



AK2003 TECHNOLOGY AND ETHICS

Who is morally responsible for fully autonomous cars?

Martin Barksten
barksten@kth.se

20th May 2015

1 Introduction

In 2012 one of Google’s cofounders, Sergey Brin, in an interview says: “You’ll ride in robot cars within 5 years”. That is in just one year from now and while it might seem as if that is a bit too optimistic, robot cars, or self-driving cars as they are more often called, are not far from realistic. Google’s self-driving cars have together driven more than 1.5 million miles.

The quick development of self-driving cars is driven by the many advantages that might be gained from getting humans off the road. In their annual report on traffic safety WHO points out that 1.2 million people die each year in traffic accidents [24]. And this is just looking at the traffic accidents that are fatal, low- and middle-income countries lose approximately 3% of their GDP as a result of traffic accidents. Studies also show that a majority of these accidents seem to be the cause “human error” [relative_trfoudas], something an self-driving car would not be prone to do.

Further motivation to the gains from self-driving cars are described by Howard in [13], where he mentions the ability for disabled to use cars and the ability to revitalize failing cities.

But self-driving cars suffers from the problem of responsibility: who is to be held responsible when a self-driving car crashes? Or makes a seemingly odd decision? Some argue that the temporary solution is to allow the driver to intervene, temporarily taking control of the car to handle the situation. However, as Goodall points out this has several problems resulting in that the probability is high the driver will unable or unwilling to take control — making the car forced to take action anyway [9].

It therefore seems that the cars must be capable of taking all decisions themselves and that the problem of responsibility remain. This is a problem that needs solving, because as pointed out by Merchant et. al.: “Cars crash. So too will autonomous vehicles [...]”.

It is not hard to imagine a scenario where a car will crash and in deciding what to do must take an ethical decision. As an example let’s use a variation of the commonly used Trolley problem [21] outlined by Goodall: The car is driving on a small bridge and in the opposite lane is a bus, which suddenly turns towards the car’s lane. The car then has three possible actions:

1. Drive off the bridge, guaranteeing a serious accident for the car;
2. Drive straight on towards a head-on collision with the bus, causing a less serious accident for both the car and bus;
3. Or attempt to drive past the bus with the possibility of a avoiding a crash, but a probability of much more serious injuries for the passengers of both the bus and the car.

It is not hard to imagine that a similar scenario will occur reality, forcing the car to make a decision with moral implications — does it sacrifice the driver in favor of the probably more passengers of the bus?

2 Essay question

This essay will attempt to find an answer to the following question: *Who is morally responsible for an autonomous car?* In order to do some distinctions have to be done.

First of all, when an entity is referred to as being morally responsible for an autonomous car it means that the the entity is responsible for the morally significant actions performed by the car. This definition follows the one defined by Eshleman in [eshleman_2014_moral_mra] who further illustrate what that entails by writing the following: “Thus, to be morally responsible for something, say an action, is to be worthy of a particular kind of reaction — praise, blame, or something akin to these — for having performed it.” While this definition suffers from some problems, they are not of any major concern for this discussion.

Furthermore, when discussing the term autonomous car is used deliberately in order to describe a car that is fully autonomous, requiring nothing from any passenger in the car. The actions of the car are therefore decided by the car alone without outside input being possible.

3 Discussion

In order to discuss who is morally responsible for an autonomous car I will go through all actors that might be potentially considered responsible — discussing what the consequences of that actor being responsible would mean in terms of duties etc.

The actors that might be considered responsible for an autonomous car are:

1. The passenger or passengers of the vehicle;
2. The autonomous vehicle itself;
3. The company responsible for programming, constructing and selling the car;
4. And the government that set the legal framework allowing the vehicles.

For most of these actors I will discuss them taking moral responsibility in one of the two following scenarios:

1. A scenario where the vehicle has the potential to cause an accident, but the possibility of avoiding it if acting correctly;
2. A scenario where the vehicle must several different wrong possible actions — all of which will yield an accident of some sort.

The second example might be considered an example of the Trolley problem, a scenario where an actor must choose between several outcomes all of which are bad but in different ways [21].

This section will therefore start with a discussion for each of the four actors, following this is a discussion taking all actors into account and the previous discussion.

3.1 The passengers

- the passenger takes one decision: choosing to use an autonomous vehicle, all consequences derive from that action.
- in the first scenario it seems not to be the fault of the passenger if the car crash, indeed the scenario could be said to occur constantly as the vehicle always has the possibility of doing a dumb action. So if there is an action that is correct, the passenger's decision to use an autonomous vehicle does not seem to be at fault.
- in the second scenario the passenger could be held more responsible as the car does the best of the situation and indeed it could be argued that the accident would not have happened had the passenger not chosen to use the autonomous vehicle.
- however, it still seems a bit odd that the passenger is held responsible for the actions of a vehicle if the passenger has no way to affect them.

