A Cooperative Human-Machine Interaction Warning Strategy for the Semi-Autonomous Driving Context

Susana Costa, Paulo Simões, Nelson Costa, Pedro Arezes
ALGORITMI Centre, School of Engineering of the University of Minho, Guimarães, Portugal
surpcosta@gmail.com, paulo.simoes@gmail.com, ncosta@dps.uminho.pt, parezes@dps.uminho.pt

Abstract—This study is a part of an ongoing work regarding the possible scenarios during autonomous driving and it takes into consideration, not only academic literature and industry updates, but also the aspects described on the standards already disclosed. As autonomous driving systems become a more tangible reality, the development of an efficient warning strategy, within the human-machine interaction (HMI), is paramount, for a range of reasons that include the trust in these emerging systems. It has been noted by several researchers that a particular moment of the semi-autonomous driving is of special interest, which is the driver’s role shift from passive monitoring of the vehicle to active control of the autonomous driving system. This study presents a cooperative approach to the vehicle-driver communication strategy, accounting for both human factors and complexity of the AD systems. A nexus diagram has been developed that in a comprehensive way, provides an alternative strategy to the conventional static warning strategy, able to be customized in some specific traits, which can later be resorted to by programmers for expeditiously implementing this much needed strategy in the real context of semi-autonomous driving.

Keywords—Warning; driver; autonomous; human-machine interaction (HMI); vehicle

I. INTRODUCTION

Autonomous cars have the potential to contribute with solutions to a number of transportation challenges, including improving road safety, optimizing traffic flow, allowing for transportation that is more efficient, new mobility models, and providing additional comfort for drivers and passengers [1], [2].

Indeed, the main motivation for implementing vehicular communication systems is safety and mitigating the rampant cost of traffic collisions. Road accidents are, according to the World Health Organization (WHO), responsible for approximately 1.2 million deaths worldwide each year, and around 50 million injured people. Unless preventive measures are taken, road death is likely to become the third leading cause of death in 2020. The American Automobile Association (AAA) reported car crashes to cost, yearly, $300 billion to the United States. Vehicular communication systems are still to be optimized, though, and only experience will allow for their refinement [3]. The stages that lead to a crash go from normal driving to crash unavoidable, passing through deviation from normal driving, emerging situation and critical situation. The importance of crashes in the warning strategy relies on the fact that each of these stages defines a set of countermeasures, which, thereafter, contribute to the establishment of the warning strategy itself [4].

The conducted studies on Human-machine interaction (HMI) that assess the impact of “self-driving” functions on a “human driver” try to understand how a person can resume active control when prompted to do so, most likely in an urgent situation. So the question remains regarding on how a driver that is not in control would be able to take over from the car if needed and whether, when and how this option must be given to him. Also the passive role of monitoring may be less satisfactory than the active role of manual control, yet it may provide additional comfort [5].

The way the automation is designed will affect the driver situation awareness. In medium level automation vehicles, it can enhance safety by reducing workload or, if poorly designed, aggravate it [6]. It is therefore important that the automotive industry takes into consideration that transfer of the control to the driver as soon as a vehicle faces a situation it cannot cope with, and where liability can be issued, may ultimately denote responsible development of autonomous driving [5].

In the autonomous vehicles panorama, the subject of warnings is unavoidable. In fact, warnings are paramount for the deployment of autonomous vehicles, as the inadequate instruction of a warning can configure a product liability claim. Indeed, “failure to warn,” can render the autonomous vehicle unsafe [7].

Safety warning systems such as Lane Change Decision Aid Systems (LCDAS), Stop & Go, and Forward Vehicle Collision Warning Systems (FVCWS) monitor the driving situation and provide the traffic situation information for drivers, whereby they may warn the driver proactively about a possible hazardous situation on the basis of the vehicle’s current position, orientation, and speed, and the road situation; besides, when facing a hazardous situation, measures can be put into practice in order to control the vehicle, as these assistance systems may use the warning information to generate the expected path and control the vehicle directly [8].

Warnings have been addressed thoroughly and extensively in literature, related either to the task of driving they are connected to, the typology of crash they were designed to avoid, their influence on the response time, or even, the modality by which they are conveyed [9]–[12].

