Gesture Based Robotic Vehicle

Trust and Autonomous Vehicles

A Thesis Project Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science

University of Virginia Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree

Bachelor of Science, School of Engineering

Ian Le

Fall, 2023

Technical Project Team Members: Ruhul Quddus, Nima Razavi, Goutham Mittadhoddi, Kenny Zhang

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signature ____

Approved _______________________________________ Date ________________________

William Davis, Assistant Professor of STS, Department of Engineering and Society

Approved _______________________________________ Date ________________________

Adam Barnes, Senior Lecturer, Department of Electrical and Computer Engineering

Overall Introduction (Comprehensive Problem Frame)

The advancement of computer technology in self-driving vehicles over the last decade has been staggering. The Google Self Driving Car Project, now known as Waymo, started all the way back in 2009 after successes in the Google X Lab. In 2020, Waymo became the first fully autonomous ride service in Phoenix Arizona and is looking to expand its serviceable area across the United States (Waymo 2023). Despite successes in autonomous technology driven by tech giants such as Google, Microsoft, and Uber, the general public is generally untrusting of autonomous vehicles. The Pew Research Center reports that 44% of Americans are against the widespread adoption of autonomous vehicles, while only 26% hold a positive view, supporting the idea that autonomous vehicles are beneficial (Center, P. R. 2017). This begs the question of why a majority of Americans are opposed to riding in an autonomous vehicle? What is the difference between a human driving the car as opposed to an autonomous system? My thesis will explore the concept of trust in relation to autonomous vehicles and examine how it plays a role in shaping public attitudes and the adoption of autonomous systems.

Technical Topic (Problem frame for technical topic)

My technical project is a small gesture driven robotic car. My group and I have created a glove that uses gyroscopic sensors to interpret hand gestures and send those instructions to the car. In turn, the car has a mounted camera and distance sensor that relay information about where the car is through video and haptic feedback. This project is aimed to be a toy for the young teen tech enthusiast. As a small toy, our car has rudimentary systems for preventing crashes compared to large scale systems such as autonomous cars. Our project has only two sensors to mitigate crashes, a simple ultrasonic distance sensor and a camera to monitor where the car is. Through driving with such limited sensing apparatus, it becomes clear why self-driving cars have such an intricate system of sensors. Through the usage of the car, people can understand some of the limitations and benefits of sensors in a robotic system through a fun and engaging platform. While the car remains fully under human control, one can experience the world similarly to an autonomous robotic system, with a view of the real world narrowed down to what is perceivable through sensors.

STS Topic (Problem frame for STS topic)

As far as technology is concerned, the vehicle perceives many times more information than a human can and can make decisions in fractions of a millisecond (Greenblatt, 2016). In an ideal world, autonomous cars would be safer, as they in theory have plenty of advantages over a human driver. A properly programmed system would always follow the law, would never tire or would be affected by intoxicants, and have inhuman reaction speed and perception. In a simulation run by Google, researchers estimated that the Waymo system could prevent 82% of collisions (Scanlon et Al, 2021). Despite the developments in autonomous technology, there are still reasons for concern over their use. Machine learning and algorithms used by autonomous vehicles are difficult to understand, and challenges with the technology lead to questions about the system's ability to handle real world events. Despite its shortcomings, the rapid pace in innovation of autonomous vehicles as well as the potential benefits and safety make a compelling case for not dismissing the technology outright. The potential benefits of autonomous autonomous vehicles are numerous, yet their adoption is generally opposed. This comes down to a lack of trust in the technology.

Part of the lack of trust in this technology is a lack of knowledge about autonomous vehicle technology itself. When discussing autonomous vehicles the idea of a "trolley problem," wherein an actor must decide to sacrifice an individual for the greater good, inevitably arises in the discussion. A popular example of this approach is the MIT Moral Machine (Awad et. al. 2018), which polled users to make decisions for a theoretical car to hit one group of individuals over another. This creates several issues, as the trolley problem assumes a deterministic set of actions, whereas the consequences of actions are not entirely certain (Nyholm 2018a). As a model, it is easiest to reduce the problem to this simple binary dilemma as opposed to how current autonomoyus systems operate. Current technologies currently rely heavily on machine learning models to determine the course of the autonomous vehicle. Machine learning uses algorithms to calculate probabilities of future events, then uses these probabilities to predict the best course of action. At the current stage of artificial intelligence, we cannot instruct a binary morality system; rather machine learning uses previous data to predict future outcomes. Machine learning is great at handling a majority of complex situations, where the most probable course of action is constantly changing and updated.

