

Theory for Designing Artifacts

Wed. 16:30-18:00

Mr. Sawaragi

Graduate School of Engineering

Mechanical Engineering

April, 2017

Student Number: 1030-29-9698

Mei Jie

The Ethics of Autonomous Vehicles

As stated in the material, an autonomous vehicle is put into a dilemma, where it should make a decision between hitting a bike with increasing the rate of causing injury to its driver, and directly passing through without any reaction so that the animal or baby in front may die because of it. This kind of problems can be seen as the extension of an ethical issue called Trolley problem but replacing the human role with a robotic agent. As a lead to the discussion of justice, there is somehow no correct answer in the Trolley problem. So what we want should not a detailed answer but a guidance in designing an autonomous vehicle, especially the one owning ethics.

Before introducing the research of embedding ethics in the autonomous vehicles, I would like to firstly discuss why this problem comes out. It seems a little mean to ask a computer or robot to answer an unsolvable problem, but this exactly implies our worries in autonomous vehicle. As a newborn, autonomous vehicle came into the public view just a few years ago, some technology it used such as deep learning still lacked of interpretability and reliability. Besides, the corresponding facilities such as law and even the moral standards have not been developed yet.

Furthermore, all our worries are based on an assumption of the "driver first" principle, which implies that the autonomous vehicle is designed to select the action of minimizing the injuries to its inner driver only. Moreover, it is also assumed that all actions the autonomous vehicle chooses are the result of calculation. Noted that in a same situation, the action from our human is based on the intrinsic reaction not the calculation. And a reaction is allowed to be imperfect but not the calculation.

So the topic of not just simply calculating the optimum within the autonomous vehicle becomes a clue in whether the ethics can be embedded while designing. Based on my understanding, some researcher focus on the definition of ethics, such as how will human choose after fully consideration and how our robot will choose based on current algorithm. In addition, some researchers prefer to predesign some rules into the algorithm so that the autonomous vehicle can take some ethical factors into account before making its decision.

And personally, as a technical optimist, I holds a strong believe in be able to prevent the autonomous vehicle from this kind of ethics dilemma through the regressive development in technology, which means our attention should to put on minimizing the risk of its happen instead of solving it.

Human vehicle collaboration

With different automated capabilities, the driving automation can be defined as different six levels as followed: no automation, driver assistance, partial automation, conditional automation, high automation and full automation in accord with the Society of Automotive Engineers. Noted that all the provided materials belong to the third level, in which the automated system intervenes into the driving activities only in an urgent situation to avoid the accidents. And the material also shows that even though with this standard, there is still a large space for different companies to select their own strategies in when and how to take this intervention.

Here is a brief summary of three different systems from Toyota, Mercedes-Benz, Volvo and the comparison in their collision avoidance strategies.

Company	Characteristic
Toyota	Lateral pedestrian avoidance
Mercedes-Benz	Vertical pre-break system, active lane keeping system, active blind spot assist, attention assist
Volvo	Collision avoidance for slow moving traffic and stationary object

Table. 1 Different characteristics of the system from above companies

	Toyota	Mercedes-Benz	Volvo
Target object	Lateral pedestrian	Vertical vehicle	Vertical vehicle
Whether alarm	Yes	Yes	No
How to alarm	Warning light	Distance display and warning light	None
How to avoid	Break and steer	Pre-break and take action	Break suddenly
others	Collaboratively steer with the driver	Calculate pressure in need and wait for driver's react	Apply to slow moving traffic

Table. 2 Comparison in collision avoidance strategy

And according to the mechanism of human-automation collaboration presented by the material, the difference between above strategies can be described as when and how the human is required to take back the driving right before a collision. As the extreme, Volvo trends to not inform its driver and break automatically and suddenly. And the nuance between the strategies from Toyota and Benz is located at their methods and goals as you can find in table 2 above.

Furthermore, the question of how to collaborate with driver should be considered. But before that, I would like to talk about the different opinions in whether collaboration is necessary for autonomous vehicle. Without doubt, the goal should be set at the fully automation, but as for how to reach it, there exists different opinions. For some software-based companies like Waymo or Drive.ai, it is believed to skip the fourth and fifth level and directly implement the fully automation, saying the combination of the uncertainty in both human and driving environment seems more difficult than the environment only. On the contrary, those companies from automotive industry trend to be more conservative and progressive. By the way, there are also some companies like Toyota bet hedging in both sides. Noted that it is an issue about not only the technology but also the traffic condition and many others. Now back to the former topic, how to collaborate with the driver. No one will deny that different strategies should be taken in accord with different situation, but in my opinion some standards are all the same. For example, how the driver will react afterward and how long will it take for this reaction should be considered before transferring the driving right back to this driver. Limited by the context, there will not be further discussion here.