Table of Content:

1. Introduction

What was only seen in science fiction movies and comics a few decades ago has become reality within the last few years. Autonomous vehicles that move by themselves or only under the observation of a human eye are now a reality in the form of self-driving cabs, self-driving container transporters or even in the form of individual passenger cars. The first autonomous vehicle dates back to the year 1926, which was still significantly different from our current understanding of autonomous vehicles. It was the radio controlled car, the Linriccan Wonder, which was introduced by Houdina Radio Control in New York City (Bimbraw K. ,2015).

Since then, a lot has happened in the field of autonomous driving and new innovations & developments have pushed autonomous driving further and further, so that today we are already talking about the possibility of truly self-driving and decisive vehicles. The autonomy, or self-governance of these vehicles, however, naturally also leads to the fact that since the beginning of autonomous driving, the question has been asked how vehicles would behave in situations in which even we humans often do not know what we should do and act according to instinct. It is precisely about the question of how autonomous vehicles would behave especially in the case of danger for the driver as well as other road users.

In the case of autonomous driving, the problem can often be described in terms of a moral dilemma (Nyholm, 2018), in which a decision has to be made between several equally unacceptable or at least unpleasant alternatives in a situation, which may also be mutually exclusive, comparable to the famous Trolley problem (Thomson, 1976, Chapter 61). Now, in the case of autonomous cars, this decision does not have to be made by a person, but by his vehicle and the underlying program. In the case of autonomous driving, this dilemma consists primarily of the question of how the vehicle will behave in a dangerous situation that may also involve other people beyond the driver (Gerdes & Thorton, 2015). The typical examples are choosing between a small child and an elderly woman, driving into a crowd of people or the wall, or even just endangering the driver himself by avoiding obstacles such as animals or the like. (TEDx Talks, 2016)

It therefore becomes clear quite quickly that acting within the framework of autonomous vehicles entails its own problems. It is therefore necessary to deal in detail with the subject of the actions of autonomous vehicles in order to understand their underlying principles and to be able to comprehend the action of an autonomous vehicle.

First of all, however, one should probably ask the question whether an autonomous vehicle is capable of acting at all? In principle, decision-making in an autonomous vehicle is only dependent on the code or program that is set up. This would mean that ultimately it is a program that decides how an autonomous vehicle should behave, and not a human entity. Does that still make it an action? To find an answer to this question, let's take a closer look at the principles and understandings of actions in the context of autonomous vehicles.

This essay aims to help to understand & name the difficulties and limitations of understanding actions in the context of autonomous vehicles. For this purpose, the difference between actions and operations will be clarified, the different stages of autonomous driving will be explained, and on the basis of this, the role of the body in the context of autonomous vehicles will be discussed. Subsequently, it will be shown on the basis of which philosophical theories of action a program, and thus also autonomous vehicles, are developed. Finally, these results will be summarized and classified in view of the current state of research and discussion, also including open questions.

2. Limitations in the understanding of actions in the sphere of autonomous vehicles

2.1 Operations and Actions

When we talk about the action of autonomous vehicles, a distinction can be made between an action and an operation, which is among the first obstacles in understanding the action of an autonomous vehicle.

As already described, the decision to be executed in our case is often based on a code or program. Taking that into consideration, this brings several factors with it that significantly influence the understanding of actions in the context of autonomous vehicles. But it should also be said that since it's inception the philosophy of action has been more concerned with a consideration of actions of human actors & not of machines or robots.

First of all, it should be noted that the written code on which the action is ultimately based comes from a human (Gerdes & Thorton, 2015). This means that every decision that the vehicle carries out in the end comes from human hands. This consideration allows the assumption that every action of an autonomous vehicle is also an action in its original sense, just via a new form of action.

However, it can also be argued that an action performed by a robot or machine, and an autonomous vehicle is nothing else, no matter how intelligent it seems to be, but an operation. This is the execution of an action by a system or a program with an unconscious action.

> *"As nouns the difference between operation and action is that operation is the method by which a device performs its function while action is something done so as to accomplish a purpose."(WikiDiff, 2022)*

Accordingly, the "action" of a machine that is methodically designed only to perform various "normal" actions can also be understood as an operation.

In the general field of action philosophy an action is understood in most cases, as an event, which an actor carries out for a certain purpose, which is led by the intention of the person (Stanford Encyclopedia of Philosophy, 2022). In the case of autonomous vehicles, however, there is no longer an actor in the original sense and the driver's intentions may have no influence on the final decision.

