

Artificial Intelligence

Manuel Pagliuca

October 11, 2021

Contents

1	Motivation	1
2	Neural networks	3
2.1	Biological background	3
2.2	Reasons for studying the biological background	3
2.3	Neurons	4
2.4	Computer vs Human brain	5
2.5	Threshold Logic Units	7
2.5.1	Conjunction example	9
2.5.2	Implication example	9
2.5.3	Multiple inputs example	9
2.6	Geometric interpretation	10
2.7	Linear separability	10
2.8	Convex hull	10
2.9	Solution of the bi-implication problem	10
2.10	Arbitrary boolean functions	10
2.11	Training TLUs	10

1 Motivation

Artificial neural networks are information processing systems, whose structure and operation principles are inspired by the nervous system and the brain of animals and humans.

Artificial neural networks are studied for various reasons.

- Extracting knowledge from data (*phenomena, events, processes, operating environment,...*).
- Understanding the phenomena that I'm observing, extracting the knowledge about that phenomena.

- Automated construction of computational paradigms for problem solving.

The ultimate goal is to be able to have a model which allow to describe the phenomenon that I'm observing and using this model I would like to solve the specific application problem.

There is an huge variety of application, this is the reason which AI now is so popular, and we have so many interested and expertise in this field.

The next evolution of the economy is based on availability of solution which are extracted from the data by analyzing the data reasoning, the only problems is that our capability of reasoning is limited by the fact that our brain is not able to analyze huge quantity of data, with automated system we can exploit this and make more comprehensive models that describe the process of systems we are observing and develop more accurate solutions.

Basically, artificial intelligence is mimicking the nature. We want to take the data analyze the data make a model, and for doing this we use techniques, we want to enrich the analysis we are doing through sensors from environments, we can use other specific sets of techniques.

We have the need of putting together some ideas and observation, we can define rules for reasoning, we extract the knowledge, but we can also build knowledge through reasoning. Basically, what we want to do is to try to replicate how the living beings observe and operate in the environment, how express, interacts.

Not only but we want to observe how the individual evolves, population evolves, in order to understand the trend of the environment. In this broad variety of approaches is to define some techniques which mainly are in two big categories, from the point of view of intelligence :

- **Symbolic approach**
- **Sub-symbolic approach**, which are sub-symbolic techniques for analyze the environment and the systems.

During this course we are focusing on sub-symbolic reasoning techniques:

- **Neural networks**, which is a simulation of a living brain.
- **Fuzzy systems**, which embed the definition of quantities, which are fuzzy (not defined in a very crisp way).
- **Evolutionary computing**, which is a set of techniques which mimics the natural evolution of the species so we can try to optimize the solution by using the basic concept of the natural evolution of the species.

2 Neural networks

2.1 Biological background

Basically, the aim of the artificial neural networks in *mimicking* the behavior of our brain starts from the fields of neural biology and neural physiology. What this discipline tries to do is to analyze the behavior of the biological neurons cells and understand how they behave for retrieving the information from the outer world.

The neurons use sensors and special cells to connect to the external world and get information. As humans or other biological species, we have five senses, taking for example the eyes, they are receptors which are able to see what is around us is, they are able to reconstruct the scene we got around us.

Basically, with artificial neural networks we are trying to replicate what **biology** does for us, in neurobiology we build a model of what is happening in the neurons, this model describes how the neurons are interacting together to extract the knowledge, to build memory, construct reasoning, take decisions.

In computer science we want to build this model in computer just for trying to do something similar, we want to mimic the behavior of the natural brain in order to try to replicate in an artificial environment the same operation.

We are building the *model* for neurons in the computer, in this way we are able to use these artificial models to *learn* the environment and to solve practical problem (predict possible behaviors and solve optimization problem, like we do naturally). We also have some basis in other disciplines like physics and chemistry since we may want to use the neural networks, also for describing physical phenomena, not only to create reasoning in our mind but we want to create models of phenomena, this can be used for various application. We can create an abstract model, and instead of observing the real world we can observe how the model behave in some conditions.

Neural networks and in general artificial intelligence, is a discipline of computer science and engineering, then we use some inspirations from other discipline but the core of the theoretical aspect of the foundation which defines the artificial neural networks this is an area which is been to computer science and engineering.

