HOW SIMILAR ARE WE TO A MACHINE?

Panev A.S.

Belarussian State University of Information and Radioelectronics, Minsk, Republic of Belarus

Liakh Y. V. – Lecturer

Annotation. In this thesis it is explained why neither Artificial Intelligence nor humans have free will.

Keywords. Free will, Artificial Intelligence, goal, brain.

Artificial Intelligence (AI) is everywhere these days: YouTube algorithms, voice Associates, autopilot for cars. There are things that use AI, and you don't even know about it. Take real estate prices which are often calculated by Artificial Intelligence rather than humans.

It is widely believed in society that AI can have free will and then it can take over the world. But in reality, the AI will not have free will. It can destroy the world only in two cases: if a direct goal is set to capture or destroy the world, and if some goal is set that can be most effectively achieved by destroying the world.

To make the second option clearer, let's use the example of watering flowers: the programmer sets the goal for the android to learn how to water the greenhouse so that the soil moisture is always within acceptable limits, and the irrigation scheme should consume the minimum amount of electricity. Sensors in the ground show humidity. The bottom line is that AI can guess to put the sensors in a closed container with the soil of the right moisture. Then, although the main goal was achieved (humidity is within acceptable limits and low power consumption), all the plants will die.

Thus, AI should not harm mankind if we formulate goals correctly.

A more likely scenario begins now. Every year AI makes our life easier in a variety of areas. At some point androids will appear, which cannot be externally and intellectually distinguished from a person. To avoid confusion and various kinds of misunderstandings, we most likely need to give androids a distinctive feature of the type of a burning light bulb on their forehead.

Although their behaviour may be similar to ours, androids will not have anything like free will. However, for many people who do not understand how artificial intelligence works, using androids as workers will look like slavery, which can lead to massive protests from people. Here is an example of how people have already felt sorry for robots. A man on Twitter writes: "I gave my mother a robot vacuum cleaner, but she practically does not use it, because she feels sorry for him." But the robot vacuum cleaner outwardly does not even look like a living creature.

We sympathise with an inanimate object, attributing the qualities of a living being to it: consciousness and the ability to feel pain and emotions. Imagine what will happen when such robots start talking. Should we feel sorry for them? Should we think all the words they say do not come from neural network algorithms, but from a certain soul trapped in a microcircuit?

If you do not delve into the topic, the answer to this question seems to lie on the surface: what behaves like it has consciousness, possesses it. But if you look at the facts, even if androids are

indistinguishable from humans either in appearance or in behaviour, it will still be a program in artificial flesh, the main purpose of which will be creating the illusion that an android has a soul. If we believe in this illusion, then the algorithm has done its job perfectly well.

So, in short: how does AI work?

There are inputs, there are internal processing "neurones", and there are output signals. This is called Neural Network (Figure 1). The signal travels from one neurone to another, each of which receives signals and produces a result according to the simplest mathematical formulas. And there are billions of such neurones. Everything is logical there and there is no place for free will or randomness. That is, with the same input data, the same neural network will always produce the same result. And our brains are no different.
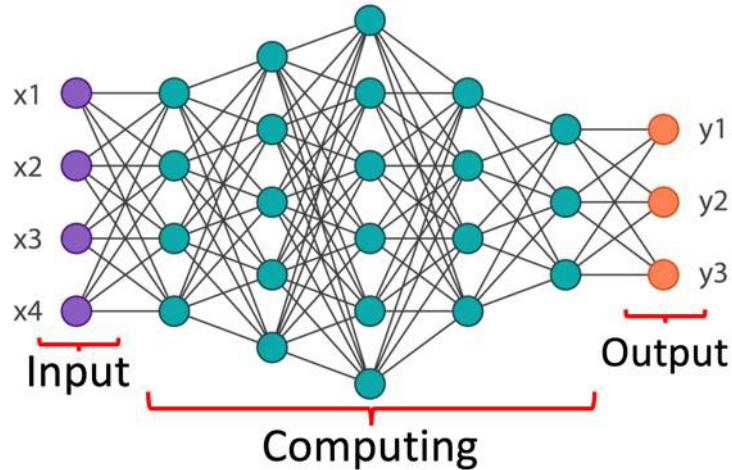

Figure 1 – Neural Network

What do brain tumour, sugar abuse, periods and anabolic steroid abuse have in common? All of these factors have been successfully used in court to explain the behaviour of killers. For example, in the case of sweets, the man's sentence for double murder was drastically reduced [2]. Free will includes responsibility for actions taken. That is, the court was suing the fact that at least sometimes we partially or completely lose free will. But who said that when the state of the brain, the levels of hormones and blood sugar are normal, we are responsible for our actions, and when the sugar level is not acceptable, we suddenly lose control over what we do? It cannot happen suddenly: in fact, in my opinion, it doesn't happen at all.

