didn’t lead to success in creating an Artificial intelligence.

In the First Chapter the author defines what is the Artificial intelligence and briefly discus why all approaches in the past have failed to implement the AI. Moreover he tells about his personal experience in researches on AI and describes the history of realizing the AI. The Key error why all attempts have failed is the approach. Everybody was just trying program a digital machines to act Intelligent. But first of all intelligence, he says, is not about behavior. Behavior is just lateral expression of Intelligence. As a quiet good Method to proof the intelligent machine is a Turing test. „if a computer can fool a human interrogator into thinking that it too is a person, then by definition the computer must be intelligent. “ p.14

In chapter 2 Hawkins introduce to the reader the neuronal networks. Once again he gives a briefly history of neuronal networks researches and applications. Author explains main difference between conventional digital computer and Neuronal networks.

“A neural network is unlike a computer in that it has no CPU and doesn't store information in a centralized memory. The network's knowledge and memories are distributed throughout its connectivity— just like real brains.” p.24

Furthermore he postulates his three criterions to understand the brain. Understand this part of human body is essential to be able to create an intelligent machine. Firstly such a creation must be aware of time. Secondly he strongly believes that a feedback plays very important role in how the Brain works. Last criterion was that any theory or model of the brain must be based on physical architecture of a brain. In this chapter he also shares his thoughts about biggest obstacles in the history to discovering the truth. He illustrates that with an example of Copernicus, who because of intuitive but wrong assumption, said that earth was stationary at the center of the universe.
The most exciting and inspiring ideas of the book was described in Chapters three to six which explains how the human brain works. Mainly because that this is important to creating a model or theory of an intelligent machine and secondly it is amazing to get a bit near to understanding how our human brain physically works. In my opinion understanding about working principles of a brain, leads in more understanding our salves as Human being and even our place in universe. Knowing that one day everything appeared out of some mystery energy and in the billions of years led to the point where we are now. Going through evolution and ending up with such a complex systems like our brain. The most important part of human intelligence is neocortex which consists of 6 layers of neuronal tissue each 2 millimeter thick. Neocortex is on the surface of brain and envelops the older part of human brain. It contains around thirty billion neurons. The most interesting point about neocortex is how the communication between the neurons works. Cortex is divided in to different parts which take care of different tasks like auditory, vision and so on. But this division is physically quite hard to see. Mostly all parts differ in interconnection of neurons so basically it is the same neocortex just adapted to some specific task. Some experiments with rats prove that actually every part of neocortex can be adapted to each function. One of the main functioning principal is a hierarchy in neocortex. Author explains this using the example of human vision. The signals from eyes come through bundle of neurons to the lowest level of the vision area, there the signal pattern changing very fast like in our eyes. The higher in hierarchy signals come more stable they gets. That leads to conclusions that as signals travel through the layers the pattern turn to be more abstract and invariant. Invariance is very important in storing, recalling and using patterns. That is the one of the most difference between brain and computers. That allows for the brain to recognize for example the face even when it is in different position or distance. This task for a computer is very complicated and even impossible for now. Same principle is found in every activities of the neocortex. The neocortex stores and recalls invariant patterns in the sequences. That's why we can learn sequences of facts much easier than random ones. Invariance enables other very important mechanism of the brain which is auto-
associativity. In cortex exist a complex mechanism which allows for some clusters of neurons to be activated if incoming pattern is already trained in the past otherwise signals goes to higher hierarchy where higher level tackle the pattern in more abstract way. That’s way if why in known environment we need less concentration in doing some tasks, because in that case decisions are made in lower hierarchal level and we can use higher levels for another activities.

In chapter 7 author discuss creativity and consciousness. Basically humans are creative in very many ways. Because transferring an experience in to the new situation is the act of creativity. He argues that genetics also plays a role in our abilities but the most important is experience. Experience “programs” in us different patterns and the more such patterns we know the more ease we can find a solution to different problems by combining them. Experience physically seems to be stored in synapses. Some scientist claims that the more often a synapse is being used the stronger a connection is builded.

In the last chapter Hawkins draws a roadmap of the future of intelligence. In my opinion the idea of using a different approach on implementing a memory is very interesting, because as the technology shrinks in size we a slowly losing control of each and every bit cell. It is becoming very hard to produce a chip wit out any error. So if it would be possible to find a different way to store data for example that in case of losing few bits will not destroy hole data package will let us use even very new theology like Organic electronic in very effective way.

So in conclusion understanding how the brain works enable us to think of implementing such mechanisms in silicon or organic materials. Though it is still a long way to do so, even the brain must be understood more precisely. But the model of a human brain could be a very good example how to build an intelligent machine. And as the author claims it should not t be copied in every aspect, because it is simply too complex for now, but rather we must take the main concepts and mechanisms of a brain. I strongly believe that such approaches will lead in step by step creating intelligent machines. Maybe starting with robust memory and ending with some very elaborated creature.