Warnings are artifacts intended to represent situations. Most warnings enclose two functions: the alerting function (iconic) and the informing function. Whilst the alerting function is emotive/motivational in nature and thus, abstract,
The technical processing stages of warnings, which shall be taken into consideration when designing warnings, are:

- Detection of object, reading data from sensor, filtering.
- Recognition of situation.
- Evaluation of situation.
- Output of warning [13].

Much work has been developed regarding the warning strategy design, considerations and requirements attending human factors [4], [12], [13], [14]–[17]. These works, however important may be, are not directly applicable to the upcoming driving paradigm of highly autonomous vehicles. As the dynamics change, so will the need to convey information to the driver (user of the vehicle). The aim of this work is, thus, to try to fill this void, by developing an adequate cooperative HMI warning strategy.

II. METHODOLOGY

The developed literature review followed the PRISMA Statement methodology [18], by which the adopted inclusion criteria included only original studies written in English and published between 2005 and February 2017. The search keywords used were “warnings”, “driver”, “HMI”, “takeover” and after the initial results the next step was to address only the topics related with “autonomous vehicles” and “transition” for those were considered the most probable to have publications of relevance to the present study.

III. RESULTS AND DISCUSSION

A study by Debernard et al. [19] presented a methodology aiming to identify and to categorize the information used by the driver in order to make “transparent” a Level 3 autonomous system. Their goal was to answer the following questions: “What shall be displayed, how and when?” in automation level 3 of the NHTSA taxonomy, which involve automated driving (AD) phases as well as transition phases where the driver has to reengage in the driving task. Each phase requires an appropriate interface to allow the driver to establish accurate situation awareness.

They then presented a series of principles to be implemented in the design of the HMI, which are:

- Principle 1: The driver should know the maximum autonomy level of the vehicle as well as the external and internal conditions that allow it to enter the autonomous mode;
- Principle 2: The driver must know which tasks the autonomous system is capable of performing, under which conditions it can perform them, and how it will perform them. The driver should know what general functions are allocated to the autonomous system;
- Principle 3: The driver must know how the system prioritizes its behavior when multiple options are possible;
- Principle 4: In the autonomous mode, the driver must be informed that the system will control the vehicle by following accepted driving practices and traffic laws (predictability of the behavior of the vehicle). Furthermore, the driver must be able to detect the actions (e.g., lane change) being performed by the vehicle and understand them;
- Principle 5: In the autonomous mode, the driver must be able to perceive the intention of the system (the maneuver it intends carrying out), why, how, and when this maneuver will be carried out;
- Principle 6: In the autonomous mode, the driver should know each maneuver that could possibly interrupt the current one. This information will help him/her avoid being surprised or frightened by what is happening;
- Principle 7: In the autonomous mode, the driver should know how a given maneuver is being carried out or why a particular behavior of the vehicle is observed;
- Principle 8: In the autonomous mode, the driver should have a sufficient understanding of what the autonomous vehicle perceives to realize its analyses and to make its decisions. The driver must be confident that the autonomous vehicle has the right information to make the right decisions and if not, he/she must be able to take control;
- Principle 9: It is important that the driver knows the boundaries of vehicle sensors, given that he/she can see information that the sensors may not receive;
- Principle 10: The driver should know what the current mode is, in order to avoid any mode confusion;
- Principle 11: The driver should clearly know how to migrate from one mode to another;
- Principle 12: The driver should know when and where the autonomous mode will be available to drive the vehicle.”

The issue of HMI design in autonomous mode is to provide the drivers with the right amount of information about vehicle operations, so that they can keep control. In this particular case, the notion of “control” does not mean that human beings drive the vehicle, but that they are aware of what is going on. In the perspective of autonomous vehicles, many projects have been conducted. The work of van den Beukel et al. [20] is part of the LRA project (French acronym for Localization and Augmented Reality), focused on designing an Augmented Reality Interface for autonomous driving at the Level 3 of automation in the NHTSA taxonomy. At this level of automation, the NHTSA specifies that vehicles enable the users to transfer full control of all safety-critical functions under determined environmental and traffic conditions, relying greatly on the vehicle to monitor for changes in those
conditions requiring transition back to driver control. This automation level involves two particular phases:

1) The AD phase, where the technical agent controls the vehicle. Consequently, the driver can carry out some secondary activities.

