

FINAL SYNOPSIS

//THESIS STATEMENT

This paper will aim to explore the ethical complications of autonomy in cars. The question of accountability of actions will be raised, in which the notion of authorship will be explored, as a result of the degradation of the role of the human being in driving.

//WHY ARE ETHICS IMPORTANT?

Driving a car entails making lots of decisions constantly, at a speed which is inherently unnatural for human beings. Often, these decisions are simple ones: whether or not to change lanes, for example. But sometimes, decisions have to be made which might concern life or death for either those in the car, or people outside of it. These decisions can be ethically challenging, and the operator of the car will have to make them based on his/her moral compass. These are decisions, which in the future will have to be made by the car. If we as human beings are to be degraded from operator of the car, to merely being passengers, the car is faced with a myriad of ethical questions when driving, for which it has to act ethically correct. But what does this entail?

In the field of autonomous vehicles, a famous example is that of the trolley example. In this, we find ourselves being at the command of a trolley on tracks, we are driving around a corner, on the other side of which we find ourselves in a predicament. Unexpectedly, five men are working the tracks ahead. We slam the brakes, but it quickly becomes apparent that these have failed – inevitably, we are going to hit, and consequently kill the men at work. Suddenly we discover a spur of tracks, taking us away from the main trail. It is possible for us to switch to this trail of tracks in order to avoid hitting them – however, there is a single man working on that section, who we are going to kill, should we take this route.

This is an excellent example of a moral/ethical predicament. None of the options are ethically permissible, but one has to be chosen. Transfer this to the world of autonomous vehicles, self-driving cars need to be equipped with a moral codex, an understanding of ethics, in order for them to make ethical decisions in difficult circumstances such as that of the trolley example. This raises some important questions: how do we decide what the ethically correct action for the car to perform is, when none of the outcomes can be seen as ethically correct? What moral code should the car act upon?

The paper will aim to highlight the importance of ethics in autonomy, by exploring what we think of as being ethically permissible in circumstances such as that of the trolley example. Furthermore, it

will seek to discuss the ethical challenges which is raised by the implementation of autonomous vehicles, and to remain critical in relation to the new technology. The baseline for the discussion will be *Why Ethics Matters for Autonomous Cars* by Patrick Lin¹, and Entanglement: *Machine learning and human ethics in driver-less car crashes* and Maya Indira Ganesh².

//A QUESTION OF ACCOUNTABILITY

As human beings, when we drive, we are solely responsible for our actions – we can be held accountable for our actions. Leading up to an accident, it can be argued, that it is impossible for an individual to calculate which option is the best, in the short space of time it has to act in. Thus, we can perhaps describe the outcome as a reactionary result, not a conscious decision. An autonomous car, however, has the computing power to calculate all of this in a split second. This means, that the actions of the cars cannot be looked upon as reaction – they are decisions made based on data and are acted out in relation to a predetermined understanding of morals programmed into the computer.

Should an accident then happen, who can be held accountable? In Walter Benjamins *The Author as Producer*³, he discussed the notion of the author as a producer when creating (in his case, writing). This notion of authorship becomes interesting in the field of software, when the software we create starts creating something for itself – as it is true of machine learning, and indeed autonomy in vehicles. If/when an accident happens, the notion of authorship then becomes an interesting perspective: is the programmers at fault, or is the car? The car might encounter situations that far exceeds those imagined by the programmers, and how it acts might be completely unexpected to the programmers, even though the actions of the car was based on the code they have written.

However, the above is true only for a fully autonomous car, of which there is none on sale in the world today. These cars are expected to do all of the driving themselves, while the people it is carrying is never expected to take control at any time. What we have on the road today can be classed as limited self-driving automation, cars that require the driver to be alert and ready to assume control of the car should it be necessary. This blurs the question of accountability even further, as has been exemplified by the recent accidents involving autonomous cars in the USA. Recently, a Tesla Model X crashed while in autonomous mode on the highway in California, taking the life of the driver. Questions has since been raised of whether the autonomous system was at fault, or whether it was the driver who failed to act as he should have. This case exemplifies the

¹ Lin P. (2016) *Why Ethics Matters for Autonomous Cars*. In: Maurer M., Gerdes J., Lenz B., Winner H. (eds) *Autonomous Driving*. Springer, Berlin, Heidelberg

² Entanglement: *Machine learning and human ethics in driver-less car crashes* (2017) by Maya Indira Ganesh <http://www.aprja.net/entanglement-machine-learning-and-human-ethics-in-driver-less-car-crashes/>

³ Walter Benjamin, "The Author as Producer," in *New Left Review*. 1/62, 1970 [1934]

questions which can be raised about accountability of actions in semi-autonomous vehicles, and I intend to use it as a baseline for further analysis on the importance of ethics in driving.