Autonomous Vehicles and Ethical Issues: Based On a Case Study

Aditi Vakeel
Masters of Computer Science
North Carolina State University
Email: avakeel@ncsu.edu

***Abstract*—Self-driving cars and autonomous vehicles are not the future anymore but the present. The cautious forays into this contemporary environment holding faith are nevertheless only partially secure. In addition to security issues, autonomous vehicles have prompted numerous ethical questions about the models' deployment, design, security, political, economic, and moral implications. This essay addresses the real-life case study of Joshua Brown's death due to the failure of the autopilot feature of the Tesla Model S car and running into a tractor-trailer. It identifies the problem in the current state of regulations of autonomous vehicles and the design of the feature "autopilot," the advanced driver assistant system of Tesla S, proposes solutions and sparks a conversation on the potential ethical and moral ramifications of autonomous vehicles in the future.**

I. Introduction and Background

Scientists and philosophers have expressed ethical and security concerns regarding the development of autonomous vehicles over time. Several discussions, debates, and even protests have been spurred on the questions like How safe is autopilot? How reliable are these systems? Who will be held accountable if the technology fails to prove the safety it claims? In the event of unfortunate accidents, which is primarily to blame? Hereafter, "autonomous vehicles," self-driving cars,"," autonomous cars," and "SDC" will be used synonymously for easy expression and understanding purposes.

Let us first wonder about the question of why ethics are critically important when designing and training an SDC. Driving has become a typical human activity. People are associated with driving over a multitude of emotions like frustration due to roadblocks, joyful long drives with loved ones, good music on a solo ride, a sense of achievement and comfort traveling with family and friends, and fear of accidents. Discussions over scenarios like the famous trolley problem have been going on for years[1][2]. For example, consider a situation where an SDC is driving on a smooth road and finds an obstacle in front. If it has to swerve on either side with no reaction time to apply brakes, what should the system choose if given there is a small boy on the left and an older woman on the right? Will choosing to save the small boy considering that he has not seen the life yet, over the old lady be okay? This is where ethics comes into the picture. Nevertheless, is this not age-based discrimination on moral grounds? The point of discussion here should not be about choosing but analyzing the path that landed on the road to the result, which is why it is essential to design a system ethically rather than only technically.

To understand this even better, consider another hypothetical situation of an SDC where now it has to choose between a fair Hindu on its left and a Muslim man in a dark color on the right side. What would happen if the SDC's design violated moral principles, practiced color discrimination, classified the dark man as an animal rather than a human(which happened in real-life [3]), prioritized the white Hindu as a human being, and then decided to veer into the path of the Muslim man, killing him? This imaginary fatal situation would spark a vast cultural and religious fight between the two groups. The company producing the cars would be claimed to bias towards certain groups, and people would question the right to equality.

II. Case Study

A case study focused on examining the actual incident involving Joshua Brown, who was killed after his Tesla S collided with a tractor-trailer while in "autopilot" mode.

A. Stakeholders

Tesla is one of the most popular electric car manufacturers in the United States. This company has a lot of stakeholders, including communities, customers, employees, and shareholders. As the case study deals with the accident of Joshua Brown's death in a Tesla S model, this section will provide details of the stakeholders of the Tesla company.

a) Investors and Shareholders: Tesla shareholders are those who invest in the company company's stocks and have a say in how it operates. It is a known fact that the company's early years were hugely dependent on investors.

b)Customers: Tesla customers are those who have bought or plan to buy a Tesla product. Customers seek quality at a reduced price. Tesla always comes with new innovations to attract customers. The company also pays much attention to gaining the audience's confidence by reducing battery prices, expanding the charging stations, etc.

c)Communities: Tesla communities are those who live in regions where Tesla operates and can buy their products. For industries like automobiles, communities are crucial. Tesla's community is very active and is quite known for its interest in clean emissions.

d)Employees: Employees are a crucial key to the success of any company apart from customers and other factors.

Tesla takes initiatives for employee benefits through internal competitions, training programs, decision-making, and more.

e)Governments: Tesla needs the support of the government to make its goal of providing sustainable automobiles possible, and government needs to keep Tesla in check regarding rules and compliance for public safety and economic growth of the country.

B. Analysis

Joshua Brown's death was pinned on both Joshua and the autopilot, which failed to see the white truck-trailer against the bright sky and chose to drive under it, possibly thinking it was a bridge. Joshua was also said to have been too distracted to intervene in the autopilot and make a decision. Nevertheless, was that not the whole point of autopilot? How reliable was the Tesla statement, "Digital control of motors, brakes, and steering helps avoid collisions from the front and sides, as well as preventing the car from wandering off the road" [4]? Who is to be blamed for this situation? Tesla should be morally responsible for this fatal situation because it designed a system that:

(i)failed to identify a huge basic tractor-trailer in the bright-day light.

