Assessing the Reliability of Self-Driving Cars: A Comprehensive Analysis

Ada Singh

Student, Mayo College Girls' School

Abstract

This review paper explores the ethical issues surrounding the deployment of self-driving cars, with a particular focus on the well-known trolley problem. As autonomous vehicles transition from concept to reality, they must be programmed to handle moral dilemmas, such as deciding whether to sacrifice one life to save several others in unavoidable accident scenarios. The paper examines various ethical frameworks and discusses how these principles can be integrated into the decision-making algorithms of autonomous vehicles. It highlights the necessity for transparency, accountability, and public engagement in shaping these technologies. Looking ahead, the paper suggests that ongoing interdisciplinary research and regulatory development will be crucial in addressing these ethical challenges, ensuring that self-driving cars are safe, reliable, and ethically sound.

Keywords: Self Driving cars, Lidar technology, Radar technology, Ethical issues

Introduction

Self-driving cars, also known as autonomous vehicles, represent a revolutionary shift in the transportation sector, leveraging advanced technologies to navigate and operate without human intervention. The concept of autonomous vehicles dates back several decades, with early prototypes emerging in the 1980s when Carnegie Mellon University's Navlab and ALV (Autonomous Land Vehicle) projects began testing computer-controlled vehicles. This era saw significant advancements in robotics and Artificial Intelligence, laying the groundwork for more sophisticated developments.

By the 2000s, companies like Google (now Waymo) entered the field, driving significant progress through extensive testing and the integration of cutting-edge sensors, machine learning, and real-time data processing. Today, numerous automotive and tech companies, including Tesla, Uber, and traditional car manufacturers like Ford and General Motors, are heavily invested in bringing fully autonomous vehicles to the mainstream market.

However, the advent of self-driving cars is accompanied by a host of ethical challenges. One major concern is the decision-making process of autonomous systems in critical situations. For instance, the dilemma of prioritizing passenger safety over pedestrian safety in unavoidable accident scenarios raises profound moral questions. Additionally, the potential for job displacement in driving-related sectors, data privacy issues given the vast amount of information these vehicles collect, and the broader societal implications of machine-driven decision-making are areas of ongoing ethical debate. As technology continues to evolve, addressing these ethical issues remains crucial to the responsible deployment of self-driving cars.

Automated vehicle technologies span a broad spectrum of capabilities, from basic driver assistance features to fully autonomous driving systems. According to the Society for Automotive Engineers (SAE) taxonomy established in 2014, there are five defined levels of vehicle automation, each representing a step towards complete autonomy.

Level 0 (No Automation): This level signifies the absence of any automation, where driving is entirely manual.

Level 1 (Driver Assistance): At this stage, the vehicle technology can take over either longitudinal control (speed and distance) or lateral control (steering), but not both simultaneously. Examples include Adaptive Cruise Control (ACC), which manages the vehicle's speed based on the flow of traffic, and Lane Keeping Assist (LKA), which helps the driver stay within the lane.

Level 2 (Partial Automation): This level integrates multiple automated control functions, such as combining adaptive cruise control with lane centering. While the vehicle can manage both steering and speed, the driver must remain attentive, monitor the roadway, and be ready to take over control at any moment.

Level 3 (Conditional Automation): At this point, the vehicle handles all driving tasks under certain conditions, allowing the driver to engage in non-driving activities. However, the driver must be prepared to regain control when prompted, typically within a specified time frame, such as 30 seconds after receiving a warning signal.

Level 4 (High Automation): Vehicles at this level can perform all driving tasks and do not require driver intervention in most situations. Nevertheless, high automation may be restricted to certain environments, such as highways or mapped urban areas.

Level 5 (Full Automation): The pinnacle of vehicle automation, Level 5, enables the vehicle to execute all driving functions and monitor roadway conditions for the entire journey. These vehicles can operate without any human occupants or with occupants who cannot drive, regardless of the environment or road conditions.

The progression through these levels highlights the gradual shift from manual to fully autonomous driving, emphasizing the increasing reliance on advanced technology to enhance safety, convenience, and efficiency in transportation. However, as vehicles become more autonomous, addressing the associated ethical and regulatory challenges becomes increasingly vital to ensure their safe and equitable integration into society.