2) The transition from automated to manual driving, where the human agent should re-engage cognitively and physically in the driving task. This phase, if not completed properly, can lead to accidents, which is why the authors focused in these particular phases, regarding the HMI design issue [20].

Each phase requires an appropriate interface to allow the driver to be aware of what is going on outside the car, i.e., to establish accurate situation awareness. The driver must also be aware of what is going on inside the car in order to understand what the technical agent can do, what it will do, and what it has done. To this end, it is necessary to determine and display the right information, in a suitable form, and at the right time. The researchers posed three fundamental research questions to orientate their interface design, including one that is of particular interest for this study:

Q1. “In autonomous mode and in handover processing, which sufficient representation should the driver maintain or establish?”

This question may be subdivided into three sub-questions:

q1.1. “What should the driver perceive?”;

q1.2. “What should he/she understand?”;

q1.3. “Which projection of the external environment and the system should he/she perform?” [20].

It is expected that, with time, information needs will adapt depending on the driver’s experience with the system, being that in the beginning of the interaction with the autonomous vehicle, drivers are expected to demand more detailed information, which will decrease with greater contact with the system through time and, consequently, higher trust in it. According to the experts, the way to deal with this issue is to resort to adaptive and configurable information displays (e.g., through selectable information profiles) [21].

To analyze the needs of information feeding to the driver/user, this new paradigm must be discriminated in three moments: the driver is active in the primary task of driving, controlling the vehicle; the driver is using the car, which is in autonomous mode; the driver is currently using the car, which is in autonomous mode but is being asked to resume the driving.

While being the actual driver of the vehicle, the task is somewhat like driving a non-autonomous car. It can be argued that the car may, however, overrun the driver whenever safety is at stake. Nevertheless, in this case, the warning strategy shall be settled based on the recommendations already existent, which were built and proven by the extensive literature supporting those.

The two other moments, however, remain to be thoroughly studied and agreed upon.

It seems reasonable to assume that, as the vehicle is running algorithms to make decisions related to the driving task, the driver will not be fed the same amount of information as he would were he the active driver of the car. In fact, two of the predictive advantages of being driven by autonomous cars are mitigating the risk of accidents due to human poor decision-making while driving and using the time of the ride to do something other than driving (i.e., non-driving-related tasks, e.g., reading a book). Therefore, the warning strategy of highly autonomous vehicles must contemplate this shift of command between vehicle-human drivers.

Apart from the system itself, one has to account for human performance when designing an effective warning system.

All-in-all, warning systems shall aim at obtaining a hazard avoidance adequate response from the driver. In order to do so, the warning signal must first attract the driver’s attention (detection) and inform him of the situation. Afterwards, the driver then needs to comprehend the signal (identification), choose an adequate response (decision) and take action (response) [4].

The system is successful if this perception-response sequence is completed before a conflict is inevitable [4]. This means that the timing of the warnings must be computed considering the human processing of warnings. The three stages of the processing are direction of attention, situational orientation and, finally, decoding of the warning. The first stage (direction of attention) comprises: recognition of the warning; mental load, vigilance and; glance duration. The situational orientation stage encompasses also glance duration, but also situational distance to warning, effects of command and complexity of the actual situation. Lastly, the decoding of the warning comprehends sensorial modality and code type. Consequently, the reaction time to will depend on the number of options, being that the fewer available options, the lower the reaction time will be. One general rule of a good warning is that they should not depend on prior learning, i.e., the warning should be self-explanatory. Hence, reaction time to warnings will also be optimized if the warnings are compatible with driver expectations [13].

As AD technology advances, the driver’s role continues to shift from active vehicle control to passive monitoring of the AD system and environment.

AD introduces new skills needs for drivers to handle manual control recovery (MCR) [22].

