International Journal of Advanced Research and Publications
ISSN: 2456-9992

Autonomous Vehicles: Ethical Dilemmas

Maulik Patel, Rushit Dave, Evelyn R Sowells Boone

Macquarie University,
Information Technology, Sydney, Australia
maulik-nareshkumar.patel@students.mq.edu.au

North Carolina A&T State University,
Department Of Computer Science, Greensboro, USA
rrdave@aggies.ncat.edu

North Carolina A&T State University,
Department of Computer Systems Technology
Es297760@ncat.edu

Abstract: With the rapid urban development, sustainable, reliable and safe transportation solutions are required, to satisfy the demand and minimize potentially negative impacts. The field of autonomous vehicles (AVs) has got lot of attention among media, researchers, politicians and industry since its introduction. It can fundamentally change road transportation and improve quality of life. Although, on one hand the benefits of AVs can bring revolution in transportation field and can solve many engineering problems. On the other hand, there are some serious concerns about social and ethical problems such as so-called trolley problem. In certain situations, AVs forced to take some decision where harming someone cannot be avoided, i.e. whether it protect its passenger or to protect other road users. Therefore, it needs proper ethical guideline and crash-optimization algorithm to make decision in certain situations and so we believe applied software engineering ethical approach can be the solution to these problems. In this paper, we will explore the regulative instruments, standards, legal and moral responsibilities of AVs to act and make decision in unavoidable situations.

Keywords: Autonomous Driving, Autonomous Vehicle, Self- Driving Vehicle, Driverless car, Automation, Morality, Ethics, Dilemma, utilitarianism, risk management, Safety, Trolley problem.

1 Introduction

1.1 Background and Motivation

In recent decade, the road vehicle automation has seen rapid progress and development. The Autonomous Vehicles (AVs) has opened a new path for mobility and have explored social and environmental benefits. Big companies like Google, Tesla, GM, BMW, Mercedes, Ford, Toyota, and Audi are working on this project to make AVs more sustainable, reliable, and safe. AVs can be further classified into following categories as per SAE International (SAE International 2014, 2016). According to WHO reports, every year, car accidents led in the deaths of around 1.2 million individuals globally.According to the ENO Transportation Center research, about 93 percent of the 5.5 million accidents in the U.S. were caused by human error [6].Most of these deaths happened owing to intoxication (e.g. alcohol, illegal drugs), speeding (30 percent), distracted drivers (20 percent), and other variables such as weather conditions or absence of driving understanding.Most scientists think that AVs can considerably decrease the total amount of accidents and fatalities.Some reports say that AVs will reduce traffic crashes by 90%. [4]. Nevertheless, autonomously managed Google car has faced 11 minor acci dents in travelling of 1.7 million miles [6]. AVs have the potential to overcome traffic congestion issue, can reduce pollution and can eliminate human error in traffic accidents.

Monitoring of driving environment	Level of automation	Description
Human driver	0: Driver only	The human driver performs all aspects of the dynamic driving task
	1: Assisted automation	A driver assistance system performs either steering or acceleration/deceleration, while the human driver is expected to carry out the remaining aspects of the dynamic driving task
	2: Partial automation	One or more driver assistance systems perform both steering and acceleration/deceleration, while the human driver is expected to carry out all remaining aspects of the dynamic driving task
Automated driving system	3: Conditional automation	An automated driving system performs all aspects of the dynamic driving task (in conditions for which it was designed), but the human driver is expected to respond appropriately to a request to intervene
	4: High automation	An automated driving system performs all aspects of the dynamic driving task (in conditions for which it was designed), even if the human driver does not respond appropriately to a request to intervene
	5: Full automation	An automated driving system performs all aspects of the dynamic driving task under all roadway and environmental conditions

