

SUBMISSION OF WRITTEN WORK

Class code: **1410001U**

Name of course: **Reflections about IT**

Course manager: **Judith Simon**

Course e-portfolio: <https://learnit.itu.dk/course/view.php?id=3005237>

Thesis or project title:

Supervisor:

Full Name:

Anders Fischer-Nielsen

Birthdate (dd/mm-yyyy):

06/05-1993

E-mail:

afin

- | | | |
|----------|-------|--------------|
| 1. _____ | _____ | _____@itu.dk |
| 2. _____ | _____ | _____@itu.dk |
| 3. _____ | _____ | _____@itu.dk |
| 4. _____ | _____ | _____@itu.dk |
| 5. _____ | _____ | _____@itu.dk |
| 6. _____ | _____ | _____@itu.dk |
| 7. _____ | _____ | _____@itu.dk |

Who Should Your Car Kill? Ethics and Autonomous Vehicles

Anders Fischer-Nielsen

May 27th 2016

Contents

Contents	1
1 Introduction	2
2 Autonomous Cars	3
2.1 Google's Autonomous Car Project	3
2.2 Motivations for Developing Autonomous Vehicles	4
3 General Ethics and Computer Ethics	5
3.1 Consequentialism	5
4 Ethics of Autonomous Vehicles	6
5 Discussion	7
6 Conclusion	8
Bibliography	9

Chapter 1

Introduction

The idea of creating mass-produced self-driving, or autonomous, vehicle has seen more and more interest from researchers and car manufacturers in the last few years.

The idea of inventing an autonomous car is not a new thing with people having dreamt up future societies where people would focus on other things while their cars drove them from A to B. In the last decade these ideas are finally being realised with research projects such as Google's ongoing Self Driving Car Project, previous research projects such as the entries of Carnegie Mellon and Stanfords in the DARPA Grand Challenge and the recent undertakings of most car manufacturers, such as BMW, Audi, Toyota, VW etc.

cite.

cite.

Certain ethical issues arise with autonomous vehicles, however. Since autonomous vehicles are programmed to handle preferably every possible scenario on the road in advance, autonomous vehicles present ethical dilemmas that human drivers in non-autonomous vehicles do not.

Human drivers are rarely blamed for acting according to their instinct in life-threatening vehicle collisions, even if this means that they inadvertently make decisions that have fatal consequences for other people.

These decisions have to be programmed into autonomous vehicles in advance in order to make the "right" choice, should they occur. This raises interesting questions, such as; who should be blamed for an autonomous vehicle hitting, and possibly killing, one person over another? Should we even program autonomous vehicles to be able to select the "preferable" target of collision, or would it be better to not do any selection at all and make the vehicle choose some random behaviour in a critical situation?

The main question I will try to answer in this essay is; "Who should determine who your future autonomous car hits in an emergency situation?" by looking into what ethical issues this question presents, and what theories can help answer this question.

Chapter 2

Autonomous Cars

Experiments attempting to automate vehicles, mainly cars, have been made since the 1920's with varying degrees of automation and success.

This essay will focus on *autonomous* cars, not just automated cars. Autonomous is, according to Thesaurus, defined as:

cite

“an autonomous republic: self-governing, independent, sovereign, free, self-ruling, self-determining, autarchic; self-sufficient.”

Automation implies that cars merely follow artificial hints in the environments, such as early experiments using magnetic strips in the road. *Autonomous* implies that cars react to their environment independently, that is they cannot depend on unnatural artefacts in their surrounding environment in order to drive properly.

Modern research has been focused on the latter, since it is unrealistic to add artificial hints on every road in the world. Rather, research has been focused on making autonomous cars adapt to environments with uncertainties, so that the same autonomous car can drive in the inner city and on mountain roads. I will therefore not look at automatic vehicles in this essay, but instead autonomous vehicles.

