Meaningful Human Control over Automated Driving Systems

November 2019

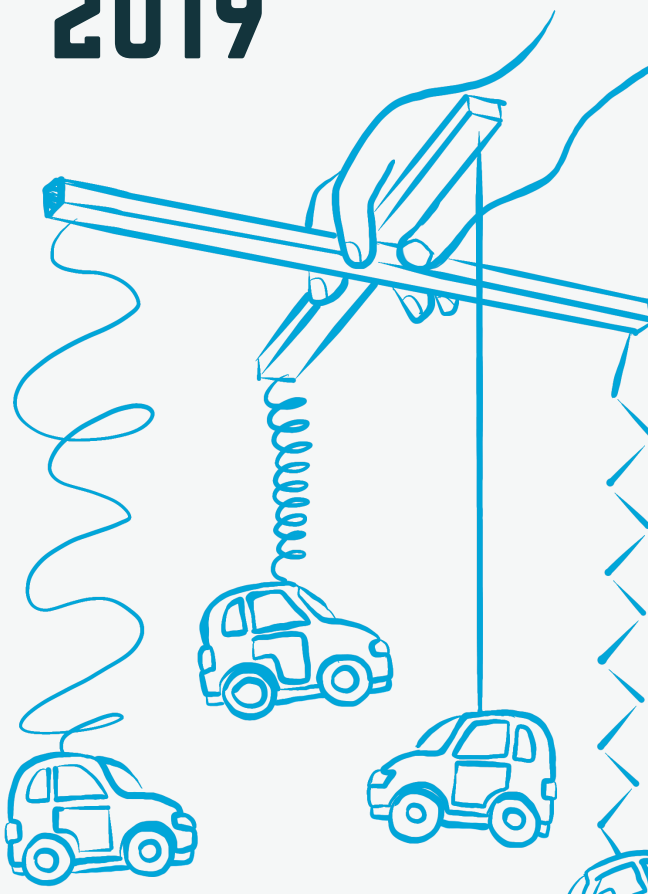
THIS ISSUE'S TOPICS

THE EDITORIAL
What We Talk About When We Talk About Responsibility for Self-Driving Cars
by *Filippo Santoni de Sio*

THE OPINION
Liability for Automated Vehicles
by *Marly de Blaeij & Jochem van Stiphout,*
Dutch Association of Insurers (Verbond van Verzekeraars)

NEWS FROM OUR PROJECT
Our updates on research and things we do in our daily life of hard (?) working researchers

PUBLICATIONS AND DISSEMINATION
A short list of our most recent publications, and some of our most noteworthy activities



COLOPHON

Meaningful Human Control newsletter
TU Delft

Editor-in-chief: Giulio Mecacci, g.mecacci@tudelft.nl
Editors: Daniël Heikoop, Simeon Calvert

THE EDITORIAL

WHAT WE TALK ABOUT WHEN WE TALK ABOUT RESPONSIBILITY FOR SELF-DRIVING CARS

by *Filippo Santoni de Sio*

Whenever a self-driving car is involved in a crash, experts and lay persons start wondering who is responsible for that. And attributions of responsibility have been one of the greatest concerns of car manufacturers, policy makers, lawyers, insurers, philosophers, engineers since the start of the process of transition to automated driving. As a philosopher who has studied, nay, has been obsessed with issues of responsibility since the (far away) time of my PhD in philosophy in Torino, I cannot help but being thrilled about this sudden public interest for my favourite topic. At the same time, as someone who has always sought conceptual clarity about responsibility in the first place, I also am sometimes puzzled about current debates about responsibility for self-driving cars. Precisely, what do we talk about when we talk about responsibility for self-driving cars?

In the aftermath of an accident, many people are usually looking at legal liability for an untoward outcome. Who is responsible (legally liable) for the death of the driver of the Tesla which tragically crashed into a lorry in 2016? Well, to this kind of questions there may be relatively straightforward answers; in fact, the investigation concluded that the driver himself was liable, as he neglected his duty to keep his hands on the wheel and his eyes on the road, in contrast to the explicit, written, recommendation of the manufacturer and to the repeated beeping of the car during the driving operation. Case closed. Or is it? Maybe what we really want to know is who is morally culpable for the accident. Here things get more complicated and looking at the law may not be enough. Certainly, the driver was reckless, and has a share of culpability, but was he sufficiently trained to perform his new tasks as controller of a partially automated vehicle? If not, aren't also governmental agencies to blame for that? And how about the car manufacturer who, according to a possibly abused but certainly effective metaphor, knowingly put a loaded gun into the hand of an incompetent user?

Moreover, moral and legal culpability are not the only things that matter. Another form of responsibility is moral accountability: the capacity of human persons to reflect upon and give reasons for their behaviour.

