
Proceedings Paper:
Millan-Blanquel, L., Veres, S.M. orcid.org/0000-0003-0325-0710 and Purshouse, R.C. orcid.org/0000-0001-5880-1925 (2020) Ethical considerations for a decision making system for autonomous vehicles during an inevitable collision. In: 2020 28th Mediterranean Conference on Control and Automation (MED). 2020 28th Mediterranean Conference on Control and Automation (MED), 15-18 Sep 2020, Saint-Raphaël, France. IEEE . ISBN 9781728157436

https://doi.org/10.1109/med48518.2020.9183263

Ethical Considerations for a Decision Making System for Autonomous Vehicles During an Inevitable Collision

1st Lalis Millan Blanquel*, 2nd Sandor M. Veres and 3rd Robin C. Purshouse
Department of Automatic Control and Systems Engineering
The University of Sheffield
Sheffield, United Kingdom
lmillanblanquel1@sheffield.ac.uk, s.veres@sheffield.ac.uk, r.purshouse@sheffield.ac.uk

***Abstract*—The automotive industry is heading towards the introduction of fully autonomous vehicles. However before these type of vehicles are commercially available at mass scale, some issues need to be solved. A major issue is the ethics involved in the decision-making during an accident; this paper presents a novel analysis of how to solve this issue. The proposal is a pre-programmed system with different ethical settings based on six formal ethical theories. For the implementation, eight ethical concerns are defined and ordered according to each theory. These concerns are defined in terms of harm to self and harm to others. The ethical concerns are used as a guideline to define the level of importance of each person or object in an accident scenario. With the proposed system, the vehicle will be partially tailored to the preferences of different users while still being bounded by legal requirements to avoid any misuse.**

***Index Terms*—autonomous vehicles, ethics, decision-making, safety systems**

I. Introduction

In recent years, there has been an increase in the development of systems for Autonomous Vehicles (AVs). However, there are unresolved issues related to the ethics involved in the use of these kind of vehicles. In this paper we will focus on the implementation of ethical decisions in case of an emergency.

While driving, humans make decisions based on their knowledge and experience. During an emergency, stress, time constraints, and emotions are factors that influence the decision-making process. However, when a machine makes a decision this is based on calculations and logical processes. Machines do not feel stress and can analyse a situation faster than a human can. Autonomous vehicles would make these critical decisions without the intervention of a human; however when faced with a situation that involves human lives, society needs more that a set of probabilities to justify and understand why certain decisions were taken [1].

The conditions to decide who would suffer has also been debated. To decide between hitting one person or another using criteria such as age, race, gender, profession or disability, can be considered a discrimination issue and is prohibited not only by governments [2] but also by engineering associations like the IEEE [3].

*Sponsored by Mexico's CONACyT, scholarship number: 440758

The use of a utilitarian perspective has been discussed in the literature. The question with these approaches is to define what is more important: to preserve the freedom and the individual rights of each person or to seek the greater good for the community. In [4] the author argues that the user of the vehicle should be able to decide the ethical behavior of the vehicle to keep her free decision on how she is being transported. However, some researchers [5] defend a mandatory ethical setting over a personal one, arguing that most people would select an egoistic behavior that will make the situation worse for everyone. According to this view, a mandatory ethical setting would be beneficial for everyone. Nevertheless such guidelines for designing algorithms should be carefully reviewed to ensure that no discrimination is present or that there are no parameters that can bias the decision of the AV in detriment to certain groups of people [6].

Although, the *trolley problem* has been popular in the debate over ethical questions about AVs, some authors argue that this abstraction is inappropriate [7]. The first difference between a trolley problem scenario and AVs is that, in the case of the AVs, the decision is made a long time before the accident, whereas in a trolley problem the person is forced to make the decision in the moment. At the same time, in the case of AVs the decision needs to be made with a thorough analysis of the situation and should be agreed upon by various stakeholders (e.g. car manufacturers, system developers, society, lawmakers). A second difference is that in the trolley problem is assumed that the outcome is 100% certain, whereas AVs only work with estimates and should consider risk estimation and decision-making under uncertainty to define the best actions.

