



Programming Ethics in Self-Driving Cars: Ethical Dilemma

Najam Ul Hassan

University of Maryland University College

Self-driving cars promise to revolutionize the automotive industry. Besides being productive and fuel-efficient, they would be significantly safer than the human-operated cars. However, in the rare case when they do get into an accident, they can calculate and opt for outcomes, based on their programming. The manufacturer, it seems, must decide the ethics that they should follow in such scenarios. In circumstances where they must choose between the lives of passengers and the pedestrians, some researchers have argued that the best solution is to choose for the safety of passengers over pedestrians. Such a strategy, they argue, not only makes sense as cars will have better control over the passenger, will help in faster adoption of self-driving cars. However, this line of thinking seems simplistic at best. The author suggests that the problem must be reexamined from various perspectives.

Keywords: Self driving cars, Ethical programming, Decision management

Introduction

Self-driving cars are a recent buzz that promises to revolutionize the automotive industry (Bonneton, Shariff, & Rahwan, 2016). With remarkable features, such as programmed to follow the law, calculating each driving move without getting distracted, communicating with other cars, serving the disabled, elderly and children, self-driving cars are creating a market that cannot wait for their arrival (Lyndon Bell, 2018). While self-driving cars are expected to be safer, fuel-efficient, and productive, they do give rise to some ethical questions.

Background

Safety

Self-driving cars are imminent due to the various advantages they have to offer. The most significant among them has to be safety. A self-driving car will be substantially safer than the one operated by a human. According to non-profit Eno Center of Transportation, since nine out of ten driving accidents are driver's fault, "if just 10% of all vehicles in the U.S. were self-driving, the number accidents each year would be cut by 211,000; 1,100 lives would be saved; and economic costs would be cut by \$22.7B" (Mearian, 2013). The study further states that, "if 90% of vehicles in the U.S. were self-driving, as many as 4.2 million accidents could be avoided, saving 21,700 lives and \$450 billion in related costs (Mearian, 2013). Based on these figures, Elon Musk, CEO of Tesla Motors, claims that discouraging the use of self-driving cars is equivalent to "killing people" (Jack Steward, 2016).

Increased productivity

Knowing that a car is safer when it is being self-driven as opposed being driven by one of its occupants, invites the occupant to take a break or use the time for other tasks. Self-driving cars will increase productivity as humans will be able to utilize the driving time for more important activities. People will be able to count the commute time towards office hours. Even if people opt for taking a break during driving time, studies suggest that it will result in better overall productivity (Lyndon Bell, 2018).

Besides making the drive time more productive for humans, self-driving cars will also increase productivity by eliminating the need for people to find parking. These cars will be able to drop off the occupants in front of the destination, then hunt for parking space on their own and be available to pick the passengers up when summoned (Lyndon Bell, 2018). People will be able to reclaim a lot of time they used to waste looking for parking in busy malls and populous downtown areas.

Aside from the increased productivity of drivers, imagine the role self-driving cars will play in the lives of handicapped people, senior citizens and children. Programming the cars to transport them to and from various destination will not only make them more independent, it will also increase the productivity of their caretakers. Aging parents' doctor's appointment, kid's football practice and grandpa's weekly Mah Jong game can all be programmed into a self-driving car (Lyndon Bell, 2018). Compromised driving abilities will cease to be an obstacle in people's lives.

Since driving abilities will not be needed and self-driving cars will insure that they abide by the traffic laws, the need for traffic enforcement personnel will be greatly reduced. Traffic police force can, therefore, be reassigned to more productive work, such as, fighting crime (Lyndon Bell, 2018). Taking it a step further, since self-driving cars will be communicating with each other and will know what to expect, higher speed limits can be on the horizon (Lyndon Bell, 2018). Reduced monitoring even with increased rate of transportation means higher productivity for enforcement agencies as well as citizens.

Reduced congestion

Even with the increased productivity, drive-time will be reduced due to reduced congestion. Traffic jams will be a thing of the past with self-driving cars. "One of the leading causes of traffic jams is selfish behavior among drivers" (Lyndon Bell, 2018). Bell argues that by not allowing enough space between cars due to selfishness, drivers cause the traffic jams. By programming the self-driving cars to leave more space between cars, road congestion will be greatly reduced (Lyndon Bell, 2018).