(ii)failed to prevent the collision from the front

(iii)decided to drive beneath the truck and did not stop until it hit the pole. Tesla was exonerated and dusted off the accusations by claiming that the driver of the car was too distracted and had at least 7 seconds before the collision. No amount of blame or proof can replace the lost life. What, then, could Tesla have done to avoid this circumstance? How would the minority be impacted by not agreeing with the majority if there is no unanimity in response to the question: "Are we really geared for self-driving cars?"

III. Solutions and Tests

A. Solution Proposols

The interaction between humans and the system is very crucial in the case of autonomous vehicles. This communication can impact the driver of the car and the people who share the ecosystem with the SDCs. In the case of Brown, he might have been too confident about the Tesla S and might have lost attention in the flow of autopilot. The Tesla S was, however, level 3 automation[5] that required the driver's intervention at appropriate request. This dependency factor cost him his life. Since model S was a beta version and was freely allowed to use a regular commute vehicle, the system should have trained for the worst-case scenarios on highways at the freeways. Three years hence another similar incident involved the death of Jeremy Beren Banner in a Model 3 on a Florida highway after colliding with a tractor-trailer. According to the article [6], the accident was so similar to Brown's case that they stated that Tesla failed to address the issue of the latter. Who takes care of such pressing issues when they are not acknowledged? What regulations are maintained and impacted in the automobile industry? In order to create a better system and avoid fatal scenarios like Brown's, the following alternatives will explore the beliefs, solutions, and different approaches that autonomous vehicle design should take to be ethical, moral, and technically appropriate.

1) Government plays an integral and influential role in road traffic and the rules of any country. The bar for the safety of autonomous vehicles must be set high and rigid since it is associated with peoples' life. Government should consider various protocols and public opinion before making any significant decisions, like should the beta version of the SDC should be granted permission to test on the actual roads with real people. Such kind of decision can create a dilemma. Nonetheless, making a decision is essential.

2) As much as other aspects of the design are discussed today, technical specifications are the first that should be discussed. There will be no point in any ethical or moral discussion if the system fails at a technical level. This was the case with Brown. Though there were other governing concerns, the primary cause of the accident was the failure to identify the object(truck trailer), which was the essential function of the autopilot. The vehicle's systems must concentrate on the worst-case scenarios and the practical and almost impractical scenarios to train the SDC with the best. Considering the above case, Model S should have been trained with scenarios on highways with freeways.

3)The public and academics studying the automobile industry are interested in ethical framework design. Is there a rule book for ethics? It is not essentially required to feed the system only with a set of ethics, but it can be left to the system itself to learn from the environment given initial primary data. However, the SDCs must follow the sole rule of protecting the human inside and outside the ecosystem minimizing the total damage. Howsoever the model might be, it definitely needs human intervention in one way or the other because it is used and made for people themselves. Thus, designing a system with specific moral standards considering the SDCs will be an inevitable part of our future.

4) As technology is growing tremendously fast, it can be expected more than the current level of autonomous driving. Is it appropriate to give the car total control of thought if it could make decisions based on ethical, social, and moral training? According to [7], 90% of accidents are caused due to human error. Suppose we could have a system whose intelligence is far beyond human capabilities and would make wise decisions. In that case, the only human intervention required is training it with ethics and morals. This could be used in cases where the driver is in a dilemma to make decisions while the system takes the decision for him to reduce the total damage in the unfortunate events. This solution is beyond the reach of this essay but not irrelevant. In the case of Brown, it was observed that the cruising speed was far above the actual speed limit for a longer time. As an

autonomous car, the system should have been designed such that at least to warn the driver when exceeding the speed limit after a particular (notably short) period.

5) Sometimes, technology is forced on people due to other pressure. In such circumstances studying the consequences should top the list rather than wasting discussions on how to stop the change. Only in the extreme situation should the change be stopped. When ford addressed the handoff problem, it backed away from producing any level 3 cars[8][9]. This meant there was a technical gap that should have been addressed first. However, Tesla was very aggressive in pushing their limits to prove themselves best and releasing the beta model S even when the handoff problem still needed to be solved.