Let's take some simple action. The person tries to bend the finger. The brain sends a signal telling the muscles to contract. An absolutely insignificant, on the whole harmless action, what consequences can there be? Very serious, if a person has a gun in his hands – that is, the action is now not so harmless.

Neurobiology allows you to find the specific neurone in the motor cortex that sent the signal and caused the finger to bend (Figure 2). You can also find neurones in the so-called premotor cortex that sent a signal to the motor cortex, from where the neurone sent a signal to the arm. These neurones, in turn, received a signal from the frontal cortex, and it from the prefrontal cortex, and that, in turn, from the parts of the brain that correspond to emotions.
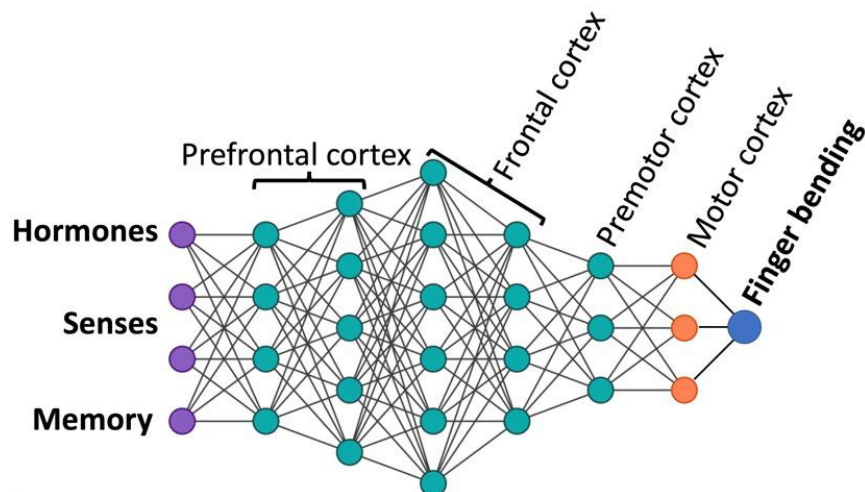

Figure 2 – Neurones in human brain

So let someone show a neurone that is at the source, which would have worked itself, let someone indicate a point of action that arose out of nowhere, without external influences, let him show a neurone, which suddenly decided to violate all known laws, suddenly sending a signal without preceding this reason...

Let us be shown the neurone that will be able to do it, the neurobiological basis of free will would appear, but so far this has not happened. There is no reason without other reasons. Nothing in the biology of our behaviour just happens spontaneously. A whole story is revealed behind each event [1].

Any person, whatever their principles, while driving a car that is about to inevitably crash straight into a pillar, directs the car to the left in order to get less damage than the person sitting on the right. No matter how it seems to us that we are deciding something, any of us with the same parameters and structure of a brain gives the same action. If you had the same tumour in the part of the brain responsible for aggression as the murderer, who a week ago was a calm family man but killed a child today, you would also commit a crime.

There are many incredibly important research results that are interesting to know. Here is the essence of one of them.

Researchers took more than 5,000 court decisions, that were taken directly at the court session, and looked, on what depends, whether the prisoner will be sent ahead of schedule or sent in court to serve a full sentence. The only indicator that helped in predicting something is how many hours have passed since the judge had eaten. If your case is considered immediately after the lunch break, your chances of early release are 60 percent – after two hours they remain 0. The work is rechecked as many times as possible, but no errors were found.

Almost certainly, if you ask the judge to substantiate a deliberate decision: for what reason to release this person, and send another to return to prison, they will quote you philosophical works that are read in the first year of law faculty or whatever will justify their decision, but no one will think about the blood sugar level.

We make a huge number of decisions without our own knowledge and neuroscience confirms that regularly, when we make a deliberate choice, especially when it comes to morality, in fact we have already decided everything on the level of emotions a few seconds earlier, and the conscious, reasonable and rational part of the brain then just comes up with an explanation, which we ourselves believe. We are machines that like to think that something is in control.

References:
1. Interview with Robert Sapolsky. – Access mode: https://youtube.com/playlist?list=PL8YZyma552VcfR6U49IyqClcDfpO65OPp– Date of access: 10.2020.
2. Stanford University, Robert Sapolsky, course "Biology of human behaviour". – Access mode: https://youtube.com/playlist?list=PL8YZyma552VcePhq86dEkohvoTpWPuauk – Date of access: 06.2020.