Safe Transitions of Responsibility in Highly Automated Driving

Rolf Johansson
SP, Technical Research Institute of Sweden
Borås, Sweden
e-mail: rolf.johansson@sp.se

Jonas Nilsson
Volvo Car Corporation
Göteborg, Sweden
e-mail: jonas.nilsson@volvocars.com

Martin Kaalhus
Semcon Sweden
Göteborg, Sweden
e-mail: martin.kaalhus@semcon.com

***Abstract*—This paper presents a method for achieving functional safety for an automated vehicle system with respect to safe transitions between a manual and an automated driver, where any single mistake of the human driver is tolerated. Safety analysis and assessment of an implementation example show how to allocate safety requirements on Human-Machine Interface (HMI) components to handle the risks of unfair transitions and mode confusion. Results from this example show that it is sufficient to allocate safety requirements on the sensor of, and the lock of, a single lever to ensure safe transitions. No safety requirements are needed on visual feedback to the driver, e.g., displays.**

Keywords-functional safety; highly automated driving; safety assessment.

I. INTRODUCTION

Presently, the most critical factor for road vehicle safety is the behavior of the driver. There are different estimates, but a common understanding is that humans directly cause significantly more than 90% of serious accidents. More advanced functionality and intelligence implemented in the vehicle means that more of the responsibility to drive safely may be shifted from the skill of the driver to the capability of the functionality implemented in the vehicle.

The potential safety benefit of increased vehicle automation is undoubtedly huge but it is important that the extra risks coming from potential failures of automation are limited to a minimum. In the discipline of functional safety, there are methods to assess risks of malfunctioning E/E implemented functionality, and to reduce these sufficiently. For road vehicles, ISO 26262, [1], is the functional safety standard.

This paper focuses on systems where the vehicle in some specific situations takes full responsibility for the driving task, i.e., level 3 (L3) automation according to the scale defined by the National Highway Traffic Safety Administration (NHTSA). Regarding the responsibility of the manual driver (MD), the precise L3 definition says: "The driver is expected to be available for occasional control, but with sufficiently comfortable transition time", and furthermore: "the driver is not expected to constantly monitor the roadway", [2].

In order to prove that a L3 vehicle is functionally safe, there are two general strategies how to consider the interaction between the manual driver and the vehicle, what in the ISO 26262 terminology is denoted controllability.

In the less conservative strategy, controllability is investigated in detail for all possible scenarios where the manual driver is "expected to be available for occasional control". In the traditional research field of human factors, this is a research question that is currently very much investigated [3], [4], [5]. A rather recent overview of what controllability assumptions that are reasonable on NHTSA L2 and L3 is found in [6].

In the more conservative strategy, there are no assumptions that the manual driver can take back control within a bounded time. One could say that we do not require the driver to be comfortable with any short transition time. This strategy is the one chosen by Volvo Cars in the DriveMe project [7], where the vehicle takes full responsibility to safely handle any critical situation during automated driving. We can call this an autopilot with full responsibility for safety, as it does not need to rely on any responsiveness from the manual driver to stay safe.

An unsolved question so far, is if the more conservative assumption about the human capability still can enable the design of a functionally safe car. The introduction of an autopilot with full responsibility leads to two new challenges within functional safety. We need to ensure safety when the autopilot is in charge, but also ensure safe transitions between the manual driver and the automated driver (AD). This paper investigates the latter.

The contribution of this paper is a method for achieving functional safety for an automated vehicle system with respect to safe transitions between a manual driver and an autopilot with full responsibility for safety. We assume that both the human driver and the autopilot are capable of safe driving, as well as judging its own ability to drive safely. Thus, neither the driver nor the autopilot are required to take control and thus the vehicle will be in a safe state if either the driver or the autopilot accept to take control of the vehicle.

There are simulator studies suggesting that human drivers may change their driving behaviour when taking back control from an autopilot, [8]. This is not considered in this paper as we focus on functional safety rather than design of the HMI or autopilot driving behaviour.

