Towards behavioral models for autonomous driving acceptance

Foteini Orfanou\textsuperscript{a,}\textsuperscript{*}, Truls Vaa\textsuperscript{b}, Eleni I. Vlahogianni\textsuperscript{a}, George Yannis\textsuperscript{a}

\textsuperscript{a}National Technical University of Athens, 5, Iroon Polytechniou, Athens GR 15773 Athens Greece
\textsuperscript{b}Transportokonomisk Institutt, Gaustadalleen 21, Oslo 0349, Norway

Abstract

The advent of autonomous vehicles (AVs) is a breakthrough innovation in the field of transportation. At the cornerstone of autonomous vehicles (AVs) research lies the challenge of ensuring that the future vehicles can react properly and efficiently in all situations and especially in emergencies. The present work analyzes autonomous vehicle’s “driver”/operator behavior and conceptualizes the changes that should be introduced to the existing behavioral driving models in order to address the requirements of autonomous vehicles and increase the acceptance of autonomous driving. For this purpose, empirical evidence and qualitative experience from over 20 relevant projects and pilots are critically reviewed. Moreover, the conceptualization and potential parameterization of AV behavioral models are analyzed based on three popular modeling alternatives: Summala’s Multiple Comfort Zone Model, Fuller’s Risk Allostasis Model (RAT), Vaa’s Risk Monitor Model (RMM). The findings are critically discussed to reveal future research directions.

Keywords: autonomous vehicle; behavioral models; comfort zone; RAT; RMM

* Corresponding author. Tel.: +000-000-000-0000;
E-mail address: forfanou@central.ntua.gr
1. Introduction

Vehicle industry has recently allocated enormous research efforts and budget to develop and test autonomous transport systems. A significant number of sensors and powerful big data algorithms to monitor vehicles moving to the digitized surrounding environment including road infrastructure, stationary and moving obstacles have emerged. However, it has recently become evident that regardless of how fast autonomous transport research and development progress, the acceptance of users to these advancements heavily relies on the understanding of autonomous mobility, the trust to autonomous technologies, as well as the relevance of the information and service provided to users with their needs and aspirations. From a transport perspective, research has focused on analyzing and evaluating the impacts of autonomous vehicles on traffic, road safety, economy and society (Atiyeh, 2012, Shladover et al., 2012), as well as the transition period when mixed traffic, consisting of both conventional and autonomous vehicles of different levels and types will become a reality for all types of road networks (Dresner and Stone, 2008, Shi and Prevedouros, 2016, Talebpour and Mahmassani, 2016, Kakimoto et al., 2018). Many simulated environments have been proposed and developed to demonstrate the impacts of automation to traffic and road safety (Pereira and Rossetti, 2012, Morando et al., 2018, Kockelman et al., 2016) with the aim to improve the trust and awareness of road users. However, the majority fail to demonstrate a realistic and feasible autonomous vehicle navigation through traffic, especially in critical conditions where a critical manoeuvre should occur (e.g. sudden lane changing due to a vehicle stop, sudden braking due to an obstacle on the road etc).

A game changing factor for increasing awareness and acceptance of the potential users is the development of autonomous vehicles that can successfully handle the risks of traffic as an “average driver”. The driver average may be around 700 msec, a level which is learned and automated as distractions for longer periods may jeopardize safety. It should be noted that the average driver is quite successful in handling the risks of traffic as it is most likely to drive through a whole lifetime without experiencing any personal injury accident. The interaction between the driver, the car and the external driving environment is close and dynamic where risks are handled continuously and automated without much cognitive (conscious) information processing. The driver acts more or less in an automated state without any longer interruption from distractors, although shorter interruptions are tolerated within the limits previously experienced as “safe” by the individual driver. Within such a scenario, the introduction of AV-functions may become a threat to safety as the dynamic and continuous risk perception may deteriorate when the AV-function lowers the perception of risk and hence the readiness to act properly, especially in emergencies that demand immediate and reflex-like actions from the driver. It is vital that the AV-technology acts rapidly and safe as the average and experienced driver would do. Modelling behaviour of the autonomous vehicle is therefore vital for safe vehicles of high driving performance. To this end, the wealth of existing behavioural models for legacy vehicles could provide solid grounds for further research on behavioural models for autonomous traffic. Although extended research has been conducted for analyzing and investigating driver behavior and various behavioral models have been developed for identifying how driver reacts under different weather and traffic conditions as well as different operating environments and critical situations, their integration to autonomous vehicles requires modeling the “driver”/operator’s behavior of the AV of different modes and levels.