Recent research projects use radar or radar-like technology in addition to GPS, odometers and computer vision in and on cars to detect the environment surrounding the car and recognise obstacles, such as people, other cars and structures, which the car will try to avoid. The addition of the radar-like LIDAR technology (a mix of the words light and radar) on the roof of the cars has provided the cars with a 200 foot-radius "view" of their surroundings, enabling them to sense the world around them in great detail.

2.1 Google's Autonomous Car Project

Google has been researching autonomous cars since 2009 . Over the years the project has been ongoing, the cars have developed a detailed view of their surroundings. In addition to the computer vision research the company has developed and implemented in the cars, LIDAR helps the car map its surroundings, enabling it to recognise smaller objects such as pedestrians and bicyclists.

cite.

2.2 Motivations for Developing Autonomous Vehicles

The main motivation for developing autonomous vehicles is that human drivers are prone to make mistakes. Human drivers get distracted, have relatively slow reaction times, do not always behave logically such as when angry or tired. Furthermore, human drivers sometimes drink while under the influence of drugs, which severely inhibits the driver's ability to react in time to avoid collisions.

Labor costs of human drivers can also be eliminated, traffic jams might be reduced, and vehicles might be able to park more efficiently, drive faster, and occupants of the car might have time to be more work while driving.

In short, a computer driving your car will probably be a better driver than you. An autonomous car can sense its surroundings 200 times per second and make just as many calculations reasoning for its next move based on the input. A human driver will never top that. Your computer will also never get tired or drunk and make mistakes because of that.

This fact presents some interesting questions. Given the amount of input and processing power, an autonomous car should always be able to make the best choice possible. An autonomous car would register the child running across the road before a human driver ever would, and should therefore always make the right choice accordingly.

But what *is* the right choice? Most people would say that the car should always try to harm as few people as possible. That would be the "ethically correct thing to do". But what happens in situations where someone *has* to get hurt? If an autonomous car is in a situation where it has the option of hitting two different people, but no option to avoid either one, who should it choose? Who should be to blame for any collisions or pay for any harm done to people or structures?

Answering these questions requires us to look at other ethical arguments described in the following sections.

Cite study
med ran-
dom kø i
alle trafik-
kredse -
YouTube?

Chapter 3

General Ethics and Computer Ethics

Computer ethics, and generally ethics of technology, is evolving along with the corresponding technologies of the time. This is one of the main problems of computer ethics, and also why it is important to discuss. As Moor says:

“Computers provides us with new capabilities and these in turn give us new choices for action. [...] A central task of computer ethics is to determine what we should do in such cases, i.e., to formulate policies to guide our actions.”

cite.

3.1 Consequentialism

In the study of normative ethics, consequentialism holds that the consequences of one's conduct are the basis of which to determine the rightness or wrongfulness of one's actions. That is, the means to which you achieve your goal pose no ethical relevance, rather the end result is what has ethical relevance. It is the idea that the end justifies the means. [1]

“Every advantage in the past is judged in the light of the final issue. — Demosthenes”

3.1.1 Utilitarianism

Utilitarianism is a form of consequentialism, founded by Jeremy Bentham, holding that the best moral action is the one that maximises *utility*, that is the well being of sentient beings. The action that maximises the well being and minimises the suffering of humans is the one most ethically correct.

Chapter 4

Ethics of Autonomous Vehicles

According to utilitarianists, no matter the ethical issues involved in possible critical situations involving autonomous cars, having autonomous cars will always be ethically sound, since autonomous cars, even with many critical situations, will save more lives overall. Even if the act of selecting individuals to hit if it is unavoidable, the end result - drastically lowering the current amount of people killed in accidents - will maximise the well-being of humans. Therefore developing autonomous cars, as long as this saves more lives, will be a good thing.

Chapter 5

Discussion

“On my view, computer ethics is a dynamic and complex field of study which considers the relationships among facts, conceptualizations, policies and values with regard to constantly changing computer technology. Computer ethics is not a fixed set of rules which one shellacs and hangs on the wall. - Moor”

Chapter 6

Conclusion

Bibliography

- [1] J. Mizzoni. *Ethics: The Basics*. John Wiley & Sons, 2009.