3.2 The autonomous vehicle

- the vehicle could be considered to be an actor taking part in the scenario as it is in practice the one taking the actions
- to say that the vehicle is morally responsible for an action is to by extension consider the vehicle capable of morality. Without going into a discussion of morality and robots, it does not seem as if the robots we will see in 5 years are sufficiently autonomous as to be capable of morality. And if indeed we say they are, we got more complex moral problems in using them to drive us around

3.3 The company

- The company will be treated as one single actor without regard for individuals in the company such as the developer who wrote the bug or similar, this as the company can be considered responsible for the final product and so by extension the actions of its employees.
- The company producing the car could be seen as to be the one most clearly at blame when an autonomous vehicle crash, as they are the ones that have programmed and built the car.
- However, it is important to consider the fact that we know already now that these traffic accidents will occur. So in allowing the vehicles we have to some extent also accepted these traffic accidents.

- in the first scenario it seems that the responsibility can be put on the company for having created a car that in a scenario where the car could reasonably have avoided the error didn't do so. The reason it didn't must be due to a fault in how it is built
- in the second scenario it does on the other hand seem odd that the company is to blame for an accident that could not have been avoided. If the car did what could be considered to be the best possible of it, then why are the company responsible?

3.4 The government

- the government here refers to some authority that allowed the autonomous vehicle on the road and set the legal framework for what might be considered a legal autonomous vehicle.
- in scenario 1 it seems that the government is not to be held responsible for the cars action, while they allowed it on the road they probably did so under the assumption that it would be able to avoid traffic accidents that could have been reasonably avoided.
- in scenario 2 it does seem as if the government is to be held at least a little responsible, they knew that the vehicles would cause accidents and still allowed them — by extension the government allowed the accident to happen.

4 Conclusion

conclusion

References

- [1] *Average Annual Miles per Driver by Age Group*. URL: <http://www.fhwa.dot.gov/ohim/onh00/bar8.htm> (visited on 05/10/2016).
- [2] Oliver Bendel. “Considerations about the relationship between animal and machine ethics”. In: *AI & SOCIETY* 31.1 (Dec. 11, 2013), pp. 103–108. ISSN: 0951-5666, 1435-5655. DOI: 10.1007/s00146-013-0526-3. URL: <http://link.springer.com/focus.lib.kth.se/article/10.1007/s00146-013-0526-3> (visited on 05/09/2016).
- [3] Jean-François Bonnefon, Azim Shariff, and Iyad Rahwan. “Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars?” In: *arXiv preprint arXiv:1510.03346* (2015). URL: <http://arxiv.org/abs/1510.03346> (visited on 05/09/2016).

- [4] Don Brutzman, Robert McGhee, and Duane Davis. “An implemented universal mission controller with run time ethics checking for autonomous unmanned vehicles—A UAV example”. In: *Autonomous Underwater Vehicles (AUV), 2012 IEEE/OES*. IEEE, 2012, pp. 1–8. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6380744 (visited on 05/09/2016).
- [5] Boer Deng. “Machine ethics: The robot’s dilemma”. In: *Nature* 523.7558 (July 1, 2015), pp. 24–26. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/523024a. URL: <http://www.nature.com/doifinder/10.1038/523024a> (visited on 05/09/2016).
- [6] Andrew Eshleman. “Moral Responsibility”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2014. 2014. URL: <http://plato.stanford.edu/archives/sum2014/entries/moral-responsibility/> (visited on 05/11/2016).
- [7] Andrew Eshleman. “Moral responsibility”. In: (2014). URL: http://pilotscholars.up.edu/cgi/viewcontent.cgi?article=1001&context=phl_facpubs (visited on 05/09/2016).
- [8] Noah Goodall. “Ethical Decision Making During Automated Vehicle Crashes”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2424 (Dec. 2014), pp. 58–65. ISSN: 0361-1981. DOI: 10.3141/2424-07. URL: <http://trrjournalonline.trb.org/doi/10.3141/2424-07> (visited on 05/10/2016).
- [9] Noah J. Goodall. “Machine Ethics and Automated Vehicles”. In: *Road Vehicle Automation*. Ed. by Gereon Meyer and Sven Beiker. Cham: Springer International Publishing, 2014, pp. 93–102. ISBN: 978-3-319-05989-1 978-3-319-05990-7. URL: http://link.springer.com/10.1007/978-3-319-05990-7_9 (visited on 05/10/2016).
- [10] *Google Self-Driving Car Project*. Google Self-Driving Car Project. URL: <http://www.google.com/selfdrivingcar> (visited on 05/10/2016).
- [11] *Google’s Sergey Brin: You’ll ride in robot cars within 5 years - CNET*. URL: <http://www.cnet.com/news/googles-sergey-brin-youll-ride-in-robot-cars-within-5-years/> (visited on 05/10/2016).
- [12] Alexander Hevelke and Julian Nida-Rümelin. “Responsibility for Crashes of Autonomous Vehicles: An Ethical Analysis”. In: *Science and Engineering Ethics* 21.3 (June 11, 2014), pp. 619–630. ISSN: 1353-3452, 1471-5546. DOI: 10.1007/s11948-014-9565-5. URL: <http://link.springer.com.focus.lib.kth.se/article/10.1007/s11948-014-9565-5> (visited on 05/09/2016).
- [13] Don Howard. *Science Matters » Blog Archive » Robots on the Road: The Moral Imperative of the Driverless Car*. URL: <http://donhoward-blog.nd.edu/2013/11/07/robots-on-the-road-the-moral-imperative-of-the-driverless-car/#.VzHclBV96bX> (visited on 05/10/2016).