Fig. 1. *SAE International level of automation*

1.2 Research Problem

Although there are many good reasons to welcome AVs, they cannot prevent accidents completely and will therefore face an ethical dilemma in which harming someone is unavoidable [9]. Software is a crucial part of modern self-driving vehicles. Thus, its duty of software engineer to consider seriously all the aspects of ethical and social issues. Typically, the ethical discussion starts with the so-called “trolley problem” which is described in many articles and journals. The general scenario of the trolley problem is discussed below: Consider a scenario where an autonomous vehicle is driving at a high speed on a street. Suddenly a group of people comes

in front of the vehicle and blocks the whole street. But the car is too quick to stop before reaching the group. If the car does not respond instantly, the entire group will be killed.But here is a scenario where car can save them by entering the pedestrian way. Here, we have following alternative scenarios: (A) Replace the pedestrian with a concrete wall or a divider, which can result into the death of passenger; (B) Vary the personas of people in the group of pedestrians or the single pedestrian or the passengers. Varying in individuals allows to include an emotional view i.e. the pedestrian could be a kid, a family member, a very old man or woman or a thief etc [7].

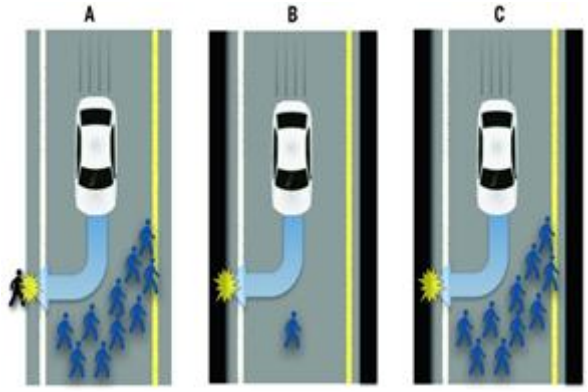

Fig. 2 *Three different traffic scenarios where harm cannot be unavoidable.*

The car must decide between (A) killing several pedestrians or one pedestrian, (B) killing one pedestrian or its own passengers, (C killing several pedestrians or its own passengers But these scenarios have limited number of possible answers, all of which are morally questionable and considered as wrong [11]. Thus, to solve these problems typical approach is to understand scenarios using various ethical approaches such as utilitarianism, other form of consequentialism or deontological ethics [1].

1.3 Research Aim

The research focuses on to explore more about complexities involved in implementing ethical approaches and legal responsibilities of AVs in certain unavoidable situations. To identify what issues, need to be addressed in order to develop satisfactory solution to the moral issues surrounding these approaches. To think about challenges involved with developing and implementing ethically good crashing algorithm.

2 Literature Review

There has been a lot of work done on the subject of "Autonomous Vehicles and ethical dilemmas" and to address my research problem We have done several literature reviews, some of them are stated below. **"Ethical and Social Aspects of Self-Driving Cars"**isone of them presented by Tobias Holstein, Gordana Dodig-Crnkovic & Patrizio Pelliccione. In this literature, they have explored various ethical and social aspects of autonomous vehicles. (Tobias, Gordana, & Patrizio, 2018) suggested that these social and ethical problems are described in the form of unsolvable decision-making problems and are highly misleading, rather than the main focus should be on integrated engineering ethical approaches for the implementation of the new technology [2]. They believe that, while an ethical dilemma has no proper solution, an engineering problem has ability to differentiate between good and bad. In support to this argument they have discussed trolley problem with different scenarios and presented how utilitarianism or any other ethical doctrine cannot solve this problem rather they can assure that whatever the solution is calculated that will be based on defined set of rules or other mechanism and thus it makes this engineering problem where they are actually defined and implemented. They believe that decision making process can be critically affected by its design and the parts used in this complex system. Such as, the quality of sensors e.g. why thesesensors are used in system—is it to reduce cost and what if we can get way better quality of information by using expensive ones.

