



Kitchener Post

Should your driverless car hit a pedestrian to save your life?

John Markoff

Waterloo Region Record | Jun 24, 2016

People say that one day, perhaps in the not-so-distant future, they'd like to be passengers in self-driving cars that are mindful machines doing their best for the common good. Merge politely. Watch for pedestrians in the crosswalk. Keep a safe space.

A new research study, however, indicates that what people really want to ride in is an autonomous vehicle that puts its passengers first. If its machine brain has to choose between slamming into a wall or running someone over, well, sorry, pedestrian.

In this week's Science magazine, a group of computer scientists and psychologists explain how they conducted six online surveys of U.S. residents last year between June and November that asked people how they believed autonomous vehicles should behave. The researchers found that respondents generally thought self-driving cars should be programmed to make decisions for the greatest good.

Sort of. Through a series of quizzes that present unpalatable options that amount to saving or sacrificing yourself — and the lives of fellow passengers who may be family members — to spare others, the researchers, not surprisingly, found that people would rather stay alive.

This particular dilemma of robotic morality has long been chewed on in science fiction books and movies. But in recent years, it has become a serious question for researchers working on autonomous vehicles who must, in essence, program moral decisions into a machine.

As autonomous vehicles edge closer to reality, it has also become a philosophical question with business implications. Should manufacturers create vehicles with various degrees of morality programmed into them, depending on what a consumer wants? Should the government mandate that all self-driving cars share the same value of protecting the greatest good, even if that's not so good for a car's passengers?

And what exactly is the greatest good?

"Is it acceptable for an AV (autonomous vehicle) to avoid a motorcycle by swerving into a wall, considering that the probability of survival is greater for the passenger of the AV, than for the rider of the motorcycle? Should AVs take the ages of the passengers and pedestrians into account?" wrote Jean-François Bonnefon, of the Toulouse School of Economics in France, Azim Shariff, of the University of Oregon, and Iyad Rahwan, of the Media Laboratory at the Massachusetts Institute of Technology.

At the heart of this discussion is the "trolley problem." First introduced in 1967 by Philippa Foot, a British philosopher, the trolley problem is a simple if unpleasant ethical thought puzzle.

Imagine a runaway trolley is barreling toward five workers on the tracks. Their lives can be saved by a lever that would switch the trolley to another line. But there is one worker on those tracks as well. What is the correct thing to do?

The research published in Science tries to quantify that philosophical quandary. "One missing component has been the empirical component: What do people actually want?" said Rahwan, who is a computational social scientist.

Each survey presented different situations, like varying the number of pedestrian lives that could be saved or adding a family member to the problem. In one survey they discovered that participants were generally reluctant to accept government regulation of artificial intelligence algorithms, even if that would be one way to solve or at least settle on an answer to this trolley problem.



A pedestrian crosses in front of a vehicle as part of a demonstration at MCity, used to test driverless and connected vehicles, on the University of Michigan campus in Ann Arbor, Mich., last July.

The number of respondents on the six surveys varied from 182 to 451.

The new research could take autonomous vehicle manufacturers down a philosophical and legal rabbit hole. And since the autonomous vehicle concept is so new, it could take years to find answers. For example, the authors write, "If a manufacturer offers different versions of its moral algorithm, and a buyer knowingly chose one of them, is the buyer to blame for the harmful consequences of the algorithm's decisions?"

The U.S. military is also trying to come to terms with the fact that advanced technology is on the cusp of making it possible for machines such as armed drones to make killing decisions. In 2012, the Pentagon released a directive that tried to draw a line between semi-autonomous and completely autonomous weapons. They are not outlawed, but they must be designed to allow "appropriate levels" of human judgment over their use.

In a companion article in Science magazine, the Harvard psychologist Joshua D. Greene suggested that the thorniest challenges in machine decision-making may be "more philosophical than technical. Before we can put our values into machines, we have to figure out how to make our values clear and consistent."

Some researchers argue that teaching machines ethics may not be the right approach.

"If you assume that the purpose of AI is to replace people, then you will need to teach the car ethics," said Amitai Etzioni, a sociologist at George Washington University. "It should rather be a partnership between the human and the tool, and the person should be the one who provides ethical guidance."

New York Times