The SAE taxonomy reveals a transformative shift in the driver's role with increasing vehicle automation, from active control at Level 0 to passive passenger at Levels 4 and 5. This evolution raises significant human factors issues, especially at Level 3, anticipated for widespread introduction by the decade's end (NBC News, 2014). At this level, drivers must resume control with sufficient transition time when the system's limits are reached or when they choose to drive manually. Ensuring safe transitions between automated and manual control involves challenges in control authority, human-machine interface design, transition strategies, and driver performance over time. Research focuses on the safety impact of secondary tasks, situation awareness, driver acceptance and trust, training, and system evaluation tools. Addressing these challenges is essential for the safe, efficient, and accepted integration of autonomous vehicles into transportation systems.

Case Study

The Trolley Problem

In the ‘switch’ case, a driverless trolley is on course to kill five people stuck on the tracks unless it is redirected to a side track. You are positioned next to a switch that can divert the trolley, saving the five but resulting in the death of one person on the side track. Despite the moral dilemma, a common response, as noted by Greene (2013), is that it is permissible to pull the switch to save the five, even though it leads to the death of the one individual.

In a variation known as the ‘footbridge’ case, saving the five requires a different approach. Here, you are on a footbridge above the tracks with a very large and heavy man whose body mass is sufficient to stop the trolley if he is pushed onto the tracks. However, this action would kill him. The moral question is whether it is permissible to push the man to his death to save the five. According to Greene (2013), a common response is that it is not permissible to sacrifice the man in this manner. Thus, while saving the five by sacrificing one person in the ‘switch’ case seems morally acceptable to many, doing so in the ‘footbridge’ case appears wrong to most.

Many people casually refer to one or both of these examples as "the trolley problem." However, some influential philosophers, such as Judith Jarvis Thomson, use the phrase to denote a more specific issue. According to Thomson (2008), the basic trolley problem is about explaining the asymmetry in our moral judgments: why is it permissible to save five people by sacrificing one in the ‘switch’ case, but not in the ‘footbridge’ case? Others adopt a broader interpretation, suggesting that the trolley problem also encompasses similar moral dilemmas that do not involve trolleys at all.

According to Frances Kamm, the fundamental philosophical issue is understanding why certain people, using specific methods, are morally permitted to kill a smaller number of people to save a larger number, while others, using different methods, are not. For example, why is it impermissible for a doctor to save five patients needing organ transplants by "harvesting" organs from a healthy individual who came in for a routine check-up? This scenario, although not involving trolleys, is considered by Kamm to fall under the broad umbrella of the trolley problem because it raises similar ethical questions about the permissibility of sacrificing one life to save many (Kamm 2015).

These various thought experiments have been employed to explore numerous normative issues in moral philosophy. They have been used to investigate the distinction between: (i) "positive" and "negative" duties, meaning duties to perform certain actions versus duties to refrain from certain actions; (ii) killing versus letting die; and (iii) consequentialism versus non-consequentialism in moral theory, distinguishing theories focused solely on promoting overall good from those considering other factors (Foot 1967; Thomson 1985; Kamm 2015). In recent years, these experiments have also been utilized to empirically study the psychology and neuroscience behind different types of moral judgments (Greene 2013; Mikhail 2013).

Literature Review

Claudine Badue (2021). This paper has done research on self-driving cars, particularly those developed since the DARPA challenges, focusing on vehicles equipped with autonomy systems classified as SAE level 3 or higher. The autonomy system of self-driving cars is usually divided into two main components: the perception system and the decision-making system. The perception system includes various subsystems responsible for tasks such as localization, mapping static obstacles, detecting and tracking moving obstacles, road mapping, and recognizing traffic signals. The decision-making system is also divided into subsystems that handle tasks like route planning, path planning, behavior selection, motion planning, and control. This survey outlines the typical architecture of self-driving car autonomy systems, reviews research on relevant perception and decision-making methods, and provides a detailed description of the autonomy system architecture of the Intelligent Autonomous Robotics Automobile (IARA) developed at the Universidade Federal do Espírito Santo (UFES). Additionally, we highlight notable self-driving car research platforms developed by both academic institutions and technology companies, as reported in the media.[1]