The present work addresses the above challenge by analyzing the aspects of autonomous vehicle’s “driver”/operator behavior and conceptualizing the changes that should be introduced to the existing behavioral driving models in order to address the requirements of autonomous vehicles traffic and road safety and increase the acceptance of autonomous driving. For the purposes of the analysis, the paper leverages empirical and qualitative evidences collected from over 20 relevant projects and pilots in the framework of Drive2theFuture project funded by European Commission, which aims at developing an ecosystem of tools for understanding, modelling and evaluating the impacts of autonomous driving. Extensive analyses in relation to user acceptance, behavior, accident/incident types, emergency situations and other risks as well training needs and Human Machine Interface (HMI) evaluations are conducted. Furthermore, the identified factors will be critically discussed in relation to the conceptual and technical aspects of existing behavioral models in order to identify the challenges in developing AV centered behavioral models. For this purpose, empirical evidence and qualitative experience from over 20 relevant projects and pilots are critically reviewed. Moreover, the conceptualization and potential parameterization of AV behavioral models are analysed based on three popular modeling alternatives: Summala’s Multiple Comfort Zone (Summala, 2007), Fuller’s Risk Allostasis Model (RAT) (Fuller, 2007) and Vaa’s Risk Monitor Model (RMM) (Vaa, 2007 and 2013) and they are briefly described in the next section.
2. Driver behavioral models

Three different behavioral models are addressed in the discussion. One justification for selecting just these three is their shared basis in neuroscience, i.e. adopting Damasio’s paradigm (1994). This paradigm shift is essential in opening “the neuroscientific door” for contemporary and future research in the field of road safety research. Further, all three has been developed further in making them well suited to state hypotheses to be tested in contexts associated with autonomous driving. The paradigm shift is mentioned and a second crude summary, now addressing this shift can be mirrored in the change of key concepts, conceptions and theories. Some have been abandoned and new have evolved as presented in Table 1.

<table>
<thead>
<tr>
<th>Abandoned concepts/conceptions</th>
<th>Evolving, alternative concepts/theories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk compensation</td>
<td>Behavioral Adaptation</td>
</tr>
<tr>
<td>Subjective risk/zero risk</td>
<td>Comfort through satisfying, comfort zone (Summala)</td>
</tr>
<tr>
<td>Homeostasis (Wilde)</td>
<td>Functional balance (Vaa)</td>
</tr>
<tr>
<td>Target level of risk (Wilde)</td>
<td>Target feeling/best feeling (Vaa)</td>
</tr>
<tr>
<td>Threat avoidance (Fuller)</td>
<td>Learning theory, operant conditioning, neuroscience</td>
</tr>
<tr>
<td>Task difficulty homeostasis (Fuller)</td>
<td>Allostasis: Maintaining levels of biological conditions (Fuller)</td>
</tr>
<tr>
<td>Information processing and decision making, unreflected cognitive (conscious) processes only, not separating between conscious and unconscious process (Wilde)</td>
<td>Separating deliberately between unconscious (emotions) and conscious (feelings) processes and the idea of a threshold between emotions and feelings (Damasio)</td>
</tr>
</tbody>
</table>

The first is the Summala’s Multiple Comfort Zone (2005, 2007) which incorporates all factors (inhibitory and excitatory) influencing drivers’ behaviour and the safety margins he/she keeps from the surrounding vehicles. According to this model, drivers tend to keep specific variables, such as time to collision, time to lane crossing, speed level and time headway, within an acceptable range so that they feel comfortable, satisfied and free while driving. This comfort zone enables drivers to react properly in any situation that may occur. Additionally, it is considered that if the surrounding conditions create negative feelings to the driver, he tries to adjust his speed so that level of comfort can be restored. Some of the factors Summala (2005) suggests affecting driver’s behaviour and his comfort zone boundaries are the sufficient space and time around him, the vehicle and road system, (roadway complexity), speed limits as well as driver experience and characteristics and his response style in various situations. Kovaceva et al. (2018) analysed driver’s behaviour while overtaking cyclists in the real-world through a naturalistic driving study. Minimum approaching gap, minimum distance in the steering phase, lateral clearance as well as time to collision were the parameters used to define driver’s comfort zone boundaries. The parameters minimum approaching gap and minimum distance in the steering phase were also used for analysing behaviour in case of pedestrian overtaking. Apart from these parameters, time headway and speed during the overtaking manoeuvre were used for implementing Summala’s model. Furthermore, Bårgman et al. (2015) tried to define and quantify driver’s comfort zone boundaries when turning left across the path of an oncoming vehicle.