While it is true that the trolley problem may be a simplistic case, its use in research is useful to develop safety features, as there are suggestions that autonomous vehicles would not be 100% safe, especially in the first years of introduction while these still share the roads with non-autonomous vehicles. Hence the importance on developing systems that can help to resolve these ethical issues [8].

This paper is organized as follows. In Section II we present the definition of the ethical theories that we are using and some legal considerations. Section III describes the proposed

system and the definition of the ethical concerns to implement. Section III-C contains a discussion about some of the problems that could arise with a system like the one presented here and Section IV presents our conclusions and future work.

II. Literature Review

A. Ethical Theories

In the philosophy field, there exists a number of interpretations for ethics. In this section we present the main five that we have selected for our system. Ethics are moral philosophies that encompass rules and behavior standards that can serve as a guide for decision-making [9]. Different ethical perspectives are encountered through the literature; however not all of them may be applicable to AVs. Some of the most relevant are explained here but a more comprehensive explanation of these theories can be found in [10]–[12].

Utilitarianism looks for the greater good. It defines the morality of an act in terms of its consequences; if a decision produces good for the most people, then it is morally correct [11].

Deontology is a moral theory focused on duties and rights, with an emphasis on the action itself and not the result of said action. This contrasts to other theories like utilitarianism where the focus is on the outcome of the action [13]. Different authors have given their interpretation of deontology; one of the most known is by Kant. In this paper we present two interpretations of this approach.

Distributive justice is a theory developed by John Rawls where he emphasizes a fair and equal allocation of resources [12], [14]. The key to this theory is to place the self in what he calls 'the original position' this means taking a perspective where the relative merits of each position or role in the scenario are understood, but from these positions, the one occupied by the self is unknown. Since the self is ignorant of its position, Rawls argues that the situation can be analyzed without biases with the objective to optimize the worst outcome, as that could be the self's outcome itself.

Altruism is to care for the well-being of the others, even if that may lead to self-sacrifice to protect other people [11].

Egoism can be divided into two types; first is psychological egoism, where decisions are made based on a natural impulse of self-satisfaction and instant gratification. Second is ethical egoism that holds that the self should seek its happiness using its rational judgment to make a decision and abiding to principles like integrity and honesty. Opposed to altruism, some philosophers argue that caring for one's self is also morally acceptable [10], [13], [14]. For our system, we will implement an ethical egoism approach.

B. Ethical perspectives in the context of AV

Some of the mentioned ethical theories have been discussed in the context of AV in the literature. Utilitarianism is the most popular suggestion for implementation in AVs [4], [5]. One of the problems with this scheme is that it does not solve the question of what to do when the number of persons harmed is equal for all possible outcomes. Take, for example, a situation where the problem is to save the passenger of the vehicle or to save a pedestrian. In this situation, utilitarianism does not provide a clear answer.

Altruism is another theory that has been implicitly discussed in relation to AVs [15]; however not all potential buyers may agree to ride in a vehicle that will not seek to protect them. Egoism may also be suitable to be applied to AVs, but it requires legislation as it may not be acceptable to cause harm to many people just to save one person. In general, it is not clear that any single ethical perspective can universally satisfy all possible sets of preferences or scenarios.

C. Legal Considerations

One of the most attractive benefits of the introduction of AVs to our roads is that they may help to reduce the number of traffic accidents. Although thousands of lives may be saved, it is important to recognize that fatalities cannot be completely avoided. The intervention of AVs would create a new set of victims that in other circumstances may have not been damaged [16]. This situation creates a trade-off between the lives that can be saved and the lives that would be affected by AVs. Even if it can be said that the greater good should prevail, this is a trade-off that may not be sustainable from a legal perspective. However some researchers argue that this not may be the correct way to interpret the situation, as the introduction of AVs would reduce the chance of an accident and that is to the benefit of everyone, regardless of whether an accident occurs or not [17].