However there are some limitations that machine learning has that raises issues with trust. Machine learning is that it is somewhat of a "black box" wherein humans are unable to understand exactly which calculations are being done to determine the predictions. We can see the outcomes of the predictions and tune the algorithms to be more precise, but we cannot change the internal algorithms themselves (Goodall 2016). This makes the idea of the trolley problem largely incompatible with machine learning training. Therefore when an accident does happen, it is difficult to understand the "thinking" of an autonomous system. In the case of the fatality in Arizona, the car's sensors were unsure if the pedestrian walking a bike was a vehicle, a pedestrian, or a bicycle (York, W. P. 2023). We may be able to understand that the algorithm was unable to figure out what the pedestrian was doing, but the internal algorithms that the car used to come to that decision are obstructed. Another cause for a lack of trust is a lack of data in niche situations. Machine learning uses past data to predict future events, and accident data is relatively uncommon compared to available datasets of normal everyday driving. Accidents happen in a matter of seconds, and out of millions of cars on the road, only a mere fraction of them experience crashes, and of those only a few are equipped to record that data into a usable format for machine learning. It is much easier to tell the car what to do in normal operation such as driving around the city than what to do in the case of a potential accident. The ambiguity around how these systems make decisions is understandibly concerning, as transparency around the decisions being made is necessary to gain trust. A human telling you the reasons for its decisions is much more approachable than an engineer explaining that the algorithm simply did not work for an unknown reason.

Transparency around the companies developing these autonomous systems is another area where trust can be formed. In a paper put forth by Waymo, they argued that autonomous vehicle safety regulations should be collaborative between lawmakers and self-driving car companies and urged that other companies follow suit (Favaro et Al. 2023). This sets forth a positive precedent for developing transparent communications around how these systems are being developed and tested. However as a company, Waymo still a for profit initiative, and have also decided to sue the California DMV to protect its crash data, citing trade secrets as well as user data as reasons for their lawsuit. As one of the largest private companies in the world, there is reason for trust and distrust to arise from the general public. On one hand, one could argue that being one of the largest companies in the world would give them the resources they need to ensure safety for their consumers, yet on the other hand their primary interest as a company is profit.

Another reason for lack of trust around autonomous vehicles is a lack of clear accountability when an accident occurs. When two human drivers crash, all parties give statements to what occurred, and the situation is determined. Once the situation is determined, blame is assigned to the parties who acted in violation of the law. Unless in the case of catastrophic failure, autonomous vehicles by their very nature gather large amounts of data during their operation. This collection of data can reproduce the scene of the crash in detail far exceeding the ability of a human. This can reduce uncertainty in determining the details of the situation drastically (Goodall 2016). While we may not entirely understand the algorithms behind the decision making of the car, the data collected is useful for a human juror to also weigh judgment. Through sifting through the data, a human can gain understanding to how the accident occured as well as the “reasoning” behind the car’s actions. The data can also be used to discover who was at fault legally to a more accurate degree. While the collection of data can resolve the details of legality and blame, it still remains unclear as to who the blame falls to in the case that it was infact the autonomous vehicles fault. In the case of a fully autonomous system the precedent for blame likely falls to the manufacturer. In the case of accidents caused by a faulty mechanical component, blame falls to the manufacture of the vehicle. Like any system, the autonomous portion of the car could be treated as a component in the system. In the US legal system, suing the company responsible for the system failure usually falls under punitive damages, or damages that exceed compensation for the victim and are awarded to punish the defendant are relatively common in the United States. These costs are much higher than settlements between two human drivers, this gives manufacturers even more incentive to improve safety (Greenblatt, 2016).

Finally, the trust in autonomous systems must stem from our expectations of these systems. When asked about the largest advantages of autonomous cars, safety was not seen as one of the biggest advantages, rather the prospect of arriving at the destination while doing other activities or relaxing was seen as the largest advantage. And when asked about how often they would use the theoretical car, they would use it primarily for short drives for convenience (Nielsen and Haustein, 2018). This frames the automated car as a luxury device, rather than a safety device. I believe that this demonstrates that autonomous vehicles are generally perceived as unsafe. Additionally, expectations of autonomous vehicles are much higher than that of regular vehicles with 87% of Americans polling that driverless vehicles should be tested using a higher standard(Center, P. R. 2017). This shows distrust in the safety of the technology as one of the foremost issues with the adoption of autonomous vehicles. This is to be expected as a new technology enters society, it is first met with skepticism, but if it performs well it generally is accepted (Hughes, 2012). Most people have not been driven by an autonomous vehicle, so do not trust these vehicles and are impressed when they can complete the same task as a human. As this technology becomes more prevalent, public opinion will improve if the technology continues to perform positively.