Ethical Dilemma

Although most researchers agree on many benefits of self-driving cars, however, they do bring technological challenges and ethical issues to the fore that have been the topic of much discussion. As you progress and move things on automation; bring a challenge to keep their software up to data. A little negligence can lead to cyber-attacks by like Medstar experienced (Hassan, 2018). Mearian (2013) states that compared to human-operated cars, self-driving cars are significantly safer, however, they can never be accident-free. Random incidents can occur due to which an accident may become unavoidable. It is in this rare scenario that a self-driving car will have to decide based on the factors included in its programming. A human driver, when encountered with

unavoidable accident situation, often does not have enough time to evaluate the available options and responds by randomly choosing an outcome which may or may not be the best option. A self-driving car, on the other hand, may not have enough time to apply the brakes but it will have enough time to weigh all its options and then make a ‘conscious’ decision of electing one of those options (Bonnefon et al., 2016).

Bonnefon et al. (2016) use the scenarios based on what is referred to as the “trolley problem” in ethical studies. The trolley problem, introduced by Foot (1967), presents a scenario where you are driving a trolley at 60 miles an hour when you see that five workers are on the track. You apply the brakes only to find out that they are not working. You observe a sidetrack with only one worker on that track. The ethical dilemma presented is should you continue on the original path that will kill five workers or should you take the sidetrack that will kill one worker?

Greenemeier (2016) explains that a typical scenario, a variant of what is referred to as the “trolley problem” in ethical studies, would unfold when, during a normal drive, a self-driving car finds at the last minute that a number of pedestrians have mistakenly stepped on to the road that it’s travelling on. The car, as it applies the brakes, calculates that it will not be able to stop in time and will end up killing the pedestrians if it continues down the same path or it can opt to turn to the right and hit the wall, but that will kill the passengers (Greenemeier, 2016). “Should it minimize the loss of life, even if it means sacrificing the occupants, or should it protect the occupants at all costs? Should it choose between these extremes at random?” (Bonnefon et al., 2016).

Save the Passenger

Jack Steward (2016), Bonnefon et al. (2016) and Mike Brown (2016) argue that the causes of the situation discussed in the scenario would likely not be known when the car has to make a decision and the reaction of other traffic can’t be predicted either. Therefore, they suggest that under such uncertain circumstances, the optimal way forward is that it should save the lives of the passengers as it has more control over taking them to safety. Christoph von Hugo, driver assistance systems manager at Mercedes, reasons that, “You could sacrifice the car. You could, but then the people you’ve saved initially, you don’t know what happens to them after that in situations that are often very complex” (Mike Brown, 2016). He further elaborates, “If you know you can save at least one person, at least save that one. Save the one in the car”. Bonnefon et al. (2016) also point out that while the pedestrians may or may not be at fault, neither the car, nor the passengers have made any error for which they should not be penalized. They conclude that the car should put up the best effort to save the pedestrians while keeping the passengers safe.

Ackerman (2016) Discusses that a survey found that most people are of the opinion that the option that saves most lives, a utilitarian approach, should be preferred. However, when the same group was asked if they would buy a car that is programmed to opt for greater good or the one that protects the passengers at all cost, they preferred the later (Ackerman, 2016). Ackerman (2016) Concludes that people want cars programmed for greater good as long as they are not the passengers. Therefore, cars not programmed to protect the passengers, will result in delay in the adoption of self-driving cars. Therefore, the contention is that since self-driving cars are safer to the point that if 9 out of 10 vehicles were self-driving, over 21,000 lives will be saved in the US alone, any delay in their adoption means condoning accident fatalities (Mearian, 2013). Elon Musk, CEO of Tesla Motors, points out that causing delays in the adoption of self-driving cars is equivalent to “killing people” (Jack Steward, 2016).

Questions and Objections

Birnbacher and Birnbacher (2017) Stated that with all the variables of the situation and immense complexity of the tasks, even if the “right” reaction can be programmed in to the self-driving cars, the following questions still remain:

- Who should have the authority to decide whether the preference rules and learning skills programmed into the system are acceptable?
- How much differentiation in ethical programming should be allowed to different stakeholders (producers, users, societies)?
- Who is responsible in case of damage? (Birnbacher & Birnbacher, 2017).