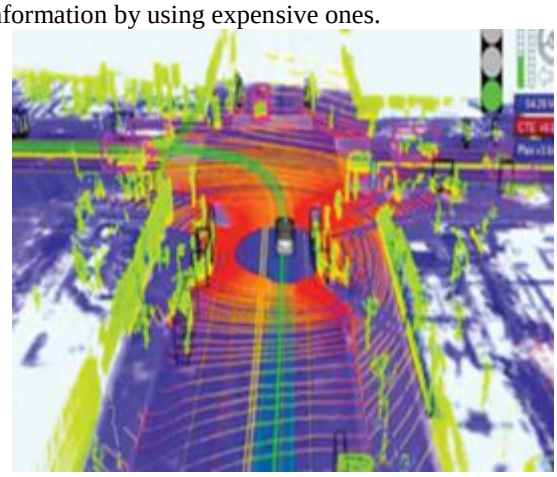

Fig. 3 Based on information gathered by a laser radar (LIDAR) [3]

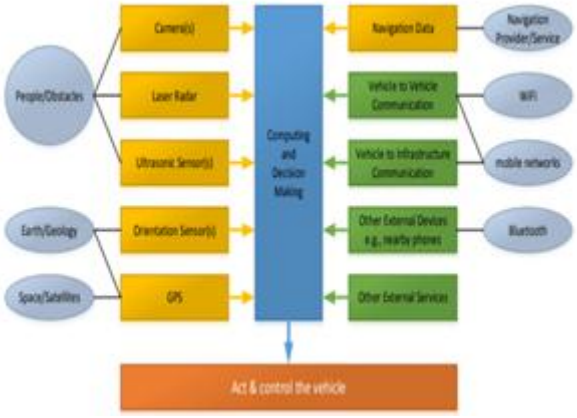

Fig.4.decision-making process in AVs.

In this article, they have discussed all the technical challenges and their ethical aspects in detailed manner. This includes:"Safety, Security, Privacy, Trust, Transparency, Reliability, Responsibility and Accountability, and Quality Assurance Process". Also, they have discussed about social challenges and their ethical aspects and thisincludes: Stakeholders – General Public Interest, Possible New Selling Points, Legislation, Standards and Guidelines. They have also provided recommendation for each of these challenges. Autonomous vehicles are considered as future of transportation systems. So, rather than focusing on unsolvable idealized problems like trolley problem, it would be if we focus on practical technological solution and their challenges.Authors have suggested that AVs, should implement ethical rules that are defined and agreed by stakeholders and should not be an autonomous black box with unspecified performance and these challenges should be addressed with software- hard-

ware engineering solutions. Therefore, they strongly believe that ethical aspects should be considered in every phase of software development, architectural and design decision process, i.e. from requirements to testing, maintenance and evolution. **"Vehicle Automation and Duty to Act by Noah J. Goodall" [4].** In this articleauthor presented an importance of self-driving vehicles and moral duty to act with different perspective i.e. Legal Perspective and Moral Perspective. The author stated that in a situation where crash is unavoidable, crash severity models could be used to determine possible safest alternative solution, with one goal to minimize the cumulative injury i.e. by utilitarian approach. But this may lead to an ambiguous solution. Further, in moral perspective author has explained two ethical approaches: Normative Ethics and Descriptive Ethics. Normative Ethics includes utilitarian and deontological ethical theories. It describes what we as an individual and society supposed to do not what we do or prefer to do in certain situation. The author stated that on one hand utilitarianism considers an action good if it produces the maximum net cumulative benefit. On the other hand, Deontology is quite opposite of consequentialism which is merely based on set of defined rules, duties or rights. So, an action will be taken solely on the solution rather than any underlying principle. On the contrary, descriptive ethics describes the study of individuals' or groups' beliefs about morality. In this approach, AVs factorize all the identity and gives value accordingly and later use these values to make decisions. In the end [10] author suggested that, there are two ways, AVs can act morally. First, where AVs can compare its current scenario and perform recommended actions. In this approach manifold scenarios are defined explicitly. And Secondly, an alternative approach is to use machine learning techniques, where AVs are continuously trained with a set of scenarios and responses. Both techniques have their advantages and disadvantages. Lastly, author believes that its society which has to decide how much amount of avoidable risk should AVs take to protect pedestrian or other vehicle, proper guidelines and regulations should be defined. **"Caught Up in Ethical Dilemmas: An Adapted Consequentialist Perspective on Self-Driving Vehicles" done by Vanessa SCHÄFFNER" [9].**The Author provided an in-depth summary of analysis that describes the problems of a potential utilitarian solution and consequentialist perspective as a feasible solution thatbuilds upon the concept of negative utilitarianism and prioritarianism. The Autor has scrutinized suitability of utilitarian approach for ethical issues in autonomous vehicles. The Author has also clearly identified the pros and cons of utilitarianism and on that basis, tried to introduce an adapted consequentialist approach.

