

A Preliminary Study of AI Ethical Duality: AI Ethics and Ethical AIs

Zhen-Rong Gan

Research Center for Humanities and Social Sciences, National Cheng Kung University
E-mail: ganrec@mail.ncku.edu.tw

Hahn Hsu

Department of Philosophy, National Chung Cheng University
E-mail: pyhhsu@ccu.edu.tw

Abstract

AI (artificial intelligence) is characterized by autonomous learning and reactive capabilities. Arguably, these characteristics mean AIs can be held liable for failing to adhere to ethical, legal and social regulations when undertaking autonomous acts. In this paper, we consider the duality of AI ethics: that AI ethics should work both for humans and Artificial Intelligences. We lay out a more inclusive ethical view which would accommodate the idea of AIs characterized by distributive agency, which posits that AIs possess only a minimal degree of autonomy or intelligent autonomy in virtue of which AIs can be given a certain agency status. The agency status of AI will differ from that of human beings, governments, and corporations, nonetheless, the inclusive view also suggests that AIs ought to be held accountable for their acts: shared responsibility. We also suggest an idea of a bottom-line view of moral obligations as the starting point for the duality of AI ethics. Finally, we suggest an implementation mechanism of AI ethics based on case evaluations.

Key Words: artificial intelligence (AI) ethical guidelines, the duality of AI ethics, a more inclusive view of AI ethics, distributive agency, a bottom-line view of moral obligations