Peng Liu (2018). Self-driving vehicles (SDVs) hold the potential to significantly reduce traffic accidents. A key question for the public, automakers, and governments is, "How safe is safe enough for SDVs?" To address this, a new expressed-preference approach has been introduced to determine the socially acceptable risk for SDVs. In a between-subject survey involving 499 participants, we assessed the respondents' risk-acceptance rates across different scenarios with varying traffic-risk frequencies, exploring the logarithmic relationships between traffic-risk frequency and risk-acceptance rate. We compared logarithmic regression models of SDVs with those of human-driven vehicles (HDVs), finding that SDVs need to be safer than HDVs. When the same traffic-risk-acceptance rates were applied to both SDVs and HDVs, we predicted and compared their associated acceptable risk frequencies. Two criteria for risk acceptance emerged: the tolerable risk criterion, suggesting that SDVs should be four to five times safer than HDVs, and the broadly acceptable risk criterion, indicating that half of the respondents desired SDVs' traffic risk to be two orders of magnitude lower than the current estimated traffic risk. This approach and its findings could aid government regulatory bodies in establishing precise safety standards for SDVs.[2]

Cyreil Deils (2016). This paper addresses the predicted increase in the occurrence and severity of motion sickness in self-driving cars. While self-driving cars promise significant benefits, including increased comfort and productivity for drivers, our findings indicate that these scenarios may lead to a higher risk of motion sickness. This increased risk could prevent the full realization of the technology's benefits, particularly for those already prone to motion sickness, potentially hindering user acceptance and adoption. This, in turn, could limit the socioeconomic advantages of this emerging technology. We discuss the causes of motion sickness in self-driving cars and provide guidelines for designing automated vehicle technologies to minimize or prevent motion sickness. The paper also explores less known consequences of motion sickness, such as postural instability and impaired task performance, and their implications for the use and design of self-driving cars. We conclude that fundamental perceptual mechanisms must be considered in the design process, and self-driving cars should not simply be treated as living rooms, offices, or entertainment spaces on wheels.[3]

Sven ove hansson (2021). The introduction of self-driving vehicles raises numerous ethical issues beyond the commonly narrow focus on improbable dilemma-like scenarios. This article provides a comprehensive overview of realistic ethical concerns related to self-driving cars. Key topics include the potential for strong opinions for and against driverless cars to lead to significant social and political conflicts. A low tolerance for accidents caused by driverless vehicles may delay the adoption of systems that could substantially reduce overall risks. There will be trade-offs between safety and other requirements within the road traffic system. Over-reliance on the quick collision-avoidance capabilities of self-driving cars could encourage dangerous behaviors, such as pedestrians stepping in front of moving vehicles, trusting in their ability to brake quickly. Children traveling alone might ignore safety protocols, like using seatbelts. Digital information about routes and destinations could be exploited to convey commercial and political messages to users. If fast passage can be purchased, it may result in socio-economic segregation in road traffic. Additionally, terrorists and criminals could hack into vehicles to cause crashes, use them to transport bombs to detonation sites, or disrupt a country's road system.[4]

Sven Nyholm (2016). The ethics of programming self-driving cars to handle unavoidable collisions invites comparison to the classic trolley problem, but there are crucial distinctions between the two that must be considered. The decision-making context in the trolley problem involves an individual making a split-second decision about whether to divert a trolley to save five people at the cost of one person's life, with immediate control and facing a direct moral choice. In contrast, for self-driving cars, the decisions are pre-programmed by engineers and developers long before any potential accident occurs, involving anticipation of numerous potential situations and outcomes in a controlled environment with time for deliberation. The moral and legal responsibility in the trolley problem lies with the person who acts, exploring personal moral intuitions and responsibilities, whereas in self-driving cars, responsibility is distributed among programmers, manufacturers, regulatory bodies, and vehicle owners, raising complex legal questions about liability for algorithmic decisions. Furthermore, while the trolley problem presents a situation framed as a certainty, where pulling the lever will definitely save five lives and kill one, real-world scenarios for self-driving cars involve significant uncertainty and risk assessment, requiring programmers to consider probabilities, imperfect sensor data, and complex environmental factors, making decisions based on risk management and statistical probabilities rather than certainties. This highlights the need for engineers to encode responses that balance safety, legality, and ethical considerations, and underscores the complexity of distributing responsibility and accountability among various stakeholders.[5]

