

Will your self-driving car be programmed to kill you?

June 12 2015, by Matt Windsor



Imagine you are in charge of the switch on a trolley track.

The express is due any minute; but as you glance down the line you see a school bus, filled with children, stalled at the level crossing. No problem; that's why you have this switch. But on the alternate track there's more

trouble: Your child, who has come to work with you, has fallen down on the rails and can't get up. That switch can save your child or a bus-full of others, but not both. What do you do?

This ethical puzzler is commonly known as the Trolley Problem. It's a standard topic in philosophy and ethics classes, because your answer says a lot about how you view the world. But in a very 21st century take, several writers (here and here, for example) have adapted the scenario to a modern obsession: [autonomous vehicles](#). Google's self-driving cars have already driven 1.7 million miles on American roads, and have never been the cause of an accident during that time, the company says. Volvo says it will have a self-driving model on Swedish highways by 2017. Elon Musk says the technology is so close that he can have current-model Teslas ready to take the wheel on "major roads" by this summer.

Who watches the watchers?

The technology may have arrived, but are we ready? Google's cars can already handle real-world hazards, such as cars' suddenly swerving in front of them. But in some situations, a crash is unavoidable. (In fact, Google's cars have been in dozens of minor accidents, all of which the company blames on human drivers.) How will a Google car, or an ultra-safe Volvo, be programmed to handle a no-win situation—a blown tire, perhaps—where it must choose between swerving into oncoming traffic or steering directly into a retaining wall? The computers will certainly be fast enough to make a reasoned judgment within milliseconds. They would have time to scan the cars ahead and identify the one most likely to survive a collision, for example, or the one with the most other humans inside. But should they be programmed to make the decision that is best for their owners? Or the choice that does the least harm—even if that means choosing to slam into a retaining wall to avoid hitting an oncoming [school bus](#)? Who will make that call, and how will they decide?

"Ultimately, this problem devolves into a choice between utilitarianism and deontology," said UAB alumnus Ameen Barghi. Barghi, who graduated in May and is headed to Oxford University this fall as UAB's third Rhodes Scholar, is no stranger to moral dilemmas. He was a senior leader on UAB's Bioethics Bowl team, which won the 2015 national championship. Their winning debates included such topics as the use of clinical trials for Ebola virus, and the ethics of a hypothetical drug that could make people fall in love with each other. In last year's Ethics Bowl competition, the team argued another provocative question related to autonomous vehicles: If they turn out to be far safer than regular cars, would the government be justified in banning human driving completely? (Their answer, in a nutshell: yes.)

Death in the driver's seat

So should your self-driving car be programmed to kill you in order to save others? There are two philosophical approaches to this type of question, Barghi says. "Utilitarianism tells us that we should always do what will produce the greatest happiness for the greatest number of people," he explained. In other words, if it comes down to a choice between sending you into a concrete wall or swerving into the path of an oncoming bus, your car should be programmed to do the former.

Deontology, on the other hand, argues that "some values are simply categorically always true," Barghi continued. "For example, murder is always wrong, and we should never do it." Going back to the trolley problem, "even if shifting the trolley will save five lives, we shouldn't do it because we would be actively killing one," Barghi said. And, despite the odds, a [self-driving car](#) shouldn't be programmed to choose to sacrifice its driver to keep others out of harm's way.

Every variation of the trolley problem—and there are many: What if the one person is your child? Your only child? What if the five people are

murderers?—simply "asks the user to pick whether he has chosen to stick with deontology or utilitarianism," Barghi continued. If the answer is utilitarianism, then there is another decision to be made, Barghi adds: rule or act utilitarianism.

"Rule utilitarianism says that we must always pick the most utilitarian action regardless of the circumstances—so this would make the choice easy for each version of the trolley problem," Barghi said: Count up the individuals involved and go with the option that benefits the majority.

But act utilitarianism, he continued, "says that we must consider each individual act as a separate subset action." That means that there are no hard-and-fast rules; each situation is a special case. So how can a computer be programmed to handle them all?

"A computer cannot be programmed to handle them all," said Gregory Pence, Ph.D., chair of the UAB College of Arts and Sciences Department of Philosophy. "We know this by considering the history of ethics. Casuistry, or applied Christian ethics based on St. Thomas, tried to give an answer in advance for every problem in medicine. It failed miserably, both because many cases have unique circumstances and because medicine constantly changes."

Preparing for the worst

The members of UAB's Ethics and Bioethics teams spend a great deal of time wrestling with these types of questions, which combine philosophy and futurism. Both teams are led by Pence, a well-known medical ethicist who has trained UAB medical students for decades.

To arrive at their conclusions, the UAB team engages in passionate debate, says Barghi. "Along with Dr. Pence's input, we constantly argue positions, and everyone on the team at some point plays devil's advocate

for the case," he said. "We try to hammer out as many potential positions and rebuttals to our case before the tournament as we can so as to provide the most comprehensive understanding of the topic. Sometimes, we will totally change our position a couple of days before the tournament because of a certain piece of input that was previously not considered."

That happened this year when the team was prepping a case on physician addiction and medical licensure. "Our original position was to ensure the safety of our patients as the highest priority and try to remove these physicians from the workforce as soon as possible," Barghi said. "However, after we met with Dr. Sandra Frazier"—who specializes in physicians' health issues—"we quickly learned to treat addiction as a disease and totally changed the course of our case."

Barghi, who plans to become a clinician-scientist, says that ethics competitions are helpful practice for future health care professionals. "Although physicians don't get a month of preparation before every ethical decision they have to make, activities like the ethics bowl provide miniature simulations of real-world patient care and policy decision-making," Barghi said. "Besides that, it also provides an avenue for previously shy individuals to become more articulate and confident in their arguments."

Provided by University of Alabama at Birmingham

Citation: Will your self-driving car be programmed to kill you? (2015, June 12) retrieved 6 May 2024 from <https://phys.org/news/2015-06-self-driving-car.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.