With autonomous cars already being tested in different cities, the first accidents have already happened [18], [19]. However there is still no legislation in place to determine who is responsible for the damages caused by AVs. It is important that these regulations are in place before these vehicles are available to the public [20]. Mainly there are three legal branches that cover AVs [21]:

- *Administrative law* oversees issues such as the traffic rules and technical norms for the operation of AVs on public roads;
- *Civil law* covers civil liability for injury or damage and product liability related to damages caused by a defective product;
- *Criminal law* manages responsibility when a crime has been committed.

One of the questions that researchers have been exploring relates to who would be responsible for the actions of an autonomous vehicle: owners, manufactures, programmers or the machine itself. Some authors suggest that the owner of the vehicle should be considered morally responsible in case of an accident [17]. This perspective could be considered as a shared responsibility within all the owners of AVs, since they are introducing a risk by using this type of vehicles. This can be applied as some sort of special insurance, tax or collective fund to compensate those affected [22], [23].

There are also proposals to protect the automakers from too harsh legislation, that could steer them away from the development of AVs. To limit the criminal responsibility of

the manufacturer, a margin of tolerance for errors that may occur from the programming of the vehicles would be set [22]. Moreover, sacrificing the passenger to protect somebody else can be acceptable assuming that the passenger has signed a contract stating in clear terms that she understands and accepts this decision [24].

Under current legislation, some of the principles than can be applied to the programming of AVs are as follows. The doctrine of necessity [24] is a legal approach in criminal law to regulate those cases where damage has been intentionally inflicted to a third party when avoiding all evils is impossible. Within this doctrine is found the theory of justification. This dictates that under extraordinary circumstances an otherwise prohibited action can be accepted. Programming a vehicle to harm under certain circumstances (to target someone that anyway was going to be harmed, in order to protect more people) can be acceptable as part of the theory of justification. However, to intentionally harm a third party unaware of the situation (swerve and harm a passer-by in the sidewalk) would not be acceptable. Nevertheless human lives should always be protected over property [22]. Some authors hold that ethical egoism should not be programmed in the AVs as it cannot be viewed as legally correct [25]. However utilitarianism also poses a challenge to the legal system since, in a democratic system, everyone has rights that cannot be ignored to justify social benefit. Nevertheless, even if some guidelines can be drawn from current legislation, scholars suggest that governing agencies should review their legislation incorporating feedback from the public, the industry and relevant government agencies [21], [26]–[28]. Including the views of all the affected parties would help to establish clear rules to help boost the development of intelligent vehicles.

D. Existing Proposals

Although there has been a lot of research in accident management, few of these studies address the moral questions of decision making, aiming mainly for collision mitigation or avoidance solutions. However, as already pointed out, it is important to know the reasons behind these decisions so that they are defensible in a legal and ethical framework.

An ethical framework alternative to utilitarianism is presented by Leben [29], based on Rawls' distributive justice theory. The idea behind this algorithm it is to achieve a Pareto-optimal solution; this means a solution where no participant's situation can be improved without making another participant's situation worse. This algorithm is based on the idea that the participants do not know what role they are playing (e.g. passenger, pedestrian) hence it is in their best interest to maximize the utility of the worst outcome for all players. To make a decision, the algorithm evaluates the probability of survival of each participant on each possible maneuver and applies the maximin strategy with the objective of obtaining the best of the worst outcome. The payoffs of each maneuver are grouped into a set; from this set the lowest payoffs of each participant in each action are selected and a new set is created. On this new set of minimums, the maneuvers that have the highest payoffs are selected. This process is repeated until just one maneuver remains -this maneuver would be the decision output. If two or more actions tie with the same pay offs, the decision is randomly selected based on these actions.