Jason Borenstein (2017). In this paper, we review Richard De George's classic article on the moral responsibilities of engineers in the infamous Pinto case and consider whether his analysis is valid in an era of pervasive and autonomous technologies. We undertake a contemporary analysis of the topic as it pertains to engineers who are designers of self-driving cars, applying the "Moral Responsibility for Computing Artifacts: The Rules," a framework developed by an ad hoc interdisciplinary group of computing professionals, engineers, and ethicists. While engineers and engineering managers are not necessarily "crazed," we argue that ethical analysis needs to be integral to the design of self-driving vehicles. Engineering and other relevant communities need to engage with the issue of what it means to uphold one's ethical and professional responsibilities in the era of these vehicles. Designers of the technology should diligently and creatively exercise their moral sensitivity capacities to uphold their obligations to the public. Integrating this activity into their decision-making process is a critical element in the realm of "anticipatory technology ethics" as defined by Brey and in the realm of "responsible research and innovation" as described by Sutcliffe.[6]

Sven Nyohlm (2018). Self-driving cars promise to be significantly safer than regular cars, but they can never be completely safe. Therefore, they need to be programmed to handle crash scenarios. Should these cars be programmed to always prioritize their owners, minimize harm, or follow some other principle? The article first explores whether all vehicles should have uniform "ethics settings." It then examines the common analogy with the trolley problem. Following this, the article assesses recent empirical research on laypeople's attitudes towards crash algorithms in relation to the ethical issue of crash optimization. Finally, it discusses what traditional ethical theories such as utilitarianism, Kantianism, virtue ethics, and contractualism suggest about how cars should handle crash scenarios. The aim is to provide an overview of the existing literature on these topics and evaluate the current state of the discussion.[7]

Conclusion

Self-driving cars represent a significant advancement in technology, promising increased safety, efficiency, and accessibility in transportation. However, their deployment brings complex ethical issues to the forefront, particularly illustrated by variations of the trolley problem. Autonomous vehicles must be programmed to make split-second decisions in critical situations, such as choosing to sacrifice one life to save multiple others. Addressing these dilemmas involves creating transparent algorithms that prioritize human safety and establishing regulatory frameworks to guide ethical decision-making. Solutions may include programming cars to minimize harm, using collective ethical input to shape decision-making algorithms, and implementing robust oversight to ensure accountability. As technology progresses, continuous dialogue among technologists, ethicists, regulators, and the public will be essential to navigate these ethical challenges and ensure the responsible integration of self-driving cars into society.

References

1. Claudine Badue et al. Self-driving cars: A survey. *Expert Systems With Applications*. 165 (2021). 113816. https://doi.org/10.1016/j.eswa.2020.113816
2. Peng Liu et al. How Safe Is Safe Enough for Self-Driving Vehicles? *Risk Analysis*. 3. 5. 2018. DOI: 10.1111/risa.13116
3. Cyriel diels et al. Self-driving carsickness. *Journal of Applied Ergonomics*. 53 (2016). 374 - 382. http://dx.doi.org/10.1016/j.apergo.2015.09.009
4. Sven Ove Hansson. Self-Driving Vehicles—an Ethical Overview. *Philosophy & Technology*. 34 (2021). 1383 -1408. https://doi.org/10.1007/s13347-021-00464-5
5. Sven Nyholm. The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem? Journal of Ethical Theory and Moral Practice. 19 (2016). 1275 - 1289. DOI 10.1007/s10677-016-9745-2
6. Jason Borenstien et al. Self Driving Cars - Ethical responsibilities of design engineers. *IEEE technology and Society Magazine*. 2017.
7. Sven Nyholm. The ethics of crashes with self-driving cars: A roadmap. Philosophy Compass. 2018. https://doi.org/10.1111/phc3.12507