B. Tests and opinions

a) Harm Test: Solutions 1,2 pose no threat to living beings. Designing a standard ethical frame is very difficult as new rules, laws, and protests constantly evolve, which can be a little challenging for option 3. Option 4,5 can pose a threat if not handled appropriately. For example, even though a few people refuse to use SDCs, they still share the same road with SDCs. Failure of the systems can pose a threat to a manual human driver.

b) Publicity Test: All the options suggested above can be publicized in newspapers, journals, etc. This not only makes people aware of new solutions but also gives them a chance to improve on them using the feedback from the common public.

c) Defensibility Test: The suggested options neither explicitly damage any existing, important, or vulnerable laws nor hurt anybody. The options should be just further scrutinized to include the interests of the minority when designing rules, laws, and frameworks. Thus, the above options pass the defensibility test.

d) Reversibility Test: All the above options were carefully suggested being in the shoes of the outsider, who could be a consumer or who shares the road space with the consumer of SDCs. Thus, all the above options pass the reversibility test, given that they do consider the sentiments of people while actually implementing them. For instance, option 3 demands the system to learn rather than feeding it with a conventional framework, and this means the system learns appropriately. If the system fails to identify humans as human, this could hurt people's feelings. Therefore, it is the designer's responsibility to build it in such a way so that it learns things ethically. If I were given a choice, I would vote to ban any beta versions like Tesla S from being sold to consumers like me.

e) Virtue Test: This is an arguable test. This test, in general, would question the acceptance of SDCs in our life. In case we succeed in building a logical, moral and ethical SDC, people will start using it recklessly. Imagine a world with no tickets for "drunk and drive" cases. As an SDC, the previous driver now becomes a passenger. This act would encourage alcohol consumption indirectly.

g) Profession Test: The above solutions might create a little disturbance with the economics of the company since the system would be scrutinized more effectively this time which would require a lot of resources and trust. Otherwise, the proposed solutions would and should not create critical changes in any other sectors. Tesla's market might take a hit if the government bans the beta- versions of SDCs from being experimented on real roads or customers, which is proposed in the first solution.

h) Organizational Test: In an organizational test, a solution is tested by the legal counsel or ethics officer of the organization. The organizational ethics officer would be highly agreeable to the solutions as all of them do not compromise ethics at any stage of implementation.

IV. Conclusion

The debate on allowing the SDCs into our everyday life is still ongoing. However, this will not alter the fact that the SDC will be our future. The question is, are we ready to accept this change? Are we ready to bring this change? Human error is the most common cause of accidents. Sometimes rare cases of accidents happen, and the SDCs are blamed. These kinds of situations might be inevitable even with the human driver. This should not stop from producing the new technology. Adopting the change, i.e., the SDC will reduce 90% of accidents[10]. Thus, even considering the worst cases accepting SDCs will reduce hundreds of accidents every day in each country across the globe. The 10% more reliability might not be in the near future, but minimal steps can be taken in the future to make the 10% possible. It would be hilarious to expect 100% efficiency in the near future because the technology that cannot understand our emotions in basic texts (produces typos on social media) cannot quickly solve ethical dilemmas.

A tentative choice for the case study and, in general, would be addressing this issue on a governmental level. Government has to establish new and more rigid rules and tests before allowing beta-versions and semi-autonomous vehicles onto the road. However, considering the sources, constraints, and pressures of today's world, it might be a suitable and choice to concentrate on vigorously training the SDC with worst-case scenarios so they succeed in the event of new scenarios like accidents near freeways on highways.

Acknowledgment

I would like to thank Dr. Timothy Menzies for giving me the chance to reflect on and write about software engineering issues that are crucial to today's world and the generations to come.

REFERENCES

[1] Foot, P. (1967). *"The problem of abortion and the doctrine of the double efect."*

[2] Judith Jarvis Thomson, *Volume 59, Issue 2 Philosophical Problems of Death*, April 1976

[3] https://www.nytimes.com/2021/09/03/technology/facebook-ai-race-primates.html

[4] https://www.tesla.com/blog/your-autopilot-has-arrived

[5] SAE International. *"Automated Driving: Levels of Driving Automation Are Defined in New SAE International Standard J3016.* 2014

[6] https://www.theverge.com/2019/5/17/18629214/tesla-autopilot-crash-death-josh-brown-jeremy-banner

[7] https://cyberlaw.stanford.edu/blog/2013/12/human-error-cause-vehicle-crashes

[8] https://www.wired.com/2015/11/ford-self-driving-car-plan-google/

[9] Alex Davies. *"Ford's Skipping the Trickiest Thing about Self-Driving Cars." Wired.* November 10, 2015.

[10] https://www.theatlantic.com/technology/archive/2015/09/self-driving-cars-could-save-300000-lives-per-decade-in-america/407956/