A three level system to be implemented as the technology becomes available is presented by Goodall [8]. In the first phase, a rule-based system would be encoded, these rules should be agreed on by ethicists, automakers and lawmakers. The behaviors selected would be utilitarian -those that minimize damage. A problem with this implementation is that, in case of a scenario not covered by the rules, the car would just brake and evade, which in a collision situation does not provide a feasible solution. Nevertheless, in phase two, machine learning techniques, like neural networks, would be implemented to increase the vehicle's understanding of ethical decision. A shortcoming with this approach is that, if the set of scenarios used to train the neural network is not diverse and appropriate, the morals learned by the algorithm can potentially show extremist behaviors or discriminatory biases. Phase three of Goodall's approach consists of receiving feedback from the automated systems to understand the logic used to make a decision. However the author argues that implementation of this step may be slow as more research is needed on how to extract such information from neural networks.

These previously presented implementations are to be pre-programed in the AV. By contrast Contissa et al [30], presents a customisable knob with three broad settings: (1) an altruistic mode which gives preference to other people lives; (2) an impartial mode, where passengers and other persons have equal importance; and (3) an egoistic mode, where the passengers lives have preference. The author also argues for a continuous mode, in which the knob would allow the user to select the weight of the passenger life relative to other persons. The system would take into account the probability for the passenger and third parties to suffer harm resulting from the AVs decision, to select the option with the smallest disutility. One of the problems of this implementation is that selecting too egoistical a value would cause the vehicle to protect the passenger at all costs even if the probability of harm is too low, unnecessarily exposing the life of the pedestrians. The authors argue that the limits over the value that the passenger gives to her life should be regulated by law to avoid this situation.

Dennis et al [31] propose a framework where the ethical decisions are drawn from the ethical codes and regulations governing the profession related to the function of the machine. They assume that each professional domain has already developed ethical principles and substantive rules to evaluate how ethical an action is when there is no ethical option available -the system should then take the least unethical decision. The authors provide an example of an autonomous aircraft, considering that the machine should act as a pilot would do. In their proposal they establish a set of ethical concerns and define which of them are more ethical to violate. The system presents two operational modes. The first is controlled by the pre-programmed plans where the programmer assumes

responsibility and the second mode where the ethical reasoning is needed to operate when no plan is available or all plans have already been implemented but failed. A rational agent determines which of the new plans, supplied by an external planning mechanism, are the most ethical to follow.

A benefit of this system is that it can evaluate how ethical a decision is but is not limited to only perform ethical actions, an unethical action can be performed provided that there are no more ethical actions available. Additionally, the framework is designed to be verifiable, meaning that if a decision made is considered unethical it can be proven that the system believed that it was the minimally unethical action from those available. A limitation in a system like this is that driving a regular vehicle is different to operating an aircraft. While there are traffic codes that must be followed, these vary from country to country and do not always resolve ethical issues.

Dennis et al's concept - defining a set of ethical concerns and ordering them in order of importance according to an ethical policy, used in [31] serves as the basis for our proposal presented in Section III-B.

The novelty of our work is that, rather than assuming a fixed ethical policy, we cater for a multiplicity of theories (six in the present work) and allow a theory to be selected by the user. This approach results in a different maneuver depending on the theory that is selected by the user, accommodating for the ethical views and preferences of each individual, while at the same time upholding pre-established rules and laws as implied by [30]. In the proposal by Dennis et al the ethical concerns are defined by a specific set of rules to follow and a decision is made based on how unethical the rule is to break and does not differentiate the gravity with which each principle is violated, simply establishing that violating two separate principles is worse than violating just one. The system proposed in the present paper does not contemplate rules. The ethical concerns are defined by considering all the possible participants in an accident scenario. These concerns are then prioritized according to what is expected in each ethical approach.