3 Approach

This research focused on decision-making process and security of autonomous vehicle. The whole decision-making process in autonomous vehicle can be outlined as follow: Decision-making process starts with obstacle detection in surrounding environment i.e. number of lanes, group of people, animals, surrounding buildings and the current location of vehicle, speed and direction of vehicle. It is very critical to have all this information about surrounding obstacles when making decision. Various sensors and other mechanism used to gather all this information precisely. Now, based on parameters set in computer program and algorithm it will calculate all the outcomes and choose best one from them.

3.1 Consequentialism

The utilitarian approach is one of the most subtle and consequently extremely relevant theories that are worthy of future evaluation. It has been proven that it is easy to compute various practical scenarios mathematically using this consequentialist approach than other ethical approaches. The utilitarian approach states that an action should be assessed ethically based solely on its implications. The utilitarian approach tends to maximize the utility and minimize the losses, even if it means to kill or harm its own passenger. In order to evaluate different outcomes of certain action practically, they should be assigned to a numerical utility value. It maps the practical scenarios which allows it to evaluate scenarios mathematically. E.g. there is a scenario where physical damage of AV and saving lives of pedestrian involved then utilitarian approach would calculate and choose saving of life over physical damage of AV. So, in case of autonomous vehicle it will choose an option which maximizes theresulting level of utility. Since the outcome is calculated in terms of caused harm, the maximum utility can be achieved by choosing least harm causing option in an inevitable situation. In a simple word, these means that in a situation where crash is unavoidable, it will take decision where the amount of harm or loss of lives are lowest, or it would try to save as many lives as possible. [9]. But, on the downside this approach has many shortcomings, when applying in a practical situation where a crash is inevitable. Basically, in order to calculate all possible outcomes, the scenarios should be quantified in some way. However, it would be ethically challenging or almost not possible to define metrics or a numerical value for functions to calculate outcomes of accident, and that is when it involves decision about to save and loss of human life. This raise many unanswerable questions, that is how a person's life's value can be objectively determine. Additionally, it is not easy to calculate the values of a particular single good outcome, especially when the sensors are used in autonomous vehicles creates a huge amount of information. This raise concerns in terms of which information can be used for making certain decision and what if the sensor is at fault and thus supplied information can be faulty as well. Many articles argue that utilitarian decisions disrespect the right of life, which cannot be traded with any other right, also it inspires the discrimination of target objects. As, utilitarian takes decision based on to minimize harm, an AV with utilitarian always choose an option with least negative outcome. E.g. an accident with smaller car can cause a less severe damage compared to bigger car.Thus, smaller cars would become safer target option in utilitarian approach. Furthermore, in some other scenarios, it raises the questions, such as whether to hit the motorcyclist with helmet or with one without helmet. Here, in this case it will chooseone with helmet because it will cause less harm compared to other one. But, isn't that unfair to hit someone who wears helmet for one's own safety? Some articles also criticize that computed outcome from utilitarian approach might actively targeting persons who can be sacrifice for the greater good, which ultimately violates the basic human principle of equality [9]. Figure 5 and 6 indicates a view of current scene from human eye and the view from the laser scanner respectively. As we can see, from figures the laser scanner provides significantly different data than the human eyes see. Thus, any moral decision or outcome chosen by an autonomous vehicle will be based on improper understanding of various objects. This raises many questions as the information about objects are

uncertain, then how it is possible to accurately evaluate various outcomes of accident.