2.2. Levels of autonomous Driving

In order to be able to differentiate between actions and operations, it is absolutely necessary to take a look at the different levels of autonomous vehicles. Today's autonomous vehicles are divided into 5 levels, where the relationship between human and machine is different for each level. With each higher level (0-5), the vehicle takes over more and more autonomous tasks, which completely detach the human from the vehicle and its decisions (Meyer, Müller & Dokic, 2015), (SEA, 2022), (NHTSA, 2022).

- Level 0 : The person in the vehicle controls the vehicle completely independently.
- Level 1: A first set of support aids accompanies the person in the vehicle. However, these aids can only be performed individually.
- Level 2: An advanced form of support aids accompanies the person in the vehicle. However, it remains to be noted that the person must continue to sit behind the wheel with full attention.
- Level 3: This is the first level at which the vehicle can drive completely independently in special situations. However, the person behind the wheel must still be absolutely alert and should be able to take over the wheel again at any time.
- Level 4: A class 4 autonomous vehicle is a vehicle that can really participate in road traffic on its own, without the need for a person to monitor or intervene.
- Level 5: In the case of the last level of autonomous vehicles, we have arrived at a really fully automatic self-driving car, in which the human being only acts as a passenger and can no longer intervene at all.

But what does this mean for the understanding of actions and operations of autonomous vehicles? Well, first of all, only in the case of a level 5 car can the vehicle be granted complete decision-making power. Before that, the human being is still part of the decision-making process, can actively and passively intervene in the events and thus determine them.

Consequently, it would be wrong to speak of actions as operations if it is at least not a level 5 vehicle. Only they make decisions purely on their code or their programming. In all other cases, the final action is still that of a human being or at least influenced by.

For demonstration purposes, this essay assumes that in practice, or everyday life, we are already dealing with level 5 autonomous vehicles whose decisions are program-based. In reality, our society is still very far from the development & use of a Level 5 vehicle for people (Oliver, 2021). Nevertheless, already these theoretical forms of autonomous driving present us with completely new challenges in the understanding of action.

2.3 Role of the Body

The decisive factor in this subdivision is also the consideration of the body. What role does the body play in the context of autonomous vehicles? After all, the ultimate goal of autonomous vehicles is to remove people, and thus the body, from road traffic and to eliminate people as a source of error (NHTSA, 2022). However, it has just been shown that the body is still needed for many of the vehicles today and also determines a large part of the actions. In this case the body is still a clear part of the action, whereby this can also be described as an action.

But what about when we consider a level 5 autonomous vehicle? The body and the vehicle that ultimately takes the "action" are completely detached from each other. First and foremost, the human being no longer has any decision-making power at all. The “action“ is then only dependent on the program, and in order to understand an action of a program, one must take a closer look at which philosophical principles a program decides for or against an action, or which philosophical principles can be built into a program at all.

2.4 The program

Basically, a code can only be characterized and built up on by two of the three main theories from normative Ethics, Deontology, Consequentialism and Virtue Ethics. All theories deal

with the question about how actions can be described and explained & and only the two theories, Deontology and Consequentialism out of the three theories of normative ethics, can be applied to a code or program. (Gerdes & Thorton, 2015), (Wernaart, 2021)

2.4.1 The Program - Deontology

In deontological ethics, an action is considered morally good, or bad, based on some property of the action itself (Stanford Encyclopedia of Philosophy, 2022). These properties can be defined in the case of a code simply in the form of variables and built into the code. Thus, every decision that the code, or even the machine, makes is based on the defined properties. However, this would also mean that some actions could always be morally impermissible, regardless of the consequences (Britannica, 2022). To better understand the principle of deontology in the case of autonomous driving, it helps to consider a small example.

Let's take the classical example, the decision between the death of a young child & the old grandma. A deontological code would be defined and would protect the person who is fixed in its code as a person worth protecting. Provided that children would be classified from the principle as more valuable than older persons, would be tried to save the child, no matter what possible consequences stand on the other side.

2.4.2 The Program - Consequentialism

In contrast to dentology, consequentialism explains actions primarily in terms of their outcome, according to the well-known maxim, the end justifies the means. The result of an action decides whether it was ultimately bad or good (Stanford Encyclopedia of Philosophy, 2022). Related to a code or a program this is probably the most understandable variant to implement. A code is ultimately only a calculation of ones and zeros and also the Consequentialism is ultimately only based on calculations (McNaughton & Rawling, 2022). These calculations do not necessarily have to occur in the typical numbers, but in the fact that certain things can be given a value and calculated with it.

