

‘GOOD’ FOR AI’S

SHIN KIM

Hankuk University of Foreign Studies, South Korea

E-mail: shinkimp@gmail.com

‘Good’ does yeoman’s job. It allows us humans to express preferences, attitudes toward a state of affairs and persons, and sometimes even objective facts of/about a state of affairs and persons. We are awfully good at figuring out, given a context, what ‘good’ means in the situation. Our abilities nonetheless do not guarantee that every individual in the same situation shall agree as to its meaning. We often talk past each other: someone might use ‘good’ to express efficiency while others might take one to express his/her liking. All these are familiar to us humans, and we do well to cope with the discrepancy. We talk to each other and balance things.

Now can robots do the same? ‘Robots ought to act ethically’ sounds okay to human ears. But how would that sound to AI? Would they be able to take the sentence and distinguish among the many different meanings that the evaluative/normative expressions? That is, would AI’s take the sentence to mean just as a human person would take its meaning? I think that it is easy to say ‘no’ to these questions. “AI’s do not have internal lives of their own. No preferences, no attitudes toward a state of affairs and persons.” But I also think it important not to be hasty in answering these questions. So, opposite to the easy path, this presentation will seek the things that are required for AI to have meta-ethics of their own.

In this presentation, we shall begin by questioning if it is required for AI to understand ‘good’ appropriate to the situation that they should have internal lives of their own. I would like to suggest that it isn’t a requirement. Instead, it is one of the dogmas of anthropocentric understanding of the mind. Cleared of the misconception, it should be possible to set up and investigate the meta-ethical question for AI. This presentation investigates the very question about the meaning of the ‘good’ and the ‘bad’, the ‘right’ and the ‘wrong’ from the Robots’ perspective.