

Epistemoral Problems

Mihály Héder

Budapest University of Technology and Economics

Presentation at Budapest Workshop on Philosophy and Technology, December 9-10, 2021.

My presentation argues that when it comes to the education of engineers in the moral dimensions of AI:

Epistemology > Ethics

(that is, epistemology is more of a challenge than ethics).

The talk discusses a notoriously hard AI Engineering problem: creating the artificial agent's internal representation of the world and itself. This representation is a precondition for discussing the usual moral dilemmas.

The core of the issue is that ascertaining the available alternatives of action, decision making, and execution of the previously decided action must happen in parallel and use the same resources and exhaust the same time allowances. Therefore the designer must decide which epistemic function must take precedence in what conditions and to what level of accuracy and confidence.

This line of thought is then generalized in the claim that moral obligations may result in epistemic duties in almost any case except for certain special circumstances. This situation is due to the unavoidable underdetermination of internal representation.

Moral dilemmas are usually represented in alternatives with definite or probabilistic outcomes. The outcomes are often undesirable, forcing the thinker to sacrifice one value for another, revealing priorities between them.

While in a thought experiment the alternatives are provided at the outset, for a generic AI application the creation of these is up to the code and therefore, to the coder itself. The challenge of internal representation building brings an epistemic component to all moral AI problems, a component that requires addressing first: hence the problems of AI are epistemoral.