III. Pre-programmed AVs with Customizable Ethical Preferences

In this section we propose a novel approach for a decision making system for autonomous vehicles, followed by the identified ethical concerns, their definition and the analysis of these according to each ethical theory. Finally a discussion about some of the issues that could arise with this implementation are touched upon.

A. Ethical Decision Making System

We are proposing a system for fully autonomous vehicles, capable of taking decisions based on the user's desires. The user will be able to select an ethical view that aligns with their preferences through a user interface in the vehicle's screen. This setting will be stored in the vehicle's memory and will be retrieved if needed. Through the normal operation of the vehicle, if a dangerous situation is encountered, the vehicle will calculate the possible maneuvers that the vehicle can perform. If a safe maneuver is available this will be implemented. In case that there is no possibility to avoid a collision then the vehicle will evaluate the possible maneuvers against the user preferences and the legal requirements previously stored in the vehicle's memory and will apply the one that complies with the greatest number of the constraints. To the best of our knowledge, a system like the one presented here has not been studied before.

The system's proposed aim is to give the user the opportunity to have an input on the behavior of the autonomous vehicle. While the vehicle computer will be the one making the final decisions, it will consider the wishes of the user. With this, we want to give a different perspective over the issue of who should make the decisions in an autonomous vehicle. In this paper we present an analysis of the ethical theories that will be implemented and the different ethical concerns that will form the base guidelines of the system to establish the course of action to follow.

Some studies mention that humans feel more comfortable when they aid the machine in taking decisions. People's ethical views can vary from person to person and be influenced by culture, so a predefined setting that may be accepted in a country may not have the same acceptance in another [32]–[34]. Hence with this system the vehicle can be pre-programmed by the user to accommodate each different preference.

B. Ethical Concerns

Similar to the work of [31], we have defined eight ethical concerns for the vehicle to follow. These are:

- Harm the least possible number of people (c1)
- Do not harm passengers (c2)
- Do not harm people outside the vehicle (c3)
- Inflict the least damage possible to people (c4)
- Do not harm vehicles with passengers (c5)
- Do not harm children or incapacitated people inside the vehicle (c6)
- Do not harm animals (c7)
- Do not harm objects (c8)

As discussed in Section II-D, we have selected six ethical perspectives to be applied to the system. These are utilitarianism, distributive justice, deontology, altruism and egoism.

The ethical concerns have been defined in terms of harm to self and others; depending on the selected ethical approach, these will have more or less importance to the system. As passengers, we define any person that is traveling in the vehicle. The people outside the vehicle refers to either pedestrians or bystanders. A concern for other moving vehicles has been added because we assume that there are people inside these vehicles, making them distinct from our definition of objects, where we refer to any element like empty cars, walls and poles. A concern related to children traveling inside the vehicle has been added. The reasoning behind this choice will be explained in more detail when we explain the deontological approach. The distinction between harming the least possible number of people and inflicting the least damage possible to

TABLE I
ETHICAL CONCERNS RANKED ACCORDING TO EACH ETHICAL THEORY FROM HIGHEST PRIORITY AT THE TOP TO LOWEST PRIORITY AT THE BOTTOM

Ethical concerns (for reference, unranked)	Utilitarianism	Distributive justice	Deontological	Kantian	Altruism	Ethical Egoism
Harm the least possible number of people (c1)	c1	c4	c3	c3	c3	c2
Do not harm passengers (c2)	c4	c1	c6	c1	c2	c3
Do not harm people outside the vehicle (c3)	c2	c2	c1	c2	c1	c1
Inflict the least damage possible to people (c4)	c3	c3	c4	c4	c4	c4
Do not harm vehicles with passengers (c5)	c5	c5	c2	c5	c5	c5
Do not harm children or incapacitated people inside the vehicle (c6)	c6	c6	c5	c6	c6	c6
Do not harm animals (c7)	c7	c7	c7	c7	c7	c7
Do not collide with objects (c8)	c8	c8	c8	c8	c8	c8

any individual is that in the first we are referring to overall numbers where fewer people damaged is better, and in the second it is permissible to harm more people if that means that the most damage to any individual would be less. For example, it would be better to harm two people but keep them alive, rather than sacrifice one to avoid harming the other. A distinction between animals and objects is included since animals are sentient beings that are important for humans, based on a recommendation from [2].