Lu et al. [23] studied the human factors of transitions in AD, thereby proposing a theoretical framework to support and align human factors research on transitions in AD. The authors described AD states (static states and dynamic states) based on the allocation of three primary driving tasks: longitudinal control, lateral control, and monitoring. A transition in AD is defined as the process by which one driving state changes to another, i.e., the state of one of the players of the human-automation system changes, for instance, from monitoring activity to active control. The authors’ concept of driving states differs from the BASt, SAE, and NHTSA levels of AD, which they consider have a major limitation because these levels of automation describe how the driver and automation should
drive, whereas their proposed driving states describe what the driver and the automation are doing at a certain moment in terms of longitudinal control, lateral control, and monitoring.

Based on the answers to the questions: “Is the transition required?”, “Who initiates the transition?”, and “Who is in control after the transition?” Lu et al. [23] defined six types of control transitions between the driver and automation:

1) Optional Driver-Initiated Driver-in-Control.
2) Mandatory Driver-Initiated Driver-in-Control.
3) Optional Driver-Initiated Automation-in-Control.
4) Mandatory Driver-Initiated Automation-in-Control.
5) Automation-Initiated Driver-in-Control.

The researchers introduce use cases per transition type. Based on previous researches, this study presented the definition of transitions in AD as either an activation or a deactivation of a function, a change from one level of automation to another, a change from one state or condition to another or the period between two different states, arguing that determining the ‘states’ based on driving tasks is a prerequisite for defining a ‘transition’ in AD. In their model for describing distribution of the primary driving tasks between driver and automation at a given moment of time, a diagram illustrated the lateral/longitudinal control and monitoring of a vehicle by the automation and the driver, where input was the state of the vehicle (e.g., velocity and acceleration) and environmental information (e.g., traffic signs and surrounding road users); and output was the state of the vehicle in the environment, one system step after the input. In their model, both the Driver decision maker (a human agent) and the Automation decision maker (a computer agent) acquire and analyze the input and determine the steering and acceleration target signals, whereby both are decision makers, higher-level information processors rather than low-level trajectory-following controllers.

In contrast to handover, which is initiated by the system, proactive takeover is initiated by the driver, whose intention for steering the car is the reason for driving manually [24].

According to Merat et al. [25] one of the major challenges for highly AD is to ensure a safe driver takeover of the vehicle guidance. This is particularly important when the driver is engaged in a non-driving related secondary task, such as reading a book. It is thus crucial to find indicators of the driver’s readiness to take over and to gather more knowledge about the takeover process itself.

Regarding traffic safety, such skills should be thoroughly researched before making this technology available on public roads, especially in what concerns to critical situations, like emergencies [22].

There are three cases when the driver may need to reclaim control from the system [26]:

1) The system detects a case it cannot deal with, and tells the driver.
2) The system does not detect that the situation is out of bounds and does not notify the driver, but does something inappropriate, and the driver has to realize this himself.

3) The system breaks down and is incapable of proceeding to do anything; the driver needs to identify the breakdown in situations the system would normally comply.

Gold and Bengler [27] defined a generic procedure for takeover situations. If the autonomous system identifies a situation where a system boundary applies, it requests the driver to takeover through a Takeover Request (TOR). The moment the driver directs his/her gaze to the traffic scene, the self-driving automation shifts to manual driving within a transition area that begins when the driver starts to steer. The time budget is the term used to define the time period between the TOR and the moment when the vehicle reaches its system boundary. The transition area, in turn, refers to the transition from AD to manual driving, starting with the driver’s gaze direction at the traffic scene, and ending with the driver taking over control entirely.

Dang et al. [28] examined the effect of different warning conditions (takeover request with time budget of 4 seconds and 6 seconds vs. an additional pre-cue, stating why the takeover request will follow) in different hazardous situations. Their results indicated that all warning conditions were feasible in all situations, although the short time budget (4 seconds) was rather challenging and led to a less safe performance. The pre-cue had the positive effect of having the participants taking over and intervening earlier in relation to the appearance of the takeover request alone. Overall, the authors’ evaluation showed that bimodal warnings composed of textual and iconographic visual displays accompanied by alerting jingles and spoken messages are a promising approach to alert drivers and to request them to take over.

Higher automation levels favor the engagement in non-driving-related tasks, in which the driver must be considered “out-of-the-loop” [29].