- [14] Aaron M. Johnson and Sidney Axinn. “THE MORALITY OF AUTONOMOUS ROBOTS”. In: *Journal of Military Ethics* 12.2 (July 2013), pp. 129–141. ISSN: 1502-7570, 1502-7589. DOI: 10.1080/15027570.2013.818399. URL: <http://www.tandfonline.com/doi/abs/10.1080/15027570.2013.818399> (visited on 05/09/2016).
- [15] *Licensed Drivers - Our Nation's Highways - 2000*. URL: <http://www.fhwa.dot.gov/ohim/onh00/onh2p4.htm> (visited on 05/10/2016).
- [16] Gary E. Marchant and Rachel A. Lindor. “Coming Collision between Autonomous Vehicles and the Liability System, The”. In: *Santa Clara L. Rev.* 52 (2012), p. 1321. URL: http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/saclr52§ion=41 (visited on 05/10/2016).
- [17] Duncan Purves, Ryan Jenkins, and Bradley J. Strawser. “Autonomous Machines, Moral Judgment, and Acting for the Right Reasons”. In: *Ethical Theory and Moral Practice* 18.4 (Jan. 30, 2015), pp. 851–872. ISSN: 1386-2820, 1572-8447. DOI: 10.1007/s10677-015-9563-y. URL: <http://link.springer.com.focus.lib.kth.se/article/10.1007/s10677-015-9563-y> (visited on 05/09/2016).
- [18] Brandon Schoettle and Michael Sivak. “A Preliminary Analysis of Real-World Crashes Involving Self-Driving Vehicles”. In: (2015).
- [19] John P. Sullins. “RoboWarfare: can robots be more ethical than humans on the battlefield?” In: *Ethics and Information Technology* 12.3 (July 7, 2010), pp. 263–275. ISSN: 1388-1957, 1572-8439. DOI: 10.1007/s10676-010-9241-7. URL: <http://link.springer.com.focus.lib.kth.se/article/10.1007/s10676-010-9241-7> (visited on 05/09/2016).
- [20] *The Ethics of Saving Lives With Autonomous Cars Is Far Murkier Than You Think* / WIRED. URL: <http://www.wired.com/2013/07/the-surprising-ethics-of-robot-cars/> (visited on 05/10/2016).
- [21] *Trolley problem*. In: *Wikipedia, the free encyclopedia*. Page Version ID: 719408859. May 9, 2016. URL: https://en.wikipedia.org/w/index.php?title=Trolley_problem&oldid=719408859 (visited on 05/10/2016).
- [22] Nicole A. Vincent, Ibo van de Poel, and Jeroen van den Hoven, eds. *Moral Responsibility*. Vol. 27. Library of Ethics and Applied Philosophy. Dordrecht: Springer Netherlands, 2011. ISBN: 978-94-007-1877-7 978-94-007-1878-4. URL: <http://link.springer.com/10.1007/978-94-007-1878-4> (visited on 05/09/2016).
- [23] M. Mitchell Waldrop. “Autonomous vehicles: No drivers required”. In: *Nature* 518.7537 (Feb. 4, 2015), pp. 20–23. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/518020a. URL: <http://www.nature.com/doifinder/10.1038/518020a> (visited on 05/10/2016).
- [24] World Health Organization. *Global status report on road safety 2015: supporting a decade of action*. Geneva, Switzerland: WHO, 2015. ISBN: 978-92-4-156506-6.

- [25] M. Zimmer. “Surveillance, Privacy and the Ethics of Vehicle Safety Communication Technologies”. In: *Ethics and Information Technology* 7.4 (), pp. 201–210. ISSN: 1388-1957, 1572-8439. DOI: 10.1007/s10676-006-0016-0. URL: [http://link.springer.com/focus.lib.kth.se/article/10.1007/s10676-006-0016-0](http://link.springer.com/focus/lib/kth.se/article/10.1007/s10676-006-0016-0) (visited on 05/09/2016).