Even without anything bad happening, we want ourselves and others to be able to explain some of their actions: why didn't you call me? Why did you move to another city? Why did I react in this way? We cannot (always) just shrug and say: I don't know, it just happened. Trying to answer "why-questions" is part of what it means to be (rational) persons in society. While delegating sensitive tasks to automated systems, we should also see to it that this delegation does not affect too much the human capacity to understand and explain their choices and behaviour, for instance by making the operation of the system too complex, cumbersome, or opaque. Moreover, whenever a complex system is set up in such a way that it can directly influence the life (and death) of people we also want some mechanisms of public accountability to be in place. We want formal mechanisms to hold designated agents to report and explain the behaviour of the system to designated fora (and we want these fora to have the power to pressure them to do so). This is a basic principle of political action in liberal-democratic society, one that we should also apply to the developers and regulators of self-driving cars (cars operate in the public space).

Finally, responsibility is not just a backward-looking game, it is also a forward-looking one. Sometimes, in asking the responsibility question we are not (only) interested in assigning liability or blame or in identifying the subject who should explain a past (untoward) event; we are rather asking who should commit themselves and see to it that that untoward events do not happen in the first place. This is sometimes called active responsibility. Who is actively responsible for the safe introduction of self-driving cars in society? In different ways, and to different extents, all of the agents involved in their design, development, regulation, use are. In order to make them aware of this responsibility and able and motivated to discharge it, I make two proposals. First, that we as a society maintain awareness of the different senses of responsibility, and we do not remain prisoners of a narrow concept of responsibility as legal liability or moral culpability; accountability and active responsibility matter too. Second, that we better clarify what kind of (meaningful) control different agents involved in the design development regulation and use should have in order to be able to discharge their different responsibilities for the behaviour of self-driving cars.

THE OPINION

LIABILITY FOR AUTOMATED VEHICLES

by *Marly de Blaeij & Jochem van Stiphout*
Dutch Association of Insurers
Corresponding author: M.de.Blaeij@verzekeraars.nl

Automation of vehicles is one of the major developments that is shaping the future of mobility. Although there are different views on what the eventual impact of automated vehicles will be, most people agree that it will change how we organize and insure our mobility in the future.

Trends in mobility
Over the last couple of years, the fleet of passenger cars in the Netherlands has remained at a stable 8.5 million vehicles. Despite the stable number of cars, the mobility landscape has been continuously been changing. Trends such as private leasing and carsharing platforms show the first hesitant signs of a shift from car ownership towards car usage. With the help of technology, a growing number of customers are starting to make use of carsharing platforms. Another trend that can be observed is an increase in the demand for multi modal transportation. Consumers wants to be able to easily choose between car, bicycle and public transport depending on their journey. Therefore, it is no coincidence that large car manufacturers are changing their business models from producing automobiles to enabling mobility in a broader sense.

Furthermore, there is the prospect of further technological development both as an enabler for the aforementioned trends as well as a major shaping factor on its own. Technological development makes it possible for vehicles to operate more autonomously than ever before. The application of smart sensors and safety instruments in cars can have a positive contribution towards road safety and significantly reduce the number of accidents. This may cause some to believe that further development in autonomous driving will in the long run remove the need for car insurance. In the next section, we will explain why this might not necessarily be the case.

An insurance perspective on automated driving
Even if we would be able to completely rule out human error trough developments in automated driving, we are still not able to prevent damage to vehicles by, for example, extreme weather or theft.

In reality, however, accidents will still occur, and the increased use of technology in vehicles also means that the value per claim is on the rise. For example, a car window is no longer an ordinary piece of glass but is now filled with technology and sensors. This is also true for other car parts making them increasingly more expensive to replace or repair.

There is also the question of liability. Nowadays, liability is often determined by the actions of the driver. If a driver skip a red light and hits another car, it is stated in the law that this driver is the liable party. It becomes more complicated when an accident occurs, for example, from a blown-out tyre: in this case the owner of the car will bare liability, since such a blowout could have been prevented if the car was well maintained. Whilst in theory the blowout could also have been caused by production fault of tyre manufacturer, in practice action is rarely taken towards a manufacturer since it is almost impossible to gather the evidence to prove manufacturer's liability.

"accidents will still occur, and the increased use of technology in vehicles also means that the value per claim is on the rise"

In a situation where cars are becoming more and more autonomous and the impact of human behavior decreases, this raises the question at what point it will no longer be the actions of the driver but the actions of the technology in the vehicle that determines liability. In other words, where does your responsibility as a driver stops and where does the responsibility of the manufacturer begins?