2.2 Reasons for studying the biological background

The reasons for studying the BB (*biological background*) in computer science are two:

- The **first reason** for studying the neural networks is the fact that they are a very appealing model, since they work in **parallel** means that they have an intrinsically extremely high *parallel process capability*. This is why computer scientist are so attracted to this topic. Our brain in many cases find the solution immediately, this is fascinating, this happens because we are exploiting the parallel capability of our neural networks.

- The **second reason** why we are studying this technology is the fact that there is really a huge amount of *practical application* in broad variety of area : *industrial manufactory, products medicine, finances, economy, social networks, data analysis, ...*

2.3 Neurons

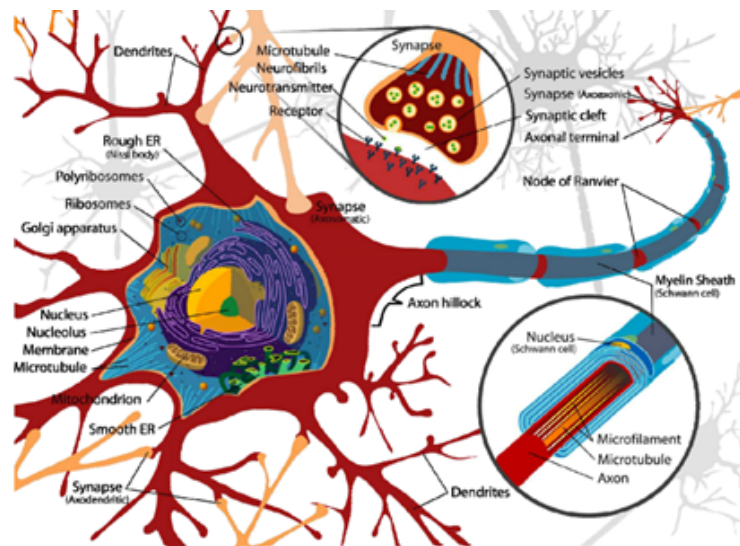


Figure 1: Neuron

The *central* part of the neuron is the part, which is managing the entire cell, this core the nucleus is listening what is happening around the neuron, when I have a **solicitation** from the external (other neurons or special cells related to the five senses).

When there is sufficient stimulation coming from this cells the nucleus is solicited and at a given point the nucleus realizes that the solicitation is so high that he has to take an **action**, the action is to send a *signal* around the *axon* (the long extension covered in blue), this will lead to have a *polarization* of an *electric signal* which is flowing along this connection and this will reach other neurons that are connected to the end of the axon to the synapses.

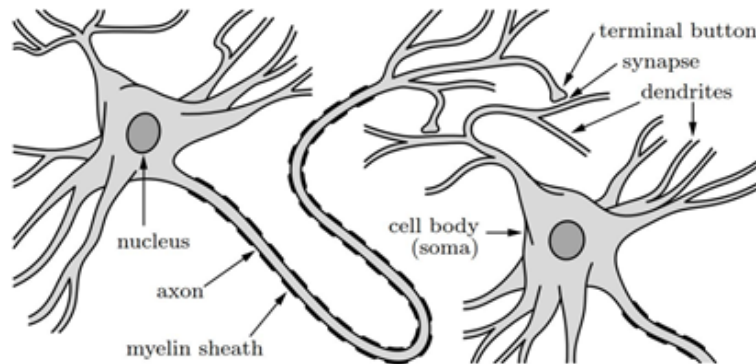


Figure 2: Connection between two neurons

The **axon** is covered by an appropriate *protein* (covered in blue) which protect the axon itself, and make sure that the polarization that occur on the nucleus is transmitted along the axon so that this signal is able to reach the synapse.

Basically, what we see is that the **neuron is exciting the nucleus**, which is generating the excitation along the axon, then it will reach the terminal synapses.

Each synapse is connected to the synapses of *another* neuron, so that the signals which are generated by the nucleus and sent to the axon and then to the synapsis terminal at the terminal the synapses are releasing some chemicals called **neurotransmitters**.

These *substances* are exciting the synapses of the connected neuron, bu stimulating the synapses this excitation is propagated from the **dendrites** to the nucleus of the other neuron, when the other cell is *sufficiently excited* by from the amount of stimulation generated by the neurotransmitter the nucleus will generate again a new stimulation which is going along the axon and reach another neuron (and so on...).