So, in the case of Consequentialism, and especially with respect to autonomous vehicles, the result would determine the actual decision of the vehicle in any case. Thus, a calculation would determine whether, in the same example from just now, it would be more beneficial to let the child, or just the grandma live. In this example, the grandma would probably draw the short straw, since the contribution that the child can still make, or the years it would still live, is a lot more than in the case of the grandma, but that depends on what would ultimately be considered as a desirable consequence.

However, the principle of Consequentialism has a major problem. In order to make a decision based on the results of a situation, absolutely complete data is required. In practice, however, this is not even possible, especially with actions concerning future events (Card & Smith, 2020). This means that every decision made on the basis of Consequentialism is made on the basis of incomplete data, and especially in the case of autonomous driving this should not be disregarded in view of the fact that the lives of road users are at stake. It must be said, however, that even without complete data, the desired results can be achieved in practice.

2.4.3 The Program - Delimitation of virtue ethics

In addition to the two theories of action just presented, general normative ethics often includes also virtue ethics (Solomon, 1995). However, in contrast to the other two approaches, this one can hardly be transferred into a code. While both rules and consequences can be incorporated into a code, this is even more difficult in the case of virtues or moral character (Wernaart, 2021).

The numerical structure of a code does not allow such complex relations, which can not be simplified by a calculation or formula & are probably based on non-numerical factors. Thus a code or a program will never be able to decide and act really like a human being. This is also one of the major problems in the development of artificial intelligence and thus also autonomous vehicles, as it is a limit that can probably never be overcome.

3. Conclusion

The limits of the scope of digitalization and mechanization often seem infinite. However, machines and their programs repeatedly come up against the fundamental limits of physics or even philosophy.

It can be stated that the understanding of actions or operations in the context of autonomous vehicles is not a simple topic and has already caused a lot of furor. Not least by the question whether such an action should be seen at all still as an action or better as an operation. In the course of this work it was explained that the limits of understanding are mainly related to the program itself, or rather to the decision theory on which they are based. It was discovered that a program can only be written based on deontology as well as consequentialism and not in the course of virtue ethics. In addition, however, the different five levels of the autonomous vehicle also influences the extent of the driver's participation in an “action“ or decision.

The research being conducted in this area is very little and also quite new. Furthermore it must necessarily adapt to ever-changing innovations and changes. In the end, many of the research questions are designed to find out how autonomous vehicles can be better optimized for users than to understand their real scope. Because of the issues surrounding the actions of autonomous vehicles, some works advocate, for example, a "Big Red Button" (Gerdes & Thorton, 2015), in which humans are still left with ultimate decision-making power, or even a software engineering solution that takes ethical and social considerations seriously (Holstein. Dodig-Crnkovic & Pelliccione, 2018)

However, when thinking about the understanding of actions and operations of autonomous vehicles, and considering the various positions, a lot of open questions remain.

For example, what changes the understanding from an action to an operation? Would an operation be evaluated differently? How would an accident be evaluated where there was actually no longer an action in the original sense? And, of course, there is also the question of who should ultimately be held responsible in the event of an accident (Li, J. et al, 2016).

Can a human being be responsible for an action that he did not perform? Or can a machine be held responsible at all? Or is it possible to hold the company responsible in the background? These are questions of law and politics, but they should be discussed on the basis of the theories just presented in the current emerging world full of autonomous vehicles. Because, never before has our society been in such a digitalized and mechanized time of change.

The limits to the actions of autonomous vehicles described in this work naturally also have a significant influence on human actions. The question arises: “How does a machine influence our human actions?“. This question is addressed by both the fields of cybernetics (Spektrum, 2022) and post-humanism (Huhtamo 2020), once with a view to technology, the other time with a view to humans. However, in order to understand the basic framework of action of autonomous vehicles, or even other technical innovations, a deeper examination of the limits of understanding and the scope of action is required as we saw it in the case of autonomous vehicles. Because it seems like we're just developing something that we don't even fully understand.