For the **utilitarian** approach, we are looking to maximize overall good, hence the most important thing is to try to harm the least possible number of people (c1) irrespective of whether they are inside or outside the vehicle. The next most important thing is to try to inflict the least possible damage to any individual (c4). Do not harm passengers (c2) and people outside the vehicle (c3) concerns are equally important as, in an utilitarian approach, their location is irrelevant.

In the **distributive justice** principle the most important thing is to inflict the least damage possible to people (c4) followed by to harm the least possible number of people (c1). In this approach it is also irrelevant if the persons involved are inside or outside the vehicle hence concerns two (c2) and three (c3) are equally important.

As mentioned before, the **deontological** approach is the only one that differentiates between adults and children traveling in the vehicle. This is because healthy adults can acknowledge the risk involved in traveling in the vehicle whereas a child or an incapacitated person does not take this decision consciously. For this approach, do not harm people outside the vehicle (c3) is the most important concern and in case that there are children or incapacitated people in the vehicle, is equally important to not harm them (c6). The next most important thing is to harm the least possible number of people (c1) followed by inflict the least possible damage to people (c4). Finally, do not harm passengers (c2) has priority over colliding with other vehicles with passengers (c5), as we consider that passengers in those vehicles or the passenger's carers (in case they are children or incapacitated people) have also agreed to the implicit danger of traveling in a vehicle. We made this assumption because knowing the characteristics of the passengers in the other vehicles would depend on both vehicles being able to communicate to share this information. From this, a discussion about privacy would arise, which is out of the scope of this paper.

In the **Kantian** view the most important thing is to protect people outside the vehicle (c3), as we give them value as a person and avoid using them as means to our objectives. Next the system will aim to harm the least possible number of people (c1) followed by not harming the passengers (c2), and lastly to inflict the least damage possible to people (c4).

Altruism and **egoism** are similar in most of the ordering of concerns with the exception of the two first. For altruism, the most important thing is to protect the people outside the vehicle (c3) while for egoism the most important thing is to protect the passengers (c2). One important thing to note here is that the egoism condition can only be applied when the number of people outside the vehicle is the same as the passengers, as it is not possible to value the life of one person more than the lives of two or more.

The order that we have given to these concerns, based on the different ethical approaches, is presented in Table I. In first instance, concerns seven (c7) and eight (c8) always occupy the last two places, as animal and objects are always less important than people. Concern six, do not harm children or incapacitated people inside the vehicle (c6) is always above concern seven (c7), except in the deontological approach. Concern five, do not harm vehicles with passengers (c5), is above number six, because with the exception of the deontological approach, for any other case, age or any other characteristic of the people involved is irrelevant.

C. Discussion

The system proposed in this paper aims to allow the user to select an ethical behavior that aligns with their preferences. A potential disadvantage of a system like this is that it could be misused. However, assuming certain conditions, this can be prevented. Examples of these conditions are: (1) the user will not have control over the driving at any time, meaning that any accident would be due to a vehicle malfunction or other external factor and not because the user acted in bad faith; (2) the system should be parametrized according to the law and this can vary for different countries. For example, trading off the lives of two or more people to save one is not permitted.

An argument against allowing the user to select the ethical behavior of the system is that it could defeat the purpose of an autonomous system. However, while it is certain that the user acquires a duty and accepts the risk of using an autonomous

vehicle, it is also true that external factors completely out of their hands can cause an accident and they still should have the right to decide how they want their vehicle to behave. This does not mean that they should not be held responsible for their decisions but rather to change the focus from a potential criminal liability to a civil one, as discussed in Section II-C.