This paper is organized as follows. Section II describes the new hazards related to the driving mode transitions introduced by NHTSA L3. In Section III, we discuss how to

 ISBN: 978-1-61208-492-3

define a safe transition and the acceptable level of fault tolerance. Section IV elaborates on possible implementations using a system example and corresponding functional safety analysis and assessment. Finally, Section V presents concluding remarks.

II. WHAT CAN CAUSE THE ROAD VEHICLE TO BE UNSAFE

One interpretation of a hazard analysis & risk assessment (HA&RA) today according to ISO26262 is that the vehicle itself is considered safe, if it only puts the driver in situations that are possible to manage safely. The driver is ultimately responsible for safe driving, and the malfunctions of the vehicle should be restricted in such a way that the driver can keep the vehicle in a safe state. The explicit method for determining the requested Automotive Safety Integrity Level (ASIL), restricting a certain hypothetical vehicle failure, is to measure three factors: exposure, severity and controllability. The two first factors are the traditional ones that are part of the definition of risk, i.e., a combination of probability and severity. The third factor is the one that takes into account that the driver may sometimes have a possibility to keep the vehicle safe, even though the ordinary (safety-related) functionality is failing.

When we shift from a situation where a manual driver has the ultimate responsibility, to highly automated driving where the manual driver and an autopilot are alternating, this will have an impact on the HA&RA. So, what will become different when going from NHTSA L2 to L3? This new challenge has partly been addressed in [9].

We require the same from an autopilot as from a manual driver. This means focusing on a safe style of driving, making the driver capable to handle also unexpected events. When programming an autopilot, this is what we cover on the tactical level [10], [11]. The autopilot should always choose to perform the maneuvers in such a way that reasonable, but still unexpected, situations could be handled safely. For example, the decision whether or not to initiate an overtaking maneuver is on the tactical level. An optimistic decision to overtake may cause the vehicle in a situation where avoiding one accident may cause another. The solution to this dilemma is of course to initiate an overtaking maneuver only when the entire operation is foreseen to be possible to fulfil in a safe manner.

Note the contrast to Advanced Driver Assistance Systems (ADAS), where the vehicle takes over only on the operational time scale, and then assumes the manual driver to continue according to the (maybe revised) tactical plan. The ADAS functionality today, does not take the ultimate responsibility to drive the vehicle safely. Firstly, it only operates on the operational time scale. Secondly, it only assists the manual driver. When the responsibility is transferred from the manual driver to the autopilot, there is no longer an assistance relation. The transfer means that from then on, the automated driver is fully responsible for driving the vehicle safely.

Given that the autopilot can drive safely once in command, the HA&RA must also cover the transitions between the driver and the autopilot. In NHTSA L3, these transitions introduce two new types of hazards, namely unfair transitions and mode confusion. These are described in detail in the following sections.

A. *Unfair transitions*

As we noted in the above section, it may be complicated for the driver to make a proper override of a failing tactical decision of the automated driver. This is because drivers may find different tactical solutions to a certain driving situation, and each of these may be correct. It may be hard for a driver to distinguish a faulty tactical decision from a one that is just different from his or her own favorite pattern. Even more, it may be very hard to continue to fulfill a tactical plan of another driver if the responsibility is transferred in the middle of the intended sequence. This difficulty is both for a manual driver to continue a plan of the automated driver, and for the automated driver to continue what has been initiated by the manual driver.

If the manual driver realizes that the automated driver has handed over responsibility, without the manual driver agreeing to this, this is a new risk to consider when entering NHTSA L3. We can say that the manual driver is put in a situation of *unfair transition*. For a driver with the same understanding of the planned tactics, the situation may be easy to handle, but an unfair transition may put the driver in a situation where driving safely will be very difficult.

The problem of unfair transitions may appear in both directions. It is reasonable to assume that the automated driver can drive safely as long as it can choose its own tactics. This is a far easier task than being able to understand and solve arbitrary situations.

