

Essay on "Is Superintelligence Impossible?"

Lorenzo Vigo

May 5, 2022

The term superintelligence refers to a system that counts with intelligence or capabilities that surpass those of human beings. These systems could be used to carry out tasks and solve problems. However, researchers and philosophers discourse about their destiny: will they surpass human kind and substitute us? Will they be used to enhance our nature? Also, will these superintelligent beings obtain conscience?

Daniel Dennett and Dave Chalmers, authorized voices in this area, are reunited in this talk to discuss some of these topics, in the event of the publishing of a new book. Both coincide in the belief that superintelligent systems are possible but that they are decades away from existence. In the same way systems nowadays explore and learn from scratch in spaces of limited complexity (a game of Go, for example), AIs could explore the space of evolutions and apply evolution to themselves. In addition, AIs could learn to design other AIs and generate better AIs than themselves.

Nonetheless, Daniel does not think we will ever create these superintelligent beings despite the fact that it is possible. There are countless possible things humans simply do not carry out because they are not worth it, or are too expensive, too hard, etc.

Sadly, I will have to disagree. Greed and pride are too present in our nature, and if these technologies may be useful for anybody to obtain their personal interests, they will become a reality at some point. Dave points out how it will be hard to avoid witnessing these developments in wars and conflicts.

Related to this concern, another worry rises: how autonomous will these systems be? We understand 'autonomy' as the capability of having goals and achieve them. In contrast, tools are used to fulfill objectives, but they do not achieve them on their own. Daniel fears that humans inevitably will be pushed out of the loop, in the same way as we do not use maps anymore (GPS, Google Maps) or do not perform simple but not trivial computations by ourselves (square roots).

Surprisingly, Dave disagrees with this point of view, claiming that humans still can take part in the decision when consulting the best path in Google Maps. However, the human brain is progressively more atrophied, and we are less and less able to perform actions that were usual in the past. Should we not forget to mention either the decrease in the average attention span.

Lastly, they discuss whether these entities will be conscious. Creating new conscious beings implies a huge number of philosophical issues. However, in general, we are interested in powerful tools rather than in artificial companions, as Daniel states. There are some scenarios where artificial companions could be helpful (the

elderly, to combat depression...).

To conclude, we will have to wait some decades to take the step and create superintelligent agents. However, we will probably be still arguing whether we should take it when somebody becomes the first to do so. For this reason, we should use our present time wisely and consider all philosophical scenarios to find out now the answers we will need in the future.