The second model is Fuller’s Risk Allostasis Model (RAT) (Fuller, 2007, 2011). Risk allostasis theory supports that drivers try to keep a feeling of risk within the preferred range and this feeling is defining, controlling and influencing their driving behaviour and their actions (Fuller, 2007). If the perceived feeling of risk exceeds driver’s tolerable limits, he will change his behaviour in order to maintain the feeling of risk within this range (Fuller, 2008). Fuller (2007) analysed driver’s ratings of task difficulty and preferred speed while watching videos of a road segment driven at different speeds. Results showed that there a strong correlation between how difficult the task is perceived by the drivers and the feeling of risk they experience and their speed choice. The task difficulty is defined by the presence and behaviour of other road users, road environment (road surface, visibility, etc) and the vehicle characteristics. The feeling of risk depends on the driver, his motivations, his capabilities, various human factors (fatigue, distraction, level of stress, etc.) as well as the environment surroundings. Speed choice and time or distance headways are some of the basic parameters the driver changes and adjust in order to maintain the feeling of risk within the desired range. Fuller’s model and theory have been used to explain driver behaviour and reaction and has been incorporated in car – following models for various purposes. Hoogendoorn et al. (2013) and
Saifuzzaman et al. (2015) have used Risk Allostasis model for analysing the effects of driver distraction. Maximum acceleration, deceleration, speed and time headway were found to be influenced by task difficulty (Hoogendoorn et al., 2013) while Saifuzzaman et al. (2015) assumed that task demand increases proportionally to speed and unproportionally to time headway. Furthermore, driver decisions have also been modelled in cases of autonomous driving and use of ACC for resuming manual control and adjusting target speed (Varotto, et al., 2017, 2018).

The third model is Vaa’s Risk Monitor Model (RMM) (Vaa, 2007, 2013), which combines many previous developed models such as the zero - risk model (Näätänen and Summala, 1974), the risk homeostasis theory (Wilde, 1982) and Damasio’s neurobiological model (Damasio, 1994) and it is considered to better predict accident scenarios. The RMM is based on neuroscience, the axiom that survival is the most basic motive for human beings and the deductions that humans should be capable of monitoring the surroundings and detecting, realizing and avoiding any danger that threaten his/hers survival (Vaa, 2007). Furthermore, it is supported that emotions and feelings play an important role in driving behaviour and that drivers adapt their actions and reactions in order to maintain a target feeling during the driving task by choosing a driving speed, which provides two purposes: a functional balance and a feeling of risk which does not jeopardize safety.

3. Challenges and the Way Forward

Modelling the behavior of an autonomous vehicle and, therefore, developing an AV behavioral model hinders major challenges that have to be faced and overcome as has been already highlighted by researchers (Bagloee et al., 2016, Martinez-Diaz and Soriguera, 2018, Bera and Manocha 2018). First of all, the lack of relevant data that could enable the observation and investigation of autonomous vehicles (of different level of automation) contributing to understanding the way they behave under various traffic and weather conditions is critical. More data should be collected aiming at analyzing how these types of vehicles interact with other autonomous or legacy vehicles, as well as motorcyclists, and other vulnerable road users (pedestrians, cyclists). This challenge also extends to observing crashes. The involvement of an autonomous vehicle in crashes is of vital importance to be deeply analyzed, as safety constitutes a priority when developing a new (in) vehicle technology. Google is the only AV-manufacturer providing information about AV accidents (Google 2015-2016, 2015; Waymo 2017). Theo and Kidd (2017) have estimated accident risk based on data provided by Google and Waymo. Accident may serve as surrogate of AV-DBM limitations. For this reason, any accident involving an AV should in principle be explained and reveal the limitation and shortcomings of the AV-algorithms governing the performance of a given AV. Accidents with AVs will serve as proxies of the algorithms developed by car manufacturers and installed in the AVs and will assist researchers in effectively developing a valid AV behavioral model able to predict and avoid an accident.

The second challenge is related to how safety of drivers and other road users is guaranteed when it comes to the use of an autonomous vehicle. Till now, the modelling purposes and the analysis of driver behavior were satisfied by observing and testing specific factors related to the driver characteristics, his habits, his driving capabilities, as well as to the environment (traffic, weather). The advent of autonomous vehicles raises the need to rethink these factors as the human component is – partially, or fully - substituted by the “driver” (the vehicle itself). The technology is now responsible for the way the vehicle will behave under various events and interact with its surroundings and other users. A critical issue putting in risk the safety of the drivers and passengers of AVs, as well as the external road users interacting with them is the risk of hacking. All computer systems, their algorithms and their individual components are vulnerable and, unfortunately, cyber security cannot be fully ensured and guaranteed. Bank computer systems, traffic lights and air traffic control systems can be easily hacked. The same risk also lurks in the case of AVs where the algorithms regulating the AV behavior can be hacked and make the machine acts inappropriately with destructive consequences for its passengers and the other road users. Furthermore, malfunctions or system failures expose people to significant danger. Therefore, it is important to formulate recommendations and future prospects towards safeguarding the security of citizens when using AVs aiming at lowering their resistance in adopting such technologies.

