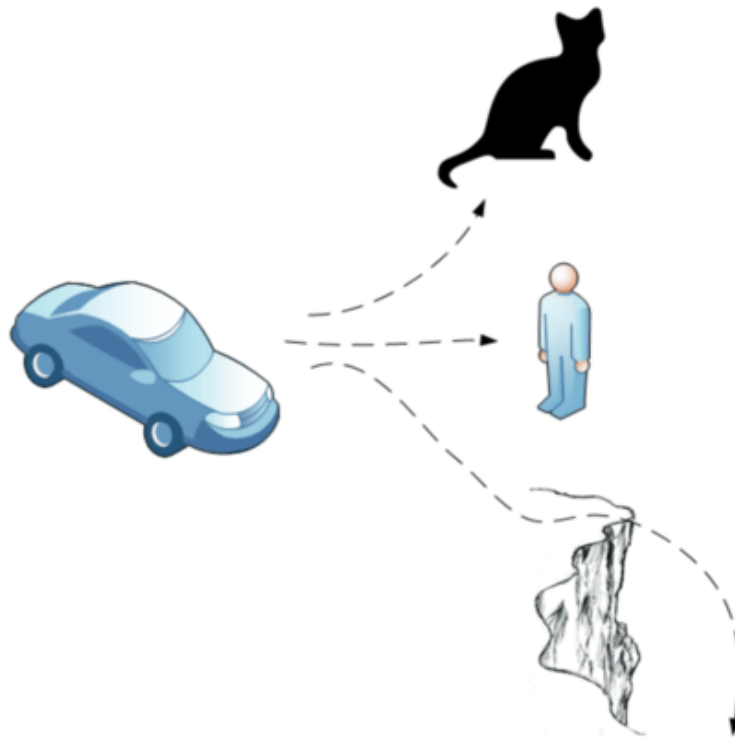


The ethical dilemmas of the driverless car

October 28 2015, by Bill Buchanan



Credit: Bill Buchanan, Author provided

We make decisions every day based on risk – perhaps running across a road to catch a bus if the road is quiet, but not if it's busy. Sometimes these decisions must be made in an instant, in the face of dire circumstances: a child runs out in front of your car, but there are other dangers to either side, say a cat and a cliff. How do you decide? Do you

risk your own safety to protect that of others?

Now that [self-driving cars](#) are here and with no quick or sure way of overriding the controls – or even none at all – [car](#) manufacturers are faced with an algorithmic ethical dilemma. On-board computers in cars are already parking for us, driving on cruise control, and could take control in safety-critical situations. But that means they will be faced with the difficult choices that sometimes face humans.

How to programme a computer's ethical calculus?

- Calculate the lowest number of injuries for each possible outcome, and take that route. Every living instance would be treated the same.
- Calculate the lowest number of injuries for children for each possible outcome, and take that route.
- Allocate values of 20 for each human, four for a cat, two for a dog, and one for a horse. Then calculate the total score for each in the impact, and take the route with the lowest score. So a big group of dogs would rank more highly than two cats, and the car would react to save the dogs.

What if the car also included its driver and passengers in this assessment, with the implication that sometimes those outside the car would score more highly than those within it? Who would willingly climb aboard a car programmed to sacrifice them if needs be?

[A recent study](#) by Jean-Francois Bonnefon from the Toulouse School of Economics in France suggested that there's no right or wrong answer to these questions. The research used several hundred workers found through [Amazon's Mechanical Turk](#) to analyse viewpoints on whether one or more pedestrians could be saved when a car swerves and hits a barrier, killing the driver. Then they varied the number of pedestrians

who could be saved.

	Weighting	Steer left	Straight Ahead	Steer right
Humans	10	3	0	1
Cats	4	1	5	3
Dogs	2	0	3	10
Horses	1	0	1	5
Score		34	27	47

Programmable ethics through a score card for driverless cars. Credit: Author provided

Bonnefon found that most people agreed with the principle of programming cars to minimise death toll, but when it came to the exact details of the scenarios they were less certain. They were keen for others to use self-driving cars, but less keen themselves. So people often feel a utilitarian instinct to save the lives of others and sacrifice the car's occupant, except when that occupant is them.

Intelligent machines

Science fiction writers have had plenty of leash to write about robots taking over the world ([Terminator](#) and many others), or where everything that's said is recorded and analysed (such as in [Orwell's 1984](#)). It's taken a while to reach this point, but many staples of science fiction are in the process of becoming mainstream science and technology. The internet and cloud computing have provided the platform upon which quantum leaps of progress are made, showcasing

[artificial intelligence](#) against the human.

In Stanley Kubrick's seminal film [2001: A Space Odyssey](#), we see hints of a future, where computers make decisions on the priorities of their mission, with the ship's computer HAL saying: "This mission is too important for me to allow you to jeopardise it".

Machine intelligence is appearing in our devices, from phones to cars. [Intel predicts](#) that there will be 152m connected cars by 2020, generating over 11 petabytes of data every year – enough to fill more than 40,000 250GB hard disks. How intelligent? As Intel puts it, (almost) as smart as you. Cars will share and analyse a range data in order to make decisions on the move. It's true enough that in most cases driverless cars are likely to be safer than humans, but it's the outliers that we're concerned with.

The author Isaac Asimov's famous [three laws of robotics](#) proposed how future devices will cope with the need to make decisions in dangerous circumstances.

- A robot may not injure a human being or, through inaction, allow a human being to come to harm.
- A robot must obey the orders given it by human beings, except where such orders would conflict with the First Law.
- A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

He even added a more fundamental "0th law" preceding the others:

- A robot may not harm humanity, or, by inaction, allow humanity to come to harm.

Asimov did not tackle our ethical dilemma of the car crash, but with greater sensors to gather data, more sources of data to draw from, and

greater processing power, the decision to act is reduced to a cold act of data analysis.

Of course software is notoriously buggy. What havoc could malicious actors who compromise these systems wreak? And what happens at the point that [machine intelligence](#) takes control from the human? Will it be right to do so? After all, in 2001, Dave has to take urgent action when he's had enough of HAL's decision-making:

Could a future buyer purchase programmable ethical options with which to customise their car? The artificial intelligence equivalent of a bumper sticker that says "I break for nobody"? In which case, how would you know how cars were likely to act – and would you climb aboard if you did?

Then there are the legal issues. What if a car could have intervened to save lives but didn't? Or if it ran people down deliberately based on its ethical calculus? This is the responsibility we bear as humans when we drive a car, but machines follow orders, so who (or what) carries the responsibility for a decision? As we see with improving face recognition in smartphones, airport monitors and even on Facebook, it's not too difficult for a computer to identify objects, quickly calculate the consequences based on car speed and road conditions in order to calculate a set of outcomes, pick one, and act. And when it does so, it's unlikely you'll have an choice in the matter.

This story is published courtesy of [The Conversation](#) (under Creative Commons-Attribution/No derivatives).

Source: The Conversation

Citation: The ethical dilemmas of the driverless car (2015, October 28) retrieved 29 April 2024

from <https://phys.org/news/2015-10-ethical-dilemmas-driverless-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.