

Researchers expand study of ethics, artificial intelligence

February 16 2021



Credit: CC0 Public Domain

The Army of the future will involve humans and autonomous machines working together to accomplish the mission. According to Army researchers, this vision will only succeed if artificial intelligence is perceived to be ethical.

Researchers, based at the U.S. Army Combat Capabilities Development

Command, now known as DEVCOM, Army Research Laboratory, Northeastern University and the University of Southern California, expanded existing research to cover moral dilemmas and [decision making](#) that has not been pursued elsewhere.

This research, featured in *Frontiers in Robotics and AI*, tackles the fundamental challenge of developing ethical [artificial intelligence](#), which, according to the researchers, is still mostly understudied.

"Autonomous machines, such as automated vehicles and robots, are poised to become pervasive in the Army," said DEVCOM ARL researcher Dr. Celso de Melo, who is located at the laboratory's ARL West regional site in Playa Vista, California. "These machines will inevitably face moral dilemmas where they must make decisions that could very well injure humans."

For example, de Melo said, imagine that an automated vehicle is driving in a tunnel and suddenly five pedestrians cross the street; the vehicle must decide whether to continue moving forward injuring the pedestrians or swerve towards the wall risking the driver.

What should the [automated vehicle](#) do in this situation?

Prior work has framed these dilemmas in starkly simple terms, framing decisions as life and death, de Melo said, neglecting the influence of risk of injury to the involved parties on the outcome.

"By expanding the study of moral dilemmas to consider the risk profile of the situation, we significantly expanded the space of acceptable solutions for these dilemmas," de Melo said. "In so doing, we contributed to the development of autonomous technology that abides to acceptable moral norms and thus is more likely to be adopted in practice and accepted by the general public."

The researchers focused on this gap and presented experimental evidence that, in a [moral dilemma](#) with automated vehicles, the likelihood of making the utilitarian choice—which minimizes the overall injury risk to humans and, in this case, saves the pedestrians—was moderated by the perceived risk of injury to pedestrians and drivers.

In their study, participants were found more likely to make the utilitarian choice with decreasing risk to the driver and with increasing risk to the pedestrians. However, interestingly, most were willing to risk the driver (i.e., self-sacrifice), even if the risk to the pedestrians was lower than to the driver.

As a second contribution, the researchers also demonstrated that participants' moral decisions were influenced by what other [decision](#) makers do—for instance, participants were less likely to make the utilitarian choice, if others often chose the non-utilitarian choice.

"This research advances the state-of-the-art in the study of moral dilemmas involving [autonomous machines](#) by shedding light on the role of risk on moral choices," de Melo said. "Further, both of these mechanisms introduce opportunities to develop AI that will be perceived to make decisions that meet [moral standards](#), as well as introduce an opportunity to use technology to shape human behavior and promote a more moral society."

For the Army, this research is particularly relevant to Army modernization, de Melo said.

"As these vehicles become increasingly autonomous and operate in complex and dynamic environments, they are bound to face situations where injury to humans is unavoidable," de Melo said. "This research informs how to navigate these moral dilemmas and make decisions that will be perceived as optimal given the circumstances; for example,

minimizing overall risk to human life."

Moving in to the future, researchers will study this type of risk-benefit analysis in Army moral dilemmas and articulate the corresponding practical implications for the development of AI systems.

"When deployed at scale, the decisions made by AI systems can be very consequential, in particular for situations involving risk to human life," de Melo said. "It is critical that AI is able to make decisions that reflect society's ethical standards to facilitate adoption by the Army and acceptance by the general public. This research contributes to realizing this vision by clarifying some of the key factors shaping these standards. This research is personally important because AI is expected to have considerable impact to the Army of the future; however, what kind of impact will be defined by the values reflected in that AI."

More information: Celso M. de Melo et al, Risk of Injury in Moral Dilemmas With Autonomous Vehicles, *Frontiers in Robotics and AI* (2021). [DOI: 10.3389/frobt.2020.572529](https://doi.org/10.3389/frobt.2020.572529)

Provided by The Army Research Laboratory

Citation: Researchers expand study of ethics, artificial intelligence (2021, February 16) retrieved 17 April 2024 from <https://techxplore.com/news/2021-02-ethics-artificial-intelligence.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.