Conclusion

In conclusion the apprehensiveness around the adoption comes partly from a lack of knowledge about autonomous vehicles and their operation. As autonomous vehicles are an entirely new field, the unfamiliarity with their intricate systems and decision-making processes contributes significantly to public skepticism. My thesis will explore further reasons for distrust for autonomous vehicles and how future efforts might change the perception of autonomous transportation.

References

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59–64. https://doi.org/10.1038/s41586-018-0637-6

Borenstein, J., Herkert, J. R., & Miller, K. W. (2019). Self-driving cars and engineering ethics: The need for a system level analysis. *Science and Engineering Ethics*, *25*(2), 383–398. https://doi.org/10.1007/s11948-017-0006-0

Center, P. R. (2022, March 17). 4. Americans cautious about the deployment of driverless cars. *Pew Research Center: Internet, Science & Tech*. https://www.pewresearch.org/internet/2022/03/17/americans-cautious-about-the-deployment-of-driverless-cars/

Du, H., Zhu, G., & Zheng, J. (2021). Why travelers trust and accept self-driving cars: An empirical study. *Travel Behaviour and Society*, *22*, 1–9. https://doi.org/10.1016/j.tbs.2020.06.012

Goodall, N. J. (2016). Can you program ethics into a self-driving car? *IEEE Spectrum*, *53*(6), 28–58. https://doi.org/10.1109/MSPEC.2016.7473149

Greenblatt, N. A. (2016). Self-driving cars and the law. *IEEE Spectrum*, *53*(2), 46–51. https://doi.org/10.1109/MSPEC.2016.7419800

Hong, J.-W., Cruz, I., & Williams, D. (2021). AI, you can drive my car: How we evaluate human drivers vs. self-driving cars. *Computers in Human Behavior*, *125*, 106944. https://doi.org/10.1016/j.chb.2021.106944

Kusano, K., & Victor, T. (2022). Methodology for determining maximum injury potential for automated driving system evaluation. *Traffic Injury Prevention*, *23*(sup1), S224–S227. https://doi.org/10.1080/15389588.2022.2125231

Nielsen, T. A. S., & Haustein, S. (2018). On sceptics and enthusiasts: What are the expectations towards self-driving cars? *Transport Policy*, *66*, 49–55. https://doi.org/10.1016/j.tranpol.2018.03.004

Nyholm, S. (2018a). The ethics of crashes with self‐driving cars: A roadmap, I. *Philosophy Compass*, *13*(7), e12507. https://doi.org/10.1111/phc3.12507

Nyholm, S. (2018b). The ethics of crashes with self‐driving cars: A roadmap, II. *Philosophy Compass*, *13*(7), e12506. https://doi.org/10.1111/phc3.12506

Nyholm, S., & Smids, J. (2016). The ethics of accident-algorithms for self-driving cars: An applied trolley problem? *Ethical Theory and Moral Practice*, *19*(5), 1275–1289. https://doi.org/10.1007/s10677-016-9745-2

Scanlon, J. M., Kusano, K. D., Daniel, T., Alderson, C., Ogle, A., & Victor, T. (2021). Waymo simulated driving behavior in reconstructed fatal crashes within an autonomous vehicle operating domain. *Accident Analysis & Prevention*, *163*, 106454. https://doi.org/10.1016/j.aap.2021.106454

Schwall, M., Daniel, T., Victor, T., Favaro, F., & Hohnhold, H. (2020). *Waymo public road safety performance data*. https://doi.org/10.48550/ARXIV.2011.00038

Waymo—Self-driving cars—Autonomous vehicles—Ride-hail. (n.d.). Waymo. Retrieved October 24, 2023, from https://waymo.com/

York, W. P., New. (2023, October 26). *Driverless Uber car 'not to blame' for woman's death*. https://www.thetimes.co.uk/article/driverless-uber-car-not-to-blame-for-woman-s-death-klkbt7vf0