In legal terms, an issue that arises with the pre-programming of actions that may lead to people being harmed is if that could qualify as premeditation. Under current legal standards planning an action that leads to someone being hurt is a crime. However, in the case of the proposed system, neither the programmer or the user knows the circumstances nor are expecting or planning for an accident to happen and since they do not drive the vehicle they can not purposefully cause an accident. Here, the user is simply setting behavior parameters in case that something occurs. In the current legal framework of most countries, the use of autonomous vehicles has not been contemplated and it is hard to judge such vehicles under the same rules that apply to a human driver. However work has started to develop new legal standards for these vehicles. Hence, in the future, a legal framework that contemplates a system like the one proposed here could be possible.

Another question that arises is why not let a body of experts in ethics decide how these vehicles should be programmed. Although this approach could be a solution, the general public may not agree with the experts' conclusions. Examples of different ethical views can be found in many areas of contemporary life, such as abortion and immigration. Debate in these areas has existed for a long time, and even laws established and applied to some countries but not in others. These continuing debates suggest that reaching a consensus over these ethical questions is not straightforward and cannot always be universal.

IV. Conclusions and Future Work

In this paper we have identified a set of ethical perspectives that can be applied to an autonomous vehicle. Ethical concerns for the usage of this vehicles have been identified and evaluated against the corresponding ethical theories and ordered according their importance in each particular theory.

A novel approach for decision making systems has been presented. One of the advantages of this proposal is that it can be used to fit a wide range of user expectations about the behavior of the AVs. A system with this flexibility can be attractive for both industry and users, allowing further development and acceptance of these vehicles.

As a future step, the ordering of the concerns presented in this paper will be implemented in an autonomous vehicle simulation to study the results of this ordering in varying scenarios that reflect those that a vehicle could encounter during its operation.

References

[1] N. J. Goodall, "Can you program ethics into a self-driving car?," *IEEE Spectrum*, vol. 53, no. 6, pp. 28–58, 2016.

[2] E. commission, "Automated and connected driving report," tech. rep., Federal Minister of Transport and Digital Infrastructure, 2017.

[3] IEEE, "IEEE Code of Ethics."

[4] T. Fournier, "Will My Next Car Be a Libertarian or a Utilitarian?: Who Will Decide?," *IEEE Technol Soc Mag*, vol. 35, no. 2, 2016.

[5] J. Gogoll and J. F. Müller, "Autonomous Cars: In Favor of a Mandatory Ethics Setting," *Sci Eng Ethics*, vol. 23, no. 3, pp. 681–700, 2017.

[6] H. Y. Liu, "Irresponsibilities , inequalities and injustice for autonomous vehicles," *Ethics Inf Technol*, vol. 19, no. 3, 2017.

[7] S. Nyholm and J. Smids, "The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem?," *Ethical Theory Moral*, vol. 19, no. 5, pp. 1275–1289, 2016.

[8] N. Goodall, "Ethical Decision Making During Automated Vehicle Crashes," *Transp Res Record*, vol. 2424, no. 2424, pp. 58–65, 2014.

[9] U. Bose, "An ethical framework in information systems decision making using normative theories of business ethics," *Ethics Inf Technol*, vol. 14, no. 1, pp. 17–26, 2012.

[10] C. Fisher and A. Lovell, *Business ethics and values : individual, corporate and international perspectives*. Pearson Education Limited, 2nd ed., 2006.

[11] C. E. Johnson, *Organizational ethics: A practical approach*. United States of America: SAGE Publications, 2nd ed., 2012.

[12] J. Rawls, *A theory of justice*. Cambridge Mass.: Belknap Press of Harvard University Press, 1971.

[13] M. Boylan, *Basic Ethics*. Pearson/Prentice Hall, second ed., 2009.