The public acceptance and the adoption of the technology of the AVs are strongly connected with the third challenge referring to the issue “human vs machine”. Since the machine will be the main component of the vehicle responsible for its behavior and reaction, it is important that it will be able to effectively scan the environment and the external stimuli and adopt its performance accordingly. The machine should be able to collect continuously and simultaneously all the “messages” from the external environment and its surroundings, decode them
appropriately and choose the best action accordingly. Establish and describe processes involved in AV’s perception of risk, information processing and decision making. The reaction time and type of the machine in emergency and extreme situations is critical when establishing an autonomous driving behavioral model. The question that arises is what attributes and characteristics should the AV Driving Behavioral Model have in order to reassure the best vehicle performance based on all stimuli received?”

The increase of the penetration rate of the AVs highly depends on the magnitude of acceptance from the humans. The last challenge faced when developing an AV behavioral model, is related to the specific attribute it should entail for “resembling” to human driving behavior. “Should the humans be adapted to the machine behavior or the machine to the different driving behaviors?” Research studies and experiments have revealed different driving profiles (timid, aggressive, eco, experienced drivers, etc) with different characteristics (e.g. different reaction times, acceleration/deceleration rates, etc). Should the machine behavior resemble the behavior and characteristics of the average normal driver, it may not be accepted by humans whose driver behavior deviate from this average profile. On the other, will it be possible, efficient and safe enough to develop different AV profiles corresponding to all existing driving behaviors? The deep understanding of the acceptance and clarification of the factors -specifically feelings of comfort and feelings of risk- forming it, is, therefore, necessary, as well as, the development of an AV model predicting acceptance of driving autonomously.

Based on these challenges, the way forward and the next steps are formulated. Firstly, accident studies should be collected and deeply analyzed for two reasons. It is claimed by car manufacturers and others, that AVs will make accidents in the road system disappear and that accidents with AVs may unmask the limits of the algorithms applied in the AV-DBM being involved in an accident, the latter might be most outspoken in fatal accidents with AVs as they might be investigated in-depth and thus provide more insight into the causes of the accidents. Hence, there are two groups of accidents: i. Aggregated studies of AV-accidents, and ii. investigation of single accidents with AVs. Furthermore, studies addressing reporting experiments or analysis of autonomous vehicles should be collected, as they will reveal important findings contributing in the development of the AV Driver Behavioural Model (Norman 1990, Endsley and Kiris, 1995, Endsley and Kaber, 1999, Norman 2015, Stanton et al. 2007, Endsley, 2017). Finally, the developed model should be tested during a pilot under real conditions in order to reveal AV’s ability to avoid accidents in critical situations. This model will predict acceptance of driving autonomously and will further predict an AV’s behaviour, when another road user is expected to be present and the subsequent interaction with the road user when s/he is present.

4. Conclusions

The era of autonomous vehicles has already started and within the next years they will be all over the world occupying all types of road networks. Their acceptance and adoption by the users is strongly related to the development of an AV behavior model ready to be adopted in all traffic and weather conditions and interact appropriately with the different types of road users. The autonomous vehicle should have increased levels of safety and comforts and be able to adopt in different (non) emergency and extreme conditions ensuring the protection and security of its passengers. The present study reviewed existing behavioral models and identified challenges and further steps required for developing an AV behavioral model predicting acceptance and testing. The lack of large amount of relevant data, the prematured scientific area on the factors that may affect acceptance of automated driving, as well as the gap between humans and machines in scanning and reacting to critical conditions are considered as the major challenges.

Evidently, in the attempt to bridge the gap between humans and machines factors influencing acceptance and trust to technologies, as well as the aspirations and needs of users/drivers with respect to comfort and road safety should be jointly considered with the requirements of vehicle industry designs and specifications for autonomous vehicles and further elaborated in relation to acceptable driving maneuvers and risk. Further, the critical discussion on how existing behavioral models could be extended to include autonomous vehicles will lead to the identification and the definition of parameters, which can be used and incorporated in a behavioral model for describing the “driver/operator model for autonomous driving”.
Acknowledgements

The analysis is conducted within the framework of Drive2theFuture project (Needs, wants and behavior of “Drivers” and automated vehicle users today and into the future” funded by European Commission under the MG-3.3.2018: "Driver” behavior and acceptance of connected, cooperative and automated transport; Research and Innovation Action (RIA).

References