To summarize, if the responsibility is transferred from one driver to the other, this must include a confirmation from the receiving driver. Otherwise, the transition may be regarded as unfair, and it is a non-negligible risk that the second driver is incapable of handling the situation, on both operational and tactical time scales.

B. *Mode confusion*

In order to make the entire trip from start to stop safe, it is critical that the two drivers always agree on which of them that currently is in charge. If they misunderstand each other, there is a risk that either there are two drivers trying to control the vehicle, or there is no one taking care of the ride. Both these potential *mode confusions* need to be addressed.

If we allow both the manual driver and the automated driver to override each other, there is an obvious risk that the resulting non-harmonized commanding of the vehicle may result in dangerous situations. This is especially probable because the two drivers most likely make different tactical decisions now and then, and as consequence regard the operative command of the other as faulty. For safe driving in NHTSA L3, it is important to reduce the risk of this reciprocal *override*.

It is perhaps even more obvious that it will become dangerous if neither the manual driver nor the automated driver regard herself as the ultimately responsible. Such reciprocal *underride* is therefore obviously important to reduce properly when performing the risk assessment for driving on NHTSA L3.

 ISBN: 978-1-61208-492-3

C. State-of-the-art comparison with other industries

This section describes technology, systems and concepts from other industries where similar problems arise caused by mode confusion and unsafe transitions. The focus has been on nuclear, rail, avionics and space since these industries deal with complex systems, is in a regulated environment and all demand active users for proper operations. Experiences from other industries give valuable insight into how to design interfaces and processes that ensure safe transitions in the context of autonomous driving. However, these inspiration sources material and solutions need to be adapted to fit into the automotive context in order to be a viable tool.

As the existing autonomous systems within the automotive industry are still in their infant stages and the majority of them still are semi-autonomous (i.e. NHTSA L1-L2) at time of writing, these systems are excluded. The interested reader may study results from several research efforts on this topic; PReVENT, HAVE IT, ADAPTIVE and INTERACTIVE to mention a few.

When reviewing earlier experiences from nuclear, avionics, rail and space industries we make one important observation. Within these industries, the technical solutions are operated by educated users, certified to use the specific equipment, often in controlled environments and in cooperation with colleagues supporting them.

In avionics, there is a system called Auto Ground Collision Avoidance System (AGCAS) that monitors the pilot's response in certain situations and if the pilot does not respond to an alarm, the system takes over and performs the necessary manoeuvre. After avoiding the threat, control is returned to the pilot. Inagaki describes this as situation-adaptive autonomy where authority over a system is transferred between human and machine agents, [12]. However, the main point of reference within avionics is that an educated pilot is responsible for operation of the airplane at all times, differing from the automotive situation.

Two major players in the avionics industry, Boeing and Airbus, apply different philosophies regarding automation. Boeing implements a strict assisting role for technology and automation, where the pilot always acts as the final authority. Airbus rather sees automation as a way of enhancing flight performance by assisting the responsible pilot. This subtle difference in philosophy causes different problems, where the Boeing strategy allows the pilot to perform errors that may cause accidents and the Airbus strategy may interfere and prevent the pilot from performing necessary maneuverers needed for safety in extreme situations [13] [14].

Within nuclear there are numerous processes to monitor. This is handled with different interfaces displaying process information. In Sweden there are different systems ensuring correct decisions regarding the operation of the nuclear plant. There are regulations stating that the plants are to be designed in such a way that operators always have a 30 minute window to perform an action. In other words, the plant is fully autonomous for 30 minutes at a time. There are also mechanisms that require several users to acknowledge an action independently in order to perform it, which can be compared to the needed protocol in automotive. However, the time constants at play are significantly different when compared to the automotive setting.

In the nuclear industry, one main control board always represents the true state of the processes. It is assured to a higher safety integrity and acts as the primary source of information should different sources provide inconsistent information. The inconsistent information does pose less of a problem since nuclear operators are well educated with the system and knows what information to depend on. However, translating this into the automotive setting is problematic, where most of the information sources primary purpose is to enhance and ease the experience rather than to provide safety-assured information on the system state.