AD gives drivers the opportunity to engage in in-vehicle tasks, frequently called secondary tasks, whether monitoring the system is mandatory or not. Performing such tasks may influence driver’s mental workload (a concept that remains hazy), and the results of neuroergonomics studies are conflicting regarding the influence of non-driving related task engagement on driving performance.

Gold et al. [30] compiled a series of studies that addressed the impact of a variety of non-driving-related tasks in takeover while aiming to quantifying the impact of traffic density and verbal tasks on takeover performance in highly AD, such as the visual Surrogate Reference Task, the cognitive n-Back Task, the conversational 20-Questions Task (TQT), and naturalistic tasks like texting, Internet search, and talking on the phone. According to the authors, the verbal task seems to deteriorate the takeover performance and that the traffic density diminished the quality of the takeover, showing delayed takeover, higher accelerations, lower TTCs, and higher crash probability in higher traffic densities.

Strand et al. [29] state that when drivers are “out-of-the-loop”, their ability to regain control of the vehicle is best if they are expecting automation to be turned off. Since disengaging the automation in a regular basis is not a practical approach for keeping drivers in the loop, the strategy should be
Based on informing drivers of their obligation to resume control of the semi-autonomous vehicle, and the best way of accomplishing it needs to be further researched.

Because no studies had yet investigated the effects of visually, cognitively and physically demanding tasks on MCR during AD mode, Young and Stanton [22] aimed at examining drivers’ actions and positions and their consequences on emergency MCR. Drivers’ involvement in visually, physically and cognitively demanding tasks that hinder the interacting with both pedals and the steering wheel was an important concern of the authors, as drivers’ being physically encumbered with an in-vehicle task could have a negative impact on MCR. In fact, controlling a car consists in mastering both longitudinal and lateral control of the vehicle, as well as being aware of the driving environment. This study proved that being engaged in an in-vehicle task could hinder drivers to quickly recover manual control of the vehicle using simultaneously hands and feet, whereas it would seem to be more appropriate than disrupted, non-simultaneous actions.

Bakowski et al. [31] establish that a functional handover assistant must enable drivers to feel comfortable in taking over control even when they are “out of the loop”, which is why the most promising strategy to compensate for system boundaries of autonomous vehicles are multimodal (auditory and visual) warnings. According to the same study, distracted drivers are capable of taking over control within a time budget of 4 seconds to 8 seconds, depending on the complexity of the situation. If drivers are provided with a longer time budget, they break less, intervene later and make less mistakes in takeover. Researchers also suggested that the vehicle should start deceleration as soon as the system alerts the driver, because results pointed to the fact that drivers might perceive this behavior as natural, especially in case of a hazardous situation.

Further research was also found necessary to investigate whether drivers generally tend to intervene before they are back “in the loop”, in which case grasping the steering wheel is not a suitable trigger for a takeover.

According to Walch et al. [32], control transitions from highly or fully AD to manual driving are difficult to perform in a safe and reliable manner and require significant amount of time. Indeed, transitions are a sensitive matter and one that needs to be thoroughly studied before being implemented, since that if done wrongly, they may affect the driver’s trust in the automation - they cannot afford to be interpreted as automation failures. At the same time, takeovers cannot be performed through having the driver state the cause for the request for control transition for it would render them too annoying. Highly automated vehicles require a better, more flexible interaction concept or they will fail to be accepted. The authors argue that automation should behave as a cooperative agent, supporting the driver to the extent of what is possible. Because the situations in which systems find their limits are usually the situations in which drivers need most support, the driver shall never be deprived of the available capabilities of the automation, in any situation.

An integrative strategy for the development of cooperative HMI is paramount, one that focus on easing communication between the human driver and the automation, and allocating tasks between the actors in a beneficial and safe manner. Shared situational awareness and bilateral understanding of intentions and actions is improved through cooperative interaction, by enabling real-time communication and maneuver planning between the two actors [33].

