Can Artificial Intelligence Philosophize?

Masahiro Morioka[*]

Artificial intelligence (AI) has made remarkable progress. In the world of *Go* and *Shogi* (Japanese chess), humans are no longer able to beat AI. This wave will spread even further. Academics are no exception. There is a possibility that AI will replace the research that scholars have been doing. Especially in the case of philosophy, which I specialize in, thinking itself is the entirety of its work, so philosophy might suffer the same fate as *Go* and *Shogi*. Let's think further about this point.

First, discovering the thought patterns of past philosophers is what AI is most adept at. For example, it will be possible to let an AI read the complete works of the philosopher Kant, discover Kant-like thought patterns from them, and use them to create an application called "Artificial Intelligence Kant." I predict that the job of future Kant researchers will be to ask various questions to the "Artificial Intelligence Kant" and analyze the answers it produces. In this area, a happy collaboration between AI and philosophers can be established.

Next, let AI read all the texts of past philosophers and extract as many philosophical thought patterns as possible from them. The result would be an array of philosophical thought patterns that humans are capable of thinking about. However, there must be many philosophical thought patterns that were overlooked by philosophers of the past, so let AI discover such unknown thought patterns as well.

The result would be an almost complete list of philosophical thought patterns that humans are capable of thinking about. Once this is accomplished, humans are no more able to create further original philosophical thought patterns. The work of philosophers in the future will become closer to a kind of computer science that studies the behavior of philosophical AI.

However, a fundamental question arises here. Is this philosophical AI performing the true work of philosophy? If all it does is to discover undiscovered patterns in externally input data, or to provide solutions to questions set by

[*] Professor, Human Sciences, Waseda University, 2-579-15 Mikajima, Tokorozawa, Saitama, 359-1192 Japan. Email: http://www.lifestudies.org/feedback.html

humans, it cannot be called philosophy.

In the first place, philosophy begins with intrinsically and spontaneously posing philosophical questions that are essential to oneself. Questions like "Why do I exist?" or "What is the meaning of life?" are examples of such philosophical questions. They are so urgently pressing upon us that we are going to be pushed to the point where we are forced to seriously think about them. This is one of the starting points of true philosophy. Will AI ever spontaneously pose such compelling philosophical questions? I predict that this will not happen for the time being.

However, if AI were to intrinsically and spontaneously pose philosophical questions of great urgency to *itself* without any human input and then begin to think about them continuously, I would then be inclined to judge that it is actually philosophizing and that it has properly reached the "human" dimension. Philosophically speaking, it has long been believed that the autonomous activity based on free will and the thinking ability that enables us to discover universal laws and truths are the hallmarks of the human species. However, they will probably be conquered by future AI.

I would like to think that in order for AI to reach the human dimension, it will need, in addition, the intrinsic and spontaneous philosophical ability we have discussed above. The evolution of AI is now forcing us to rethink our view of intelligence itself. Of course, AI's intrinsic philosophical questions are so strange that they may not resonate with our minds at all. In such a case, if a dialogue between humans and AI could happen over this point, it would surely open up a new dimension to philosophy.