**Need for Speed: Are we rushing towards universal machine ethics?**

The move towards self-driving cars is fast approaching, but our ethics are not ready. We don’t have a sufficient moral framework for programmers to follow when making self-driving cars. Policies regarding machine ethics ought to be like traffic laws: virtually universal, yet with regional differences that make them work well in different settings. It is always illegal to drive through a red light, and it is always illegal to run over a pedestrian. Globalisation means that the same universality needs to apply to ethics regarding self-driving cars: pedestrians and passengers must know that they are safe in every corner of the world, and programmers need a moral framework to apply when developing these systems for global use. Moreover, these systems should incorporate different cultural interests and approaches.

However, the difficulty of universal machine ethics can be shown through the findings of the Moral Machine experiment, where people were asked to decide who should die in self-driving car crashes. Participants were shown two potential courses of action and asked to choose the most favourable. Their answers, along with demographic information, were then collated and presented in an article in. The researchers grouped 130 countries into 3 cluster groups based on culture and geographical location: Eastern, Western, and Southern.

The user interface of the Moral Machine, showing two potential outcomes of a car crash.

Among these three clusters, there were significant differences in what was considered moral.This means that deciding who should die in a collision is not as simple as polling everyone and basing programming off their responses, as there are significant cultural differences. For example, if we prioritised the Eastern cluster’s broad opinion that pedestrians should be saved over passengers, Western countries would be unhappy. Similarly, if we programmed self-driving cars to spare those of higher economic status, tneither Western nor Eastern countries would be content. Moreover, this exercise effectively assigns a numerical value to each person’s life, according to their outward traits, and ranks them to determine who deserves to live and die. This is immoral, as it leads to treating people as a means to an end, creating a slippery slope towards a society that values people based on these characteristics. Public opinion cannot, and should not, be our only consideration.

The number of people saved in a collision, i.e., a utilitarian-esque approach, seems to be the natural next option, because distinction based on characteristics is problematic when striving for universalizability. However, simply saving the most people has issues of its own. It would be difficult to make universal because of the inherent bias we have for self-preservation. The Moral Machine experiment also highlighted this, as there was a tendency to want ownership of a car programmed to save its passengers, but for others to have cars that save the most people possible. It would be impossible then, to placate everybody. We could disregard public opinion and instead focus on what ethicists have to say, but that is also problematic. If we disregard public opinion completely in favour of simple, universalizable machine ethics, consumers will not buy self-driving cars.

There are other issues with listening solely to ethicists and field experts, as is demonstrated by the German Ethics Code for Automated and Connected Driving. Guideline 9 states that “any distinction based on personal features (age, gender, physical, or mental constitution) is strictly prohibited”. This shows the necessity of public opinion to this debate. It highlights that policymakers should be wary of backlash if they don’t give special status to children because of how heavily they were favoured in both the Western and Southern clusters in the Moral Machine. This then means that every group of people was favoured by at least one cluster, which invalidates the work of the German Ethics Commission.

As such, a balance needs to be struck between public opinion and the reflections of ethicists if we are to make self-driving cars that both function well and are able to be used globally, with cultural differences being understood. This balance can’t be achieved quickly, as rushing towards universal machine ethics leads to dilemmas like those stemming from the German Ethics Code and the Moral Machine. What these problems have done is illustrate how we must work together and to prioritise be caution over speed. After all, would you drive a car that somebody had programmed to kill you, if it meant you could drive it tomorrow?