4. Bibliography

Bimbraw K., (2015), *Autonomous Cars: Past, Present and Future: A Review of the Developments in the Last Century, the Present Scenario and the Expected Future of Autonomous Vehicle Technology*, 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO-2015), pages 191-198, DOI:10.5220/0005540501910198

Britannica, *deontological ethics,* [Online]. Available at: https://www.britannica.com/topic/deontological-ethics (Accessed 31.01.2022)

Card D., & Smith N. A., (2020), *On Consequentialism and Fairness*, Frontiers in Artificial Intelligence , May 2020 , Volume 3 , Article 34, DOI: https://doi.org/10.3389/frai.2020.00034

Gerdes J. C. & Thorton S. M., (2015), *Implementable Ethics for Autonomous Vehicles,* DOI: 10.1007/978-3-662-48847-8_5, in M. Maurer et al. (Hrsg.), *Autonomes Fahren*

Huhtamo E., (2020), *The Self-Driving Car: A Media Machine for Posthumans?* In: Andrés Burbano; Ruth West (coord.) «AI, Arts & Design: Questioning Learning Machines». Artnodes, no. 26: 1-14

Holstein T., Dodig-Crnkovic G. & Pelliccione P., (2018), Ethical and Social Aspects of Self-Driving Cars, IEEE Transactions on intelligent transportation systems, Vol. 18, No. 6

Li, J., Zhao, X., Ju, W., & Malle, B. F., (2016), *From Trolley to Autonomous Vehicle: Perceptions of Responsibility and Moral Norms in Traffic Accidents with Self-Driving Cars,* Conference Paper April 2016, SAE 2016 World Congress and Exhibition, DOI: 10.4271/2016-01-0164

McNaughton D. & Rawling P., (2022), *2. Criticisms of act-consequentialism,* 10.4324/9780415249126-L013-1 [Online]. Available at: https://www.britannica.com/topic/deontological-ethics (Accessed 31.01.2022)

Meyer G., Müller B. & Dokic J., (2015), *European Roadmap, Smart Systems for Automated Driving,* EPoSS - European Technology Platform on Smart Systems Integration, Version 1.2

NHTSA, (2022), *Automated Vehicles for Safety* [Online]. Available at: https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety (Accessed 31.01.2022)

Nyholm, S. (2018), *The ethics of crashes with self‐driving cars: A roadmap, I*, Philosophy Compass 13(6293):e12507 DOI:10.1111/phc3.12507

Oliver J., (2021), *3 Reasons We Haven't Reached Level 5 AV Deployment*, Embedded Computing Design, [Online]. Available at: https://www.embeddedcomputing.com/application/automotive/adas-autonomous-drive/3-reasons-we-havent-reached-level-5-av-deployment (Accessed 31.01.2022)

SEA, (2022), *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, [Online]. Available at: https://www.sae.org/standards/content/j3016_201806/ (Accessed 31.01.2022)

Solomon D., (1995), *Ethics: III. Normative Ethical Theories,* [Online]. Available at: https://www.encyclopedia.com/science/encyclopedias-almanacs-transcripts-and-maps/ethics-iii-normative-ethical-theories (Accessed 31.01.2022)

Stanford Encyclopedia of Philosophy, (2022), *Action,* [Online]. Available at: https://plato.stanford.edu/entries/action/ (Accessed 31.01.2022)

Stanford Encyclopedia of Philosophy, (2022), *Consequentialism,* [Online]. Available at: https://plato.stanford.edu/entries/consequentialism/ (Accessed 31.01.2022)

Stanford Encyclopedia of Philosophy, (2022), *Deontological Ethics,* [Online]. Available at: https://plato.stanford.edu/entries/ethics-deontological/ (Accessed 31.01.2022)

Spektrum, (2022), Kybernetik, spektrum.de, Lexikon der Neurowissenschaften [Online]. Available at: https://www.spektrum.de/lexikon/neurowissenschaft/kybernetik/6831 (Accessed 31.01.2022)

TEDx Talks, (2016), *The Social Dilemma Of Driverless Cars | Iyad Rahwan | TEDxCambridge,* TedTalk Youtube Chanel [Online]. Available at: https://www.youtube.com/watch?v=nhCh1pBsS80&t=110s (Accessed 31.01.2022)

Thomson J. J., (1976), *Ethical Theory, An Anthology*, Second edition, edited by Russell Shafer-Landau, Wiley-Blackwell

Wernaat B., (2021), *Developing a roadmap for the moral programming of smart technology,* Technology in Society. Volume 64, February 2021, https://doi.org/10.1016/j.techsoc.2020.101466

WikiDiff, (2022), *Operation vs Action - What's the difference?,* [Online]. Available at: https://wikidiff.com/operation/action (Accessed 31.01.2022)