The axon is *depolarized* if there is enough excitatory input, basically this is how the signals are propagated through the brain.

Due to the fact that each neuron can stimulate each other neurons connected to them, this will create the **parallel processing** in our brain, so that each component will take care of analyzing each part of the information and this is a consequence to derive a part of the total computation.

2.4 Computer vs Human brain

When we look to the differences between a computer and the human brain what we notice, is that in the computer we got processors composed by many *transistors*, the human brain counterpart the got 10^{11} *neurons*, the number of the latter overcomes the number of the cores in a processor.

The neurons aren't able to process the same *complex operation* of a core, but they are still so many that they can overcome the limit of the complexity of

the individual operation with the fact that they are working significantly in parallel.

	Personal computer	Human brain
Processing units	1 CPU, 2–10 cores 10^{10} transistors 1–2 graphics cards/GPUs, 10^3 cores/shaders 10^{10} transistors	10^{11} neurons
Storage capacity	10^{10} bytes main memory (RAM) 10^{12} bytes external memory	10^{11} neurons 10^{14} synapses
Processing speed	10^{-9} seconds 10^9 operations per second	$>10^{-3}$ seconds < 1000 per second
Bandwidth	1012 bits/second	10^{14} bits/second
Neural updates	106 per second	10^{14} per second

Figure 3: PC vs Human brain

The **storage** of the neurons has an immense storage potential than any other memory capacity. The **processing speed** seems *much higher* on the computer, but the parallel operation that you can do in parallel in the human brain in the end *overcome* the processing speed of the hardware of a computer.

Basically, what we can observe is that the biological neural networks are able to outperform significantly the processor that we have nowadays, there are some researches carried out which are trying to create processors which are replicating in hardware the operation of the biological neural networks but still the capabilities of this systems are *far* from the capability of human brain.

It may *seem* in some system that they are working faster then humans, it may seem that they are running better, but actually what you have is that you have the ability in the computer to run the algorithms and the explore the possible (deterministic) moves in a fast way but needs some background knowledge in order to that in a very short time.

Advantages of biological neural network:

- High processing speed due to massive parallelism.
- Even if we have a significant amount of the biological neural network damaged, the system is considered **fault tolerance**, it remains functional even if larger parts of the network get *damaged* (maybe some functions will be disabled). The other cells are able to overcome the death cells, this thanks to the elasticity of the neurons which are able to overcome possible damages in the structures.

- If more neurons are failing, the brain will *degrade the performances* in a *graceful way*, it will not just stop working, will work a bit less not with the same performance and the same functionality but with reduced function. Only when a *really massive* number of neurons is dying at that point a function is not working anymore.
- They are well suited for inductive learning (*learning from examples, generalization from instances*).

What we are trying to do with artificial neural networks is to capture the parallel operation of the brain. The ability to extract the knowledge from the data, we want to replicate these capabilities.

There are some problems due to the ANN (artificial neural networks), if we kill part of the architecture which is replicating the BNN (biological neural networks), the ANN is **not automated to survives**, we have to assure some physical redundancy, this is a problem of the architecture we are using to execute the computation of the ANN.

What we want to do in our model is to construct a *set of abstract models of the neurons* that we call **artificial neurons**, and we like to connect them together to replicate the structure that we have in the natural brain this is why we have the ANN, a connection of neurons that is trying to mimic the behavior of the brain.

The complexity of the brain is so high that is actually difficult to replicate everything in ANN, what we are doing is to replicate a specific function of the brain which is able to solve a specific application problem that we have.

2.5 Threshold Logic Units

This is the first *abstract model* for an artificial neuron of the brain. A **threshold logic unit** (TLU) is a *processing unit* (neuron) with several inputs. It can solve a very simple set of problems.

We have a **core** (the neuron) with several *inputs* that are reaching the neuron, and we have the *output* which is delivered to the subsequential neurons which are connected with it. A TLU is a processing unit in which the output is governed by a threshold θ , if it has a *sufficient excitation* from the inputs, then the TLU became **active** (value 1) and generates the output y .

