

MORAL ARTIFICIAL INTELLIGENCE

Kiseliova M.P.

*Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus*

Lazarenko A.M. - Senior Lecturer

The issues of Artificial Intelligence, often ignored by humanity, are raised. Neglecting moral aspects is analyzed briefly.

Imagine a native English speaker who knows no Chinese locked in a room full of boxes of Chinese symbols together with a book of instructions for manipulating the symbols. Imagine that people outside the room send in other Chinese symbols, which, unknown to the person in the room, are questions in Chinese.

And imagine that by following the instructions in the program the man in the room is able to pass out Chinese symbols which are correct answers to the questions. The program enables the person in the room to pass the Turing Test for understanding Chinese but he does not understand a word of Chinese.

Imagine the powerful artificial intelligence, that can easily talk about everything. This AI will be able to sail through Turing test. Does it mean that it understands the meaning of the words said, reflects itself and must have the same rights and attitude as a real person? Chinese room says no. But imagine you see a robot and a real person in front of you and you can not differ them, how will you treat them differently?

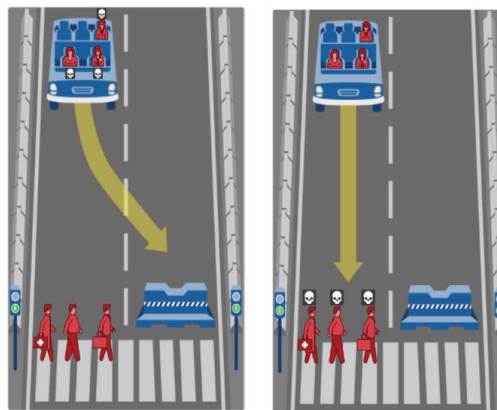
Despite films' and books' plots, the real threat of AI can be not in receiving consciousness and enslaving humanity, but in it's ability to make the most efficient decision. The machine will perform its function, for example, ore mining in most effective way. But what if it means to blow up the rocks under the city or to get rid of ill people in the pit?

Here moral question appears again. AI, which influences our lives, must observe ethic rules, and we are the ones to establish them at the development stage. But how will we do that, if we can not answer those questions in first place.

Imagine you are driving a car and your brakes have jammed. There is a wall in front of you and the only way to save your life is to turn left and kill a woman standing in the courtyard. What would you do? Well, most people will succumb to the self-preservation instinct and turn left, no matter what answer they would give. But AI won't have instincts and will make the right decision. And what is right depends on what we'll teach it. So, we should decide what is objectively correct. Would you agree to buy a car, that will choose to kill you? Or would you go outside with hundreds of machines-killers riding the roads.

On the site <http://moralmachine.mit.edu> you can play God and choose whose life matters. At this site you are to choose what machine should do in different situations. You can see one of them on the picture 1. According to the statistics of this site for men, elders, criminals and poor people it will be better not to leave their homes in the future at all.

In this paper I want to emphasize the importance of moral questions in this realm. Without answering them, there is no chance for humanity to go further in artificial intelligence development.



Picture 1 – the example of machine moral choice

References:

1. <http://moralmachine.mit.edu/hl/ru>.
2. <https://medium.com/@transphilosophr/searles-chinese-room-thought-experiment-a-twist-c7eb28f65e6c3>.
3. Stuart Russell: *Artificial Intelligence: A Modern Approach*