Within the space industry, interfaces and systems are often complex and users need to understand how to use many different systems at the same time for proper operation. Because of this, a lot of effort is put into mental models of the systems; the users do not need to understand how and why the systems work only have a basic understanding of what situations the system can be used in and what the outcome would be. This could be translated into the automotive setting where the situation is similar, although on a smaller scale and users need to quickly gain a basic understanding about the autonomous systems capabilities and limitations.

Studies from the rail industry have analysed operator workload and the possibilities of it causing human errors. Two main ways of managing human performance have been formulated, through either technology or human resource management. Assessment of individual possibilities to manage the required workload has been performed through psychometric testing, as well as limiting workdays and issuing regular breaks [15].

As the automotive setting makes it difficult to limit usage periods, the technology and interfaces must be designed to ensure safe usage under these circumstances. Adaptive interface features linked to specific task requirements with consistency in interface design across different modes of system operation is recommended in order for the users to effectively apply mental models [16].

III. WHAT MAKES A TRANSITION SAFE

In the previous section, we have listed new categories of risks to handle related to the dual driving modes when going up in automation degree to NHTSA L3. In the following sections we summarize our current understanding on how to handle these.

A. General Strategy

A main strategy to eliminate *unfair transitions* is to introduce a fair procedure for handover. This means that the current responsible driver (manual or automated) stays responsible until there is an agreement for a handover. If we can find out how to design safe handover of responsibility, this will then solve the problem of unfair transitions. For a handover to be regarded as safe, we need to address both what is reasonable to assume of a driver, and what safety

requirements we need to allocate on the elements implementing the vehicle part of the handover protocol.

The problem of *Mode confusion* can be solved by combining safe handover mechanisms with requirements on each of the two candidate drivers to remember who is currently in charge. When the automated driver is responsible, the manual driver should then try to avoid interfering with the AD. This can be solved by not allowing the MD to have any impact on the vehicle, if not first going through a handover procedure. If we want, we can transfer part of that manual responsibility to the vehicle by putting safety requirements on ignoring any try from the manual driver to control the motion of the vehicle. Furthermore, we require of the manual driver to stay responsible once becoming in charge. In a similar way, we put safety requirements on the vehicle to remember who the agreed responsible driver is.

B. Fault Tolerance

As stated in the previous section, we require from a safe transition that the two candidate responsible drivers (MD and AD) regard the transitions fair and have a common understanding who has received the responsibility. This implies that both drivers need to explicitly confirm that a transition is possible and fair to perform. Furthermore, it implies that both drivers really are aware of what has been agreed.

Already today, we have a substantial amount of serious traffic accidents caused by driver lapses. There is no reason why not to regard the manual driver of a highly automated vehicle as prone to mistakes in any HMI, including the one for transition of responsibility. Because of this, we need to have a procedure where the manual driver has to perform several and coordinated actions, in order to allow a transition. Every single action can be assumed as performed by mistake, but the more of coordinated multiple actions that are required the less probable it is that the driver is not aware of what she or he is doing.

For the vehicle, we assume that safety requirements are allocated to all elements critical for achieving a transition in such a way that it can be considered as fair and consistently understood by both drivers. We make a conservative assumption that the ASIL attribute to use is the one that is representing the highest ASIL among the possible induced hazards. In practice, this means that we need ASIL D on guaranteeing freedom from mode confusion and from unfair transitions. Of course, redundancy patterns may be applied allowing the ASIL D to be decomposed onto different elements of the implementation.

A way to argue that a transition is safe is to check what happens if there is either a manual mistake or an E/E failure, or combination of these. This must be checked for any state in the transition protocol. For any hazardous consequence, it must be shown that the corresponding E/E failure is prevented with an appropriate safety requirement. If a manual failure may lead to a hazardous consequence even in a fault free case, the protocol implementation is obviously not