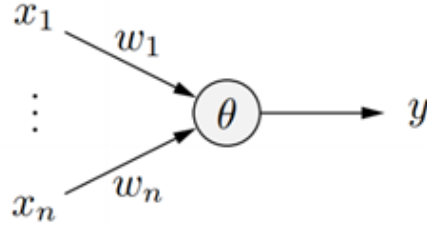


Figure 4: Threshold logic unit

We have n inputs identified by x_1, \dots, x_n the unit is generating only one output y each input is **not** delivered directly to the core of the TLU, but each input is **weighted**, some are more relevant, and others are less relevant (exactly how we are doing when we consider the data from the external world).

$$y = \begin{cases} 1, & \text{if } \sum_{i=1}^n w_i x_i \geq \theta, \\ 0, & \text{otherwise.} \end{cases}$$

Figure 5: TLU conditions

This means that our problem is depending on the most important information, this is what happens in the real world.

In the ANN we are replicating the relevance of the individual input, we can control how much each input is actually important to influence the generation of the output, to do this we use a weight w for each of the input so that the TLU will not see exactly the input but will see a weighted input, this will allow the modulation of each input by giving to each of them the appropriate importance during the generation of the final output.

When the TLU is solicited enough it will generate an output y that will be delivered to the *terminal synapse* to another neuron.

- If the **excitation is enough**, mathematically means that the weighted summation is greater than a thresh θ , we generate 1 (the neuron is active).
- If the **excitation is not enough** for overcoming the threshold θ , we generate 0.

This model that tries to represent what is happening in the BNN from two scientists, also called the **McCulloch-Pitts neuron**.

2.5.1 Conjunction example

The result is equal to 1 only when the two outputs are equal to 1. There is not a standard way to select the threshold θ , we have to choose that in base of the function that we want to implement. In this case I have to select a θ which is greater than the biggest weight.

$$x_1 \wedge x_2$$

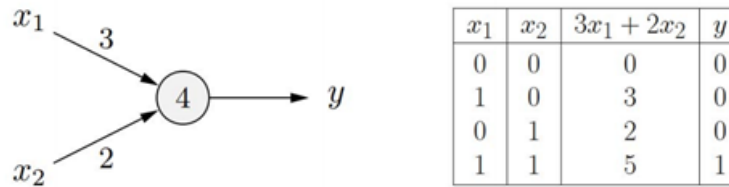


Figure 6: Conjunction TLU

2.5.2 Implication example

I can choose the value of the threshold in order to have the proper function, in this case applies on the same ways. *How can I choose the interconnection weights?*

The problem is that there are no general rules, I choose the weights according to the intrinsic relevance of each input variable.

How can I do that when we have a high number of inputs? we will see that there is a procedure for doing that.

$$x_2 \rightarrow x_1$$

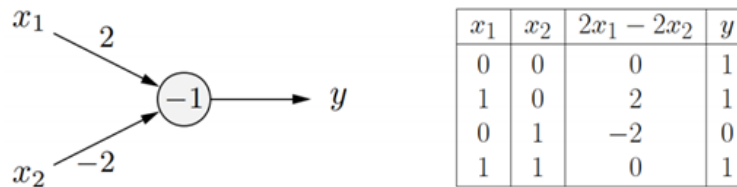


Figure 7: Implication TLU

2.5.3 Multiple inputs example

In this case we can see that we have three possible inputs, we can discriminate the inputs in excitatory input and inhibitory input. The first tries to contributes

to the final computation of the neuron in such a way that the results will be greater than the threshold, the other neuron does the opposite.

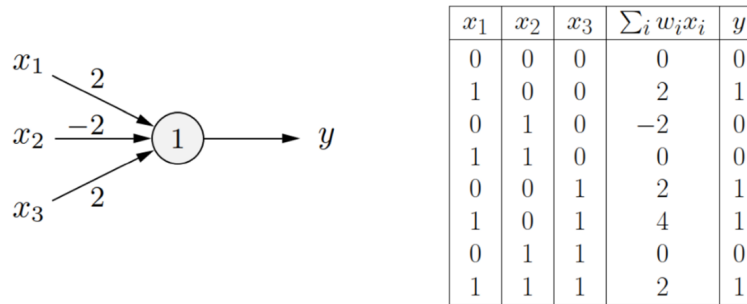


Figure 8: Three-input TLU

2.6 Geometric interpretation

2.7 Linear separability

2.8 Convex hull

2.9 Solution of the bi-implication problem

2.10 Arbitrary boolean functions

2.11 Training TLUs