The Society of Automobile Engineers (SAE) developed a classification of five stages of automation in vehicles. For every stage they describe the level of automation as well as the level of attention still required by the driver. According to this classification, a level 3 automated vehicle would allow the driver to leave the handling of the vehicle and the monitoring of the environment up to the vehicle, acting only as a back-up system in an emergency or unexpected situation. However, scientists have already warned that in practice drivers will be significantly less alert in these situations. This is due to drivers being more easily distracted when the vehicle takes over key driving functions. Unfortunately, yet significantly, recent accidents during the testing of level 3 automated vehicles showed these concerns were justified.

These examples show the complexity of determining liability in automated driving and open the door to possibly lengthy discussions and procedures of determining liability in case of an accident involving an automated car. In addition, consumers that want to reclaim their damages based on product liability will be faced with a significant imbalance in power versus large car manufacturers. These global companies have the means to run long and costly legal cases, reducing the chance for consumers to prove they are in the right and potentially eroding the level of protection that consumers enjoy in the current liability framework.

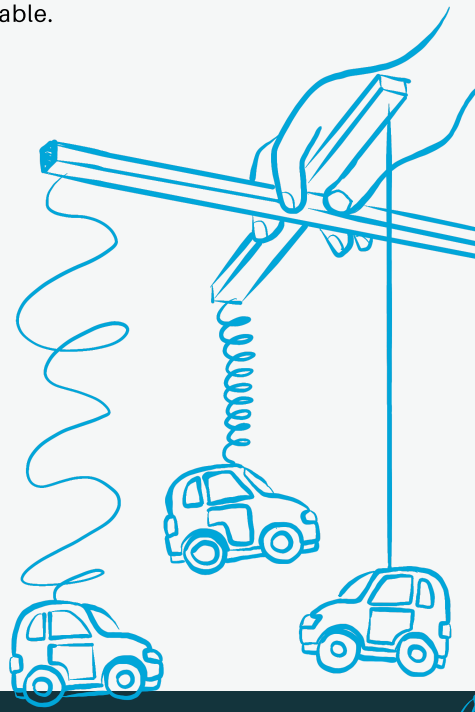
"recent accidents during the testing of level 3 automated vehicles showed these concerns were justified"

How can we solve these issues?
There are three elements that are crucial in solving these challenges. First of all, to extend the use of event data recorders which are fitted in most new cars. These operate similarly to a black box in an airplane and record data such as speed, steering, braking and acceleration. This data should be available to law enforcement agencies in order to help them quickly identify the circumstances under which the accident took place and to help to determine whether the actions of the driver or the actions of the vehicle led to the accident.

Currently, privacy provisions prevent the data from these recorders to be traceable to individual vehicles and their users, making them unsuitable for the purpose of reconstructing accidents and possibly determining liability. Providing the police unrestricted access to this data would significantly reduce the time and costs related to determining liability after an accident. Secondly, we need to establish an independent and international framework to determine liability for different levels of automation in vehicles. This will provide more clarity for manufacturers, drivers and insurance companies on where their liability begins and ends in case of an accident involving an automated vehicle.

"we need to establish an independent and international framework to determine liability for different levels of automation in vehicles"

Finally, we need to change how motor liability insurance works, by going from a third party to a first party insurance. This means that after an accident a policy holder can turn to their own insurer to claim damages or bodily injury instead of the insurer of the liable party. Which also means that consumers do not have to wait until the liability is fully determined before they can claim their damages. Such an approach would also allow for compensation of drivers in cases where there is reduced or shared liability between them and the car manufacturer. For consumers, first party insurance also means that their own insurance company, which they know, will be handling their claims in a fast and transparent way. The Dutch Association of Insurers therefore believes that this is the way to go to keep the future of mobility insurable.



NEWS FROM OUR TEAM

Dr. Simeon Calvert, traffic engineer

Recent months have involved two main ongoing efforts on the engineering side of the project. The first concerns a joint effort with Kostas Ampountolas from the University of Glasgow on developing a control design for a traffic system based on MHC. This is part of the challenging, but very necessary and relevant step of operationalisation of MHC. A large amount of focus in the coming half year of the project will be aimed at further operationalising the concept for traffic and transportation such that MHC does not remain an abstract concept, but rather the benefits and potential applications will become evident by the end of the project. Parallel to this effort, we continue to refine the theory of MHC together with the philosophers in the project, and applying this to the contact of urban traffic with its many complex interactions. The second main effort lies in model development to allow explicit cases to be carried out. Based on the previously constructed simulation framework, we are further investigating ways to perform case studies and demonstrate the power of MHC and feasibility of the models. We are also investigating further options to perform a final case study, besides those already initiated. We very much look forward to an impressive conclusion of the project through this.