After extensively reviewing recent research on control transitions from highly automated vehicles at system boundaries, Shen and Neyens [33] concluded that car-driver handover concept is only a solution for a very narrow problem and a more adaptive and advanced approach is necessary for driver-vehicle interaction in highly AD. Echoing the results of other studies, the author advocates for driver-vehicle cooperation as the solution to overcome human factor issues. The authors state that there is a high need to implement directability, mutual predictability, shared situation representation, and trust in future HMI concepts. They also present systems that already implement parts of these requirements, such as the works of Payre et al. [34] and Naujoks et al. [35] and conclude that recent research is dedicated to investigating control authority transitions between automated system and human driver at system boundaries, mirroring the scientific needs motivating their own work.

Khan [36] described transitions between human control and automation as high level “design” challenges. The stance of the researcher towards autonomous vehicles is that the proponents of these initiatives are technology developers and there is a lack of convincing evidence that there is a market demand for autonomous vehicles. These experimental autonomous vehicles are undergoing tests in terms of “proof of technology” whilst initial public policy steps have been taken in a few jurisdictions to allow drivers to use their autonomous vehicles on public roads. New generation driving assistance system (NDAS) should have cognitive features that mimic non-distracted and nonaggressive driving tasks. In line with the findings of Larsson [37], the researcher also finds that these systems are intended to assist the driver and, when necessary, take corrective active safety action should the driver be incapacitated or highly distracted or if the driver selected the automation option. However, driving the cognitive vehicle does not take the driver out of the loop. The design attributes of NDAS should be influenced by human factors in driving. According to a recent news article, development of ‘human-like’ self-driving technologies is attracting investor capital [2]. Also, because the design of the driving assistance, if guided by human factors, is likely to enhance driver acceptance, safety benefits will be achieved.

It is necessary to investigate how human and technology factors can be conjugated so that the transition between human control and automation is seamless and to surpass shared authority issues in increasingly automated vehicles [5].
Regarding the out-of-the-loop state of the driver, Diels and Bos [38] have a very important vision of the matter. According to the researchers, Vehicle automation has the potential to provide significant benefits to not only the driver but also society at large. However, current concepts and scenarios put forward for self-driving cars fail to take into account basic perceptual mechanisms and run the risk of causing occupant discomfort, i.e. self-driving carsickness. As such, this may prevent the driver from activating the automation or engage in non-driving tasks. Consequently, the benefits of this technology may not be capitalized on, which may negatively affect user acceptance, technology uptake, and ultimately, the assumed positive socioeconomic impact. This work provided research questions and design guidelines to aid the design of future automated vehicle technology and avoid the occurrence of self-driving carsickness and associated negative side effects to facilitate its successful introduction. In short, self-driving cars cannot be thought of as living rooms, offices, or entertainment venues on wheels and require careful consideration of the impact of a moving environment.

Thus, the development of HMI s and control algorithms for safely transferring control between automation and drivers is a challenge for human factors researchers in AD.

Adaptive allocation of control between humans and automation can promote effective HMI. As such, Lu et al. [23] propose the creation of a switching agent that allocates tasks to the driver and/or the automation and that can determine whether transitions should happen.

Gathering all knowledge collected on the matter, one can conclude that future cooperative systems should be multimodal and adaptive. Moreover, researchers, engineers, and designers should keep the big picture of human–machine cooperation in AD in their eyes rather than focusing only on small parts to design integrative interaction concepts and HMI s.

Also, it seems logical that the adequate HMI warning strategy is one that adapts to the situation, meaning that, as opposed to classical static taxonomy, the new taxonomy of warnings has to evolve to a flexible one, in order to encompass not only the stage of the evolution of the autonomous cars, but also, one that accompanies the evolution of the relationship autonomous car – user (driver), according to the specificity of the moment.

As such, a decision tree (which is partially presented in Fig. 1) has been developed, aiming to establish as many predictable situations as possible, according to whether the driver, the vehicle or both are capable, willing or obligated to assume the control of the car. This decision tree is thought to be resorted to by programmers when attempting to compute the algorithms for the establishment of the warning strategy.

**ACKNOWLEDGMENTS**

This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/MOU1/CEC/00319/2013 and by the Portugal Incentive System for Research and Technological Development Project in co-promotion nº 002797/2015 (INNOVCAR 2015-2018).