Philosophy and Ethics C&IS Assignment 1
Implementing self-driving cars into everyday life

Menno Bartels s1007797
Lucas van der Laan s1047485
Marco Post s1046670
Elwin Tamminga s1013846
Ernst Hamer s1047547

November 2021

Introduction

Recently, the field of self-driving cars has become really popular [3, 2]. In the ideal world, the self-driving cars would transport humans from one location to another, without interferrence from the humans in the driving itself. Unfortunately, error-proof self-driving cars have not yet been developed, because of this there are still critical accidents that happen with self-driving cars [5, 6]. This is the central ethical issue of self-driving cars, and the choices the cars make during an accident. Regarding accidents that can happen, is it ethically justified to implement self-driving cars into everyday life? Do the pro's outweigh the cons? This is what we will discuss using various kinds of ethics.

Utilitarianism

The general opinion of utilitarianism is as follows: "The right action is the one that promotes the greatest happiness of the greatest number(maximizes social utility)". In this approach it is good to specify what happiness is, and where it comes from. When we look at the consequentialism approach: what are the results following from the action "using self-driving cars" that we classify as "promoting the greatest happiness". For now we regard the positive outcome "More safety". A well-functioning self-driving car should not be involved in any accidents at all. From a basic utilitarianism point of view this means that: if there are more cases of self-driving cars that are not involved in traffic accidents than cases that are involved, then implementing self-driving cars in everyday life would be an ethical thing to do. This standpoint does not take into account that the cases where no traffics accidents happen, may not weight up against the possible death in traffic accidents.

Ethical egoism

But what when an accident does occur? Ethical egoists would state that their vehicle within the incident should act in a way that it preserves its own well being first and foremost. Contrasting, utilitarianism states that the machine should act in a way that maximizes happiness for all participants, and thus the least injury overall. Assuming self-driving cars will also contain systems to handle accidents, the egoist would be in favor of having the machine act in such a manner that the injury of the occupant is minimized, only then should the machine minimize for total casualties. If every car does this however, a prisoners dilemma occurs, possibly making a very unpredictable situation. The common argument would be that this is still preferable over the Utilitarian approach, because the it's far more likely to be a participant in a accident than the sole instigator. Most accidents contains more than just two people, and thus it's morally justified to fend for yourself, if you're not the sole cause of the incident. Ultimately, if the individual has more positive outcomes from having an automatic vehicle, the pro's outweigh the con's and as much vehicles should be on the road as possible, as this increases personal happiness the most.

Kantianism

Unlike versions of consequentialism like utilitarianism and ethical egoism, if we want to apply Kantian ethics to self-driving cars we don't look at the consequences at all. Instead, the categorical imperative of Kant's ethics states the car should "act only according to that maxim whereby you can, at the same time, will that it should become a universal law" [4] and must treat rational agents as ends, not mere means. An example of a maxim rule is that a self-driving car should not injure anyone, because the universal rule is to not injure people. We cannot live in a world where all self-driving cars injure people, so it must be wrong. But when an accident occurs and the car has to make a choice between injuring an adult passenger or injuring a child on the road, then the car will not take action to prevent injuring the child, because according to Kantian ethics we are all rational being so a child is not "worth" more than the passenger. So a self-driving car with Kantian behaviour will only take action if it complies with universalizable rules, without looking at the consequences.

Aristotelianism

Aristotle believed in that humans are special because they are rational creatures and have a common understanding of good, which could be generalized as "common sense". Aristotle also believed that every human has the potential to be good and can learn to be good based on practice and upbringing. This is not done by learning general rules, but instead by learning and applying practical wisdom, which was Artistotle's most important Virtue. Applying this to self-driving cars means that we need to look into what we have learned about

people driving cars and about the self-driving cars that have been testing already. We know that humans make mistakes, because they are tired, because they are bored and not paying attention, or simply because they are not skilled in driving. This leads to a simple answer, would taking away the responsibility for humans to drive a car lead to fewer accidents, and if they happen, less sever accidents? If this is the case, this a then "common sense", a well understood fact of good and then not implementing self-driving cars would be bad, since that decision would lead to more harm.

Confucianism

The golden rule of the ancient Chinese belief system Confucianism is "Do not do unto others what you would not want others to do unto you.". Confucianism is considered a humanistic approach and focuses on virtues and ancestor worship. Here we explore this approach to ethics when translated to a more modern issue, as the question if it is ethically justified to implement self-driving cars into everyday life. It was found by Edmond Awad, et al (2016) [1] that when faced with ethical dilemma's regarding self-driving cars there are profound differences between the western countries and eastern countries, where Confucian values are far more prominent. The Confucianist cultural groups in this study strongly preferred saving pedestrians over car passengers. This is reflected in the Confucianism ethic upholding of righteousness and the moral disposition to do good. So when self-driving cars are implemented on a grand scale, it would be best from a Confucianism point of view to program the car to favor pedestrians /other people over the passengers in equal situations. From this point of view it would be an ethical choice to implement self driving cars into everyday live because machines would uphold Confucian values to an ever greater extent and consistency than human drivers ever could and maybe willing to do. Although the extremes in this situation could be undesirable for some, thinking in this direction could help set the socially acceptable principles and boundaries for machine ethics in self-driving cars.

Conclusion

We have looked at the topic of implementing self-driving cars into everyday life. From the several types of ethics we have discussed we see that there are many choices a self-driving car can make during accidents and that it depends on the ethical perspective of society if it is ethically justified. We think that "self-driving cars handle well in accidents" and "self-driving cars are more skilled than human drivers", but one can argue that such assumptions are not always true, and that in reality we value human lives more than the given that "a lot of people" get to their destination safely. So there is not just one ethical perspective that is suited for self-driving cars. Other ethical questions concerning self-driving cars could have different answers. However, Some questions that must be answered before we can ethically apply self-driving cars could be: Who

is responsible when two self-driving cars are in an accidents? What are the consequences in the divide between richer and poorer cars owners, can only rich people afford self-driving cars, and is travelling therefore safer for richer folks, and is that fair? Many more of these complex ethical questions must be answered. To achieve this we must not only grow our technologically capabilities, but also our moral capabilities.

References

[1] Edmond Awad, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. The moral machine experiment. *Nature*, 563(7729):59–64, 2018.

[2] Göde Both. What drives research in self-driving cars? (part 1+2). 04 2014.

[3] Noor Zainab Hussain REUTERS GRAPHICS. American drivers warming to self-driving cars - survey, 2019.

[4] Marcus G Singer. The categorical imperative. *The philosophical review*, 63(4):577–591, 1954.

[5] Kelsey Piper VOX. It's 2020. where are our self-driving cars?, 2020.

[6] ore Petrović, Radomir Mijailović, and Dalibor Pešić. Traffic accidents with autonomous vehicles: Type of collisions, manoeuvres and errors of conventional vehicles' drivers. *Transportation Research Procedia*, 45:161–168, 2020. Transport Infrastructure and systems in a changing world. Towards a more sustainable, reliable and smarter mobility.TIS Roma 2019 Conference Proceedings.