Dr. Daniël Heikoop, behavioral scientist

The simulator experiment is now finally close to getting started. Recruitment has taken place (twice), students are assigned (two more and one replacement), and the selection procedure has been completely worked out. Close to 100 participants have already applied, and more continue to apply in the coming weeks. Next to that, the focus group paper is taking shape, and is already in a shape worthy of submitting it to the FISTS conference, to be held here at Delft, in June 2020. Two more papers are being worked on to submit to the FISTS conference: one that describes a literature review on HMIs in automated vehicles, and one that writes about an online questionnaire that asks 120 participants about their opinion regarding their trust in and acceptance of automated driving systems. Two other papers will be submitted to the ICTTP conference in Goteborg: one on how the recruitment and selection procedure of the participants based on the Big Five Inventory has taken place, and one on the data we retrieved from the recruitment phase, covering among others demographics, and experience with automated driving systems. Last but not least, the literature review on HMIs in automated vehicles is currently being expanded, covering over 20 different types of details, and we now also have two more students working on an on-road experiment, using our instrumented Toyota Prius: one who will be looking at the general driving behaviour in an automated vehicle based on their personality (again, the Big Five Inventory), and one who will be looking at an empirically assessable measure for trackability, one of the two cornerstone conditions of MHC.

Dr. Giulio Mecacci, philosopher of mind and technology

The philosophy side of this project have been hectically working on further refining the conceptual basis of meaningful human control, to allow clear access to operationalization. We have been continuously assisting our fellow engineers in the extremely delicate matter of translating normative values into somewhat quantifiable dimensions. As you can nicely read in our month's editorial, a notion like that of responsibility, far from being objectively observable and measurable, has a very complex, abstract and socially determined nature. Nonetheless, it should be sought and represented by the artificial systems we design, that operate within, and for, a society made by values, norms and ultimately people.
We have been succesfully collaborating with Kostas Ampountolas in discussing how control should be maintained and distributed in urban, mixed traffic scenarios, where traditional and automated cars have to coordinate with vulerable road users. A paper is almost complete where we explore how MHC can provide a clearer understanding of the relations between the multiple responsible agents in urban traffic.
Another couple of papers, more philosophical, are being currently on the work. There, we spell out different senses of responsibility and of being responsible agents. We also suggest how MHC principles can help designing systems that maximally preserve and operationalize all of them.

RECENT PUBLICATIONS AND DISSEMINATION

Journal articles (published)

- S. C. Calvert, G. Mecacci, B. van Arem, F. Santoni De Sio, D. D. Heikoop, & M. Hagenzieker, (2020). "Gaps in the control of automated vehicles on roads" IEEE Intelligent Transportation Systems Magazine.

- G. Mecacci & F. Santoni de Sio (2019). "Meaningful human control as reason-responsiveness: the case of dual-mode vehicles". Ethics and Information Technology.

S. C. Calvert & B. van Arem (2019). "A generic multi-level framework for microscopic traffic simulation with automated vehicles in mixed traffic" In Transportation Research Part C

- S. C. Calvert, D. D. Heikoop, G. Mecacci & B van Arem (2019), "A human centric framework for the analysis of automated driving systems based on Meaningful Human Control". Theoretical Issues in Ergonomics Science.

- G. Keeling, K. Evans, S. M. Thornton, G. Mecacci, F. Santoni de Sio (2019). "Four Perspectives on What Matters for the Ethics of Automated Vehicles". In: G. Meyer, S. Beiker (eds) Road Vehicle Automation 6. AVS 2019. Lecture Notes in Mobility. Springer.

General news

- The first international workshop on Meaningful Human Control over Intelligent Machines was a success.

- Filippo is now member and rapporteur of the EU expert group on ethics in automated driving

- Giulio and Filippo are invited speakers at the "2nd International Symposium on Symbiotic Intelligent Systems", in Osaka, Japan, organized by Minoru Asada. 31 Jan - 1 Feb.

- We also give plenty of guest lectures and do lots of outreach activities. A few examplse, in the last six months we've talked at CBR's Rijvaardigheidsdag, at TUDelft's masterclass "Digital Ethics by Design", twice at Radboud University, at TUDelft's AITech, at a Groenlinks' event, and even in Zurich!

Abstracts and conferences

- F. Santoni de Sio G. Mecacci. (2019) "Meaningful Human Control over Robots in the Security Sector", Autonomous Security Systems conference, May 2-3, Zurich.

- D. D. Heikoop & M. Hagenzieker (2019). "Menselijk gedrag met automatische rijsystemen: Een kwantitatief raamwerk voor zinvolle menselijke controle",
10e VerkeersGedragDag, April 3, Utrecht, The Netherlands.

Media appearances

- Filippo interviewed in VPRO Gids 24, 2019

- Our Animation! "Who is in control of self-driving cars?"


Osaka, Japan