Passenger Prioritization in Autonomous Vehicle Collisions

Josh Miao

12/18/2025

ECE 472 - Deep Learning, Cooper Union

Contents

Introduction

Autonomous Vehicles

The integration of Deep Learning into automotive vehicles(AV) has yielded great advancements in the field of AVs. It was only a few years ago when AVs transitioned from science fiction to active modes of transportation in global transit. Advocates of AVs assert the improved their increased safety as they eliminate the "human factor" present in an overwhelming majority of motor vehicle collisions The National Traffic Safety Administration theorize a 94% reduction of vehicle crashes with fully autonomous cars (National Highway Traffic Safety Administration 2015). However, humans drivers still a significant amount of control at the wheel and the 94% reduction in collisions apply to fully autonomous vehicles. In practice, humans still retain the ability to avoid accidents or be distracted at the wheel in our commercially available automation levels.

Our society is trending in the direction of increasing autonomy and as these artificial intelligence(AI) systems take more control, the design of crash-optimization algorithms becomes paramount. Subsequently, ethical dilemmas in unavoidable collision scenarios remain a point of contention. While some argue for a utilitarian framework minimizing total harm, a broader scope reveals the value of prioritizing passenger safety.

Discussion Hypothetical

The Society of Automotive Engineers classifies autonomous vehicles into 5 levels ranging from momentary assistance like emergency breaking to full autonomy in all conditions (International 2021). This unavoidable collision scenario will assume a level 5, fully autonomous, motor vehicle so that the collision avoidance algorithm is in complete control of the scenario. The hypothetical in the current discussion proposes the general scenario of a unavoidable car collision with a pedestrian. Other considerations common in trolley hypotheticals such as demographic and quantity of individuals are irrelevant.

The most obvious stakeholders in this scenario are the pedestrian(s) and vehicle occupant(s). However, the vehicle manufacturer has stake in their brand reputation, potential legal consequence, and future sales. Additionally, all road users and the general public has stake due to the broader implications of crash optimization algorithms that this discussion will explore.

Assertions of the Current Discussions

This discussion ultimately asserts that AVs should prioritize the safety of the vehicle occupants. This position is supported by three primarily pillars: legal duty of cared owed by manufacturers to users, the prevention of moral hazard in pedestrian behavior, and the adoption paradox where utilitarian programming prevents the widespread use of the very technology intended to save lives.

Ethical Considerations

The trolley problem is often utilized as an abstraction to AV ethics. However, it is frequently criticized as an inadequate analogy. The context of motor vehicles and their societal prevalence demands a discussion of the broader context of norms and role-based responsibilities (Nyholm and Smids 2016). Subsequently, the vacuum of the trolley problem immediately diminishes its relevance in this discussion.

Manufacturer Duty of Care

The relationship between a car manufacturer and its passenger(s) is innately fiduciary. Similar to a doctor and their patient or a lawyer and their client, the professional context creates an obligation to the client that cannot be circumvented for any potential greater good (Gurney 2016). Just as HIPAA forbids a physician from sacrificing a healthy patient others or dispelling potentially incriminating information to law enforcement, an AV manufacturer is obligated to not program a vehicle to sacrifice its occupant to save even a much larger number of bystanders.

Additionally, any vendor of goods or services where a customer's safety is at risk demands the prioritization of customer well-being. The existence of federal agencies such as the Food and Drug Administration(FDA) or Consumer Product Safety Commission(CPSC) imply the obligation of customer protection in a product. In the context of motor vehicles, the NHTSA establishes safety standards that guarantees a baseline of safety for vehicle occupants. Consequently, AV manufacturers have an obligation to car owners to protect their physical well-being. A crash-optimization algorithm choosing the harm of the vehicle occupants in any case is a clear violation of the manufacturer-customer obligation.

Moral Hazard and Pedestrian Behavior

A strict pedestrian-policy creates backwards incentives that could ironically decrease overall traffic safety through risk compensation.

Strategic Interaction

Millard-ball utilizes game theory to show that if AVs are known to be "altruistic" and avoid a pedestrian at the expense of occupants, pedestrians will adopt riskier behaviors (Millard-Ball 2016). Under the expectation the the AV is programmed to avoid the pedestrian at all costs, pedestrians may jaywalk, obstruct traffic, or perform generally unsafe behaviors, leading to an increase in dangerous situations. This is in clear opposition to the original intention of road safety.

If pedestrian safety was prioritized despite this, cars would need to perform more emergency maneuvers yielding further unsafe road habits. Alternatively, lower speed limits could be enforced to compensate but this would lead to less efficient transportation and a reinforcement of bad pedestrian behavior, further facilitating the opposite of pedestrian safety. Therefore, the prioritization of pedestrian safety in this way would create a feedback loop of increasingly unsafe pedestrian behavior.

This behavior is grounded in the Peltzman Effect (Hedlund 2000). When safety regulations, like AV sacrifice algorithms, decrease the perceived risk of an activity, individuals respond by increase their risky behavior. If the cost of riskily entering a roadway is shifted entirely to the car's occupant, the number of unavoidable scenarios will likely proliferate, negating any and all safety gains provided by the algorithm.

Predictability is a Safety Mechanism

The predictability of a transit system and its consequences to risky behavior facilitate universal efficiency in transportation.