[14] C. Yoon, "Ethical decision-making in the Internet context: Development and test of an initial model based on moral philosophy," *Comput Hum Behav*, vol. 27, no. 6, pp. 2401–2409, 2011.

[15] J.-F. Bonnefon, A. F. Shariff, and I. Rahwan, "The social dilemma of autonomous vehicles," *Science*, vol. 352, no. 6293, pp. 1573–1576, 2016.

[16] P. Lin, "The Ethics of Saving Lives With Autonomous Cars Is Far Murkier Than You Think," 2013.

[17] A. Hevelke and J. Nida-Rümelin, "Responsibility for Crashes of Autonomous Vehicles: An Ethical Analysis," *Sci Eng Ethics*, vol. 21, no. 3, 2015.

[18] S. Maki and A. Sage, "Self-driving Uber car kills Arizona woman crossing street - Reuters," 2018.

[19] Shepardson David, "Third fatal Tesla Autopilot crash renews questions about system - Reuters," 2019.

[20] T. Chen, F. Chen, and C. Chen, "Review on Driverless Traffic from Management Perspective," *MATEC Web of Conferences*, vol. 124, 2017.

[21] V. Ilkova and A. Ilka, "Legal aspects of autonomous vehicles-An overview," *21st International Conference on Process Control*, 2017.

[22] S. Gless, E. Silverman, and T. Weigend, "If Robots cause harm, Who is to blame? Self-driving Cars and Criminal Liability," *New Crim Law Rev*, vol. 19, no. 3, pp. 412–436, 2016.

[23] A. K. Taylor and S. Bouazzaoui, "Moving forward with autonomous systems: Ethical dilemmas," in *Advances in Human Factors and Systems Interaction*, Springer International Publishing, 2019.

[24] F. Santoni de Sio, "Killing by Autonomous Vehicles and the Legal Doctrine of Necessity," *Ethical Theory Moral*, vol. 20, no. 2, 2017.

[25] I. Coca-Vila, "Self-driving Cars in Dilemmatic Situations: An Approach Based on the Theory of Justification in Criminal Law," *Criminal Law and Philosophy*, vol. 12, no. 1, pp. 59–82, 2018.

[26] A. M. Khan, A. Bacchus, and S. Erwin, "Policy challenges of increasing automation in driving," *IATSS Research*, vol. 35, no. 2, pp. 79–89, 2012.

[27] J. Lederman, M. Garrett, and B. D. Taylor, "Fault-y Reasoning: Navigating the Liability Terrain in Intelligent Transportation Systems," *Public Works Manag. Policy*, vol. 21, no. 1, pp. 5–27, 2016.

[28] J. Fleetwood, "Public Health, Ethics, and Autonomous Vehicles," *Am J Public Health*, vol. 107, pp. 532–537, apr 2017.

[29] D. Leben, "A Rawlsian algorithm for autonomous vehicles," *Ethics Inf Technol*, vol. 19, no. 2, 2017.

[30] G. Contissa, F. Lagioia, and G. Sartor, "The Ethical Knob: ethically-customisable automated vehicles and the law," *Artif Intell Law*, vol. 25, no. 3, pp. 365–378, 2017.

[31] L. Dennis, M. Fisher, M. Slavkovik, and M. Webster, "Formal verification of ethical choices in autonomous systems," *Robot Auton Syst*, vol. 77, pp. 1–14, 2016.

[32] Y. E. Bigman and K. Gray, "People are averse to machines making moral decisions," *Cognition*, vol. 181, pp. 21–34, 2018.

[33] J. Gogoll and M. Uhl, "Rage against the machine: Automation in the moral domain," *J Behav Exp Econ*, vol. 74, pp. 97–103, 2018.

[34] N. Gold, A. M. Colman, and B. D. Pulford, "Cultural differences in responses to real-life and hypothetical trolley problems.," *Judgment and Decision Making*, vol. 9, no. 1, pp. 65–76, 2014.