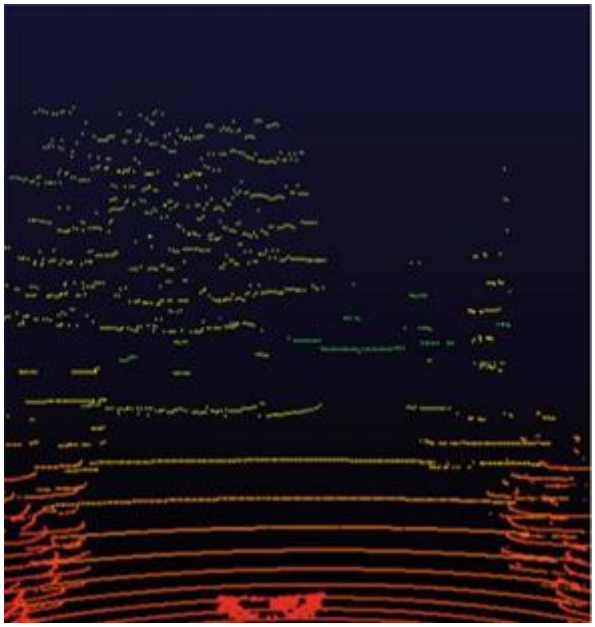

Fig. 5. *A Human eye perspective*

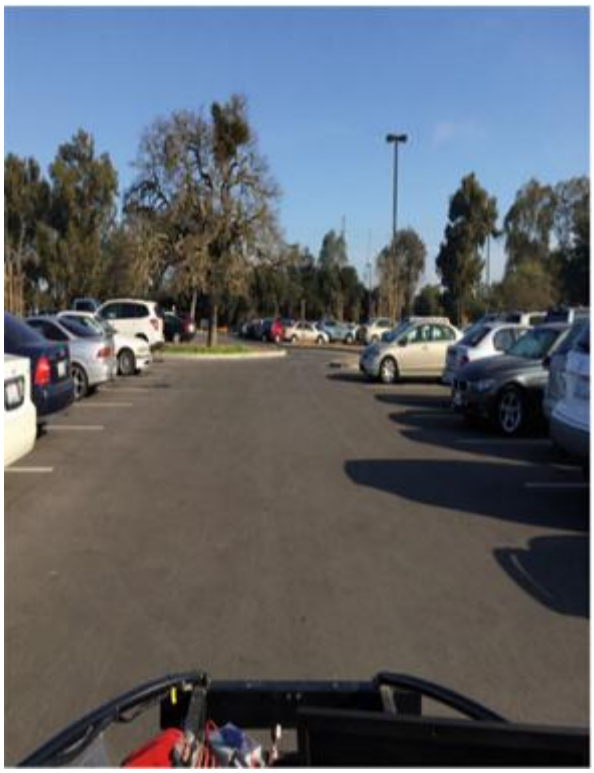

Fig. 6. *Laser Scanner perspective*

Applying these ethical frameworks and moral principles into algorithms does not solve the problem but it can give assurance that whatever the decision will be made it will be based on a given set of rules and mechanism.