Rail Transit as an Analogy

Consider the norms and obligations present in rail transportation. Globally, trains are generally efficient due to their predictability in timing and routes. They move tens of millions of people daily (Ivanov 2023). One component of their predictability is that pedestrians do not challenge the train in its routs. The physical and algorithmic outcome is certain; the train will stay its course, regardless of any obstacle made of flesh. The few unfortunate cases of train-pedestrian collisions while tragic, reinforce the predictability of what happens if an individual obstructs a train's course. Consequently, pedestrians have effectively learned to avoid any chance of collision with a train by staying off of the tracks. Similarly, roadway pedestrians can learn to avoid dangerous situations by taking extra precaution when deciding to obstruct an AV's path.

Platooning

As a result of predictability in rail transit, train traffic controllers may make the assumptions necessary to automate and increase density of rail transport with minimal fear of pedestrian obstruction. The future of car transportation lies in vehicle-to-everything(V2X) communication, enabling cars to travel in "platoons" with minimal following distances (Alam et al. 2016). In such a system, one unpredictable pedestrian actions could result in a cascading collisions, resulting in significantly higher damages than a single localized collision with a pedestrian.

Algorithmic Uncertainty

AVs operate with finite information that will be a complete representation of the physical environment. This uncertainty has serious consequences in the context of roadway safety.

Entropy is an appropriate analogy for the uncertainty of road transportation. The low entropy choice is an accident minimizing unforeseen consequences. One such low entropy choice is striking the pedestrian to avoid dramatic maneuvers.

Tolerance for Uncertainty

AVs operate on probabilistic data. A sensor will inevitably detect incorrectly detect a "pedestrian", a false positive (Goodall 2014). An algorithm programmed to sacrifice the health of the passengers for the pedestrian will sacrifice the well-being of its passengers in vain. A tolerance must therefore exist to account for situations such as these.

Unforeseen Consequences

Evasive maneuvers will yield of consequences of great uncertainty that are impossible to predict in entirety. Limited information and the limitations in the ability to extract meaningful insights from that data means unavoidable risk in swerving or braking. The vehicle may hit an undetected cyclist, structural support, or necessitate evasive maneuvers for other vehicles. Evasive maneuvers are therefore high entropy decision that may result in much greater loss than the low entropy decision of minimal action. The most conservative and reliable safety maneuver in high-uncertainty conditions is to stay the course.

Adoption Paradox

Utilitarian critics argue that the pedestrian has much less protection and is therefore worthy of the protection in a collision-optimization algorithm Lin 2016. However, this argument is counterintuitive even from a utilitarian perspective where harm should be minimized.

Empirical studies have demonstrated that while the general population supports utilitarian in this hypothetical, they would personally refuse to purchase car that may sacrifice them when placed in a consumer role (Bonnefon, Shariff, and Rahwan 2016). This isn't just hypocrisy, it's a public safety backfiring. The refusal to adopt level 5 AVs results in greater human error and greater risk on roadways due to human inferiority.

Consequently, a utilitarian approach will result in greater public danger, yielding the opposite effect derived from the utilitarian ideology. Therefore, the passenger-first approach is the true utilitarian choice.By making AVs more attractive to consumers, we accelerate the transition to a system that eliminates 94% crashes casued by human error, saving many more on a larger scale.

Conclusion

In unavoidable accident scenarios between a pedestrian and autonomous vehicle, the ethical and systematic choice is to prioritize the survival of the vehicle occupants. This approach honors the manufacturer's duty of care, prevents the hazard of pedestrian recklessness, and ensure the predictability necessary for safer and more efficient transportation. It also satisfies a utilitarian viewpoint by ensuring life-saving technology reaches the masses. A safer world for everyone lies therefore in a a passenger-first approach to collision-optimization algorithms.

Bibliography

References

Alam, N., et al. 2016. "V2X communication in autonomous driving systems." *IEEE Communications Standards Magazine.*

Bonnefon, Jean-François, Azim Shariff, and Iyad Rahwan. 2016. "The social dilemma of autonomous vehicles." *Science* 352 (6293): 1573–1576.

Goodall, Noah J. 2014. "Ethical decision making during automated vehicle crashes." *Transportation Research Record* 2424 (1): 58–65.

Gurney, Jeffrey K. 2016. "Crashing into the unknown: An examination of crash-optimization algorithms through ethics and law." *Albany Law Review* 79 (1).

Hedlund, James. 2000. "Risky business: safety regulations, risk compensation, and individual behavior." *Injury Prevention* 6 (2): 82–90.

International, SAE. 2021. "SAE Levels of Driving Automation: Clarity and Refinements." SAE International. Accessed: 2025-12-18. https://www.sae.org/news/blog/sae-levels-driving-automation-clarity-refinements.

Ivanov, Dmitry. 2023. "Supply Chain Resilience: A Review and Research Agenda." Accessed: 2025-12-18, *Transportation Research Part E: Logistics and Transportation Review* 176:103201. https://doi.org/10.1016/j.tre.2023.103201. https://www.sciencedirect.com/science/article/pii/S0967070X23001816.

Lin, Patrick. 2016. "Why ethics matters for autonomous cars." In *Autonomous Driving: Technical, Legal and Social Aspects,* 69–85. Springer, Berlin, Heidelberg.

Millard-Ball, Adam. 2016. "Pedestrians, autonomous vehicles, and cities." *Annals of the American Association of Geographers* 108 (1).

National Highway Traffic Safety Administration. 2015. *Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey.* Technical report DOT HS 812 115. Traffic Safety Facts.

Nyholm, Sven, and Jeroen Smids. 2016. "The ethics of accident-algorithms for self-driving cars: an applied trolley-problem?" *Ethical Theory and Moral Practice* 19 (5): 1275–1289.