Scheutz (2017) Argue that self-driving cars’ inability to coordinate with humans through other methods, such as hand gestures or eye contacts, make them a hazard in the traffic system. He suggests that features should be developed that would make the self-driving cars more ‘humanlike’ and still offer the benefits such as reduction of accidents.

Reassessment

JafariNaimi (2018) Implores that we must step back and reassess our assumptions regarding self-driving cars. He argues that the justifications provided for the implementation of ‘algorithmic morality’ are due to the incorrect view of the trolley problem among other aspects that is muddying our ability to see the inbuilt limitations of cars that use such algorithms.

JafariNaimi (2018) contents that Bonnefon et al. (2016) use experimental ethics when presenting the scenarios to the subjects which resulted in acceptability of moral algorithms in the study. On the other hand, in the trolley problem, he argues, the scenarios are ‘understood to be far-fetched or even hyperboles’ and are understood not to be taken literally. Furthermore, JafariNaimi (2018) explains that the trolley experiment uses quandary ethics in which ‘the parameters are predefined and fixed and the choices are clear,’ however, in with the scenarios of self-driving cars parameters are not defined or fixed and the choices are not as clear. He argues that the situations would uncertain and organic while the subject are being asked questions which are based on situated knowledge and relational. Also, he argues that the broad and long ranging effects of the decisions are not being considered. He concludes that ‘succumbing to algorithmic morality in the name of increased safety would be a grand failure of both our ethical and technical imagination.

Conclusion

Rapid innovation in self-driving cars presents some tough ethical questions. A cursory look at the issues suggest that the assertion of Bonnefon et al. (2016) where they contend that the passenger should be saved when it comes down to it, is worth pursuing. However, a deeper look begs that the need of the hour is to take a step back and rethink the problem all over again where ‘safety of pedestrians and bikers, the livability of cities, and environmental sustainability—all center stage’ (JafariNaimi, 2018). The rate of innovation in the automotive industry indicates that self-driving cars will roaming around the roadways in the very near future. With heightened safety, increased productivity and diminished congestion among other scores of benefits, their introduction should not only be allowed, but encouraged and facilitated. The debate over the ethical decision

should be addressed by developing “cars don’t drive into situations where that could happen and [will] drive away from potential situations where those decisions have to be made” (Mike Brown, 2016). These cars will not only positively impact the society, but will also deliver tremendous boost to the development of self-driving busses, trucks and rails.

References

- Ackerman, E. (2016). People Want Driverless Cars with Utilitarian Ethics, Unless They’re a Passenger.
- Birnbacher, D., & Birnbacher, W. (2017). Fully Autonomous Driving: Where Technology and Ethics Meet. *IEEE Intelligent Systems*, 32(5), 3-4.
- Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, 352(6293), 1573-1576.
- Greenemeier, L. (2016). Driverless cars will face moral dilemmas. *The Scientific American*. <http://www.scientificamerican.com/article/driverless-cars-will-face-moral-dilemmas>.
- Hassan, N. U. (2018). Ransomware Attack on Medstar: Ethical Position statement. *SEISENSE Journal of Management*, 1(4), 33-35.
- Jack Steward. (2016). ELON MUSK SAYS EVERY NEW TESLA CAN DRIVE ITSELF.
- JafariNaimi, N. (2018). Our bodies in the trolley’s path, or why self-driving cars must* not* be programmed to kill. *Science, Technology, & Human Values*, 43(2), 302-323.
- Lyndon Bell. (2018). 10 Benefits of Self-Driving Cars, *autobytel*. Retrieved from <https://www.autobytel.com/car-ownership/advice/10-benefits-of-self-driving-cars-121032/#>
- Mearian, L. (2013). Self-driving cars could save more than 21, 700 lives, \$450 B a year. *Computerworld*.
- Mike Brown. (2016). Mercedes-Benz's self-driving car will kill pedestrians over drivers. *Inverse*. Retrieved from <https://www.inverse.com/article/22204-mercedes-benz-self-driving-cars-ai-ethics>
- Scheutz, M. (2017). The case for explicit ethical agents. *AI Magazine*, 38(4), 57-64.