3.2 Deontology Approach

This decision-making process may vary depending on the ethical framework. Deontology framework presents ethics as a fixed defined set of inflexible rules, duties or rights and following them makes action morally correct, otherwise makes it immoral act. It is an opposite of consequentialism, where a moral action is considered upon calculated outcomes rather than defined principle. According to Kant, in deontology, "the sole feature that gives moral worth is not the outcomes that is achieved by an action, but the motive that is behind the action." [8]. Its human's duty to act what is morally right even if the result is bad. It's better to program an autonomous vehicle to try to attempt avoid collision rather than trying to calculate harm or injury to human beings. These means that by limiting the responsibility to collision avoidance in autonomous vehicle, the car need not to be programmed to sacrifice itself to protect the human being in which it would otherwise not have been engaged [5]. In general, the cost function, calculates the impact of various actions on manifold surrounding objects and optimal controller would prioritize these objectives with higher cost than other objectives. Such objectives or set of rules for an autonomous vehicle can be defined by imposing constraints on the optimization problem. So basically, we can program the autonomous vehicle in a way which avoid collision if there are already set of rules/ prioritize or control inputs existed to prevent or handle incidents [5]. The following set of rules are my interpretation from Asimov's Three Laws of Robotics:

1. An automated vehicle may not harm or minimize the harm to pedestrians or cyclist.
2. An AVshould try to avoid any accidental situation with another vehicle, except where orders would clash with the first law.
3. An AV should not clash with any other asset or property in the surrounding, except where orders would clash with first 2 rules.
4. An automated vehicle should obey the traffic law of its current location, if orders does not clash with first 3 law.

These set of rules can be implemented straightforward in AVs and prioritized according to this hierarchy with minor error on constrain violation. To implement such ethical rules, they just need to categorize the various objects and does not need for finer calculated value about any injury or harm.However, these rules do not completely define ethical framework, they can simply address many ethical issues start with first law. This law acknowledges the importance of a persons' lifein self-deriving cars and duty to protect it. The reason to giving them fist priority would be an AV is a product and it should be AV's first responsibility to protect fundamental value of human life, otherwise no one would prefer to buy such product. To understand the cost function and constraints in decision-making process of autonomous vehicle consider following example. Figure below shows the simple implementation of ethical rules. As shown in figure below, the vehicle's primary objective is to travel safely down the road without any interruption at a desired speed.

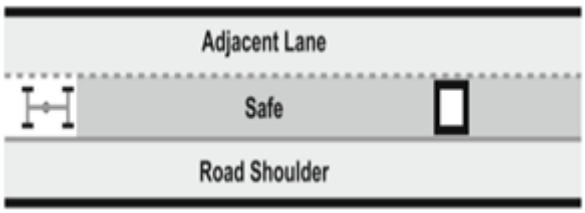

Fig.7. *Basic driving scenario, where car is traveling on a straight-two lane road with the road shoulder on right side and approaches an obstacle blocking the lane [5].*

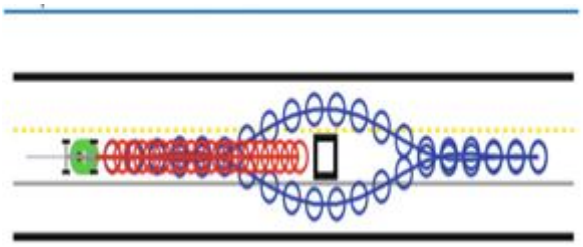

Fig.8. *Three possible choices to avoid an accident with obstacle – use of brake before colliding with obstacle in red, and navigate through either side of the obstacle in blue [5].*

Upon finding an obstacle in its way it starts searching alternatives and as shown in figure 8 it has three options i.e. either it can use brake to stop before colliding with obstacle or it can navigate to either side of the obstacle. AV will try to evaluate lowest cost from available options based on their weights and constraints using optimization-based controller. So, if the lane boundaries have assigned high cost values to cross, then AV will use brake and stop before colliding with obstacle as this action would produce lowest cost. But, in case if lane boundaries are encoded as soft constraints then AV can have other possibilities. Now, AV need to choose among available options i.e. into the lane on which traffic is coming or to travel onto the road shoulder, whichever has lowest cost and that will also depend traffic laws of current location of AV. Such as, if the section of the road does not simply allow to navigate through oncoming traffic lane, then it might have high cost constraints and, in such cases, use of shoulder would be more appropriate in terms of cost and constraints. Thus, we can handle various collision situations, hazardous situations and goals such as traffic flow and continuous mobility by simply defining different priorities and behaviors with different cost into a program.

4 Outcome

Consequentialism proves to be not trustworthy and coflicting with the ethical goals of autonomous vehicles. Furthermore, it's not just unreliable but also unfeasible and impractical for a programmer to program and implement it. For autonomous vehicles to be successful in the market, they need to guarantee the safety of those using it. On the other hand, through careful design of control system and by defining and implementing set of rules, duties or priorities an autonomous vehicle can prevent any incident that lies within the limitations imposed by the laws of physics. Apart from this, in cases where crash is unavoidable, society need to accept the suboptimal outcomes even if it is not in favor of human being also societies need to accept that AVs possess a clear respect for human life above any other priorities.

5 Conclusion

According to various reports, fully autonomous vehicles are expected to enter our road by 2030. By looking at their widespread acceptance, establishing an ethical system that provides guidance to handle dilemma situations is a pressing issue. A consequentialist approach in AVs may sound more attractive and appealing, as everyone love to save human lives as many as possible, but it is impractical, inconsistent and unfeasible.We found a hope in deontology approach where ethical framework can be programmed using different priorities and behavioral values of autonomous vehicle. If we want autonomous vehicle to respond to an action ethically, then we may need to provide more controlto them. The transparency in defining and programming set of priorities will ultimately increase the trust of society, and a faith that whatever the decision made in an unavoidable situation are sensible, ethical and acceptable to society.

References

[1]. Barbara, M. & Andrew, F., 2014. Ethics: Theory and contemporary issues. s.l.:Nelson Education.

[2]. Bonnefon, J.-F., Shariff, A. & Rahwan, I., 2016. The social dilemma of autonomous vehicles. Science, 352(6293), pp. 1573-1576.

[3]. Earth Imaging Journal(EIJ), 2012. LiDAR Boosts Brain Power for Self-Driving Cars. [Online] Available at: https://eijournal.com/resources/lidar-solutions-showcase/lidar-boosts-brain-power-for-self-driving-cars

[4]. Goodall, N. J., 2014. Vehicle Automation and the Duty to Act. [Online] Available at: https://www.academia.edu/22479919/Vehicle_Automation_and_the_Duty_to_Act [Accessed 03 2019].

[5]. J., C. G. & Sarah, M. T., 2016. Implementable Ethics for Autonomous Vehicles. In: Autonomous Driving. Berlin(Heidelberg): Springer, pp. 88-102.

[6]. Jan, G. & Mu¨ller, J. F., 2016. Autonomous Cars: In Favor of a Mandatory Ethics Setting. Springer Science+Business Media Dordrecht, Issue 23, pp. 681-700.

[7]. Lim, H. S. M. & Taeihagh, A., 2018. Autonomous Vehicles for Smart and SustainableCities: An In-Depth Exploration of Privacy andCybersecurity Implications. MDPI, 11(5).

[8]. Matt, M., n.d. Immanuel Kant: Metaphysics. [Online] Available at: https://www.iep.utm.edu/kantmeta/[Accessed 05 2019].

[9]. SCHÄFFNER, V., n.d. Caught Up in Ethical Dilemmas: An Adapted Consequentialist Perspective on Self-Driving Vehicles. [Online] [Accessed 03 2019].

[10]. Taeihagh, A. & Lim, H. S. M., 2018. Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. Taylor and Francis Group, pp. 103-128.

[11]. Tobias, H., Gordana, D.-C. & Patrizio, P., 2018. Ethical and Social Aspects of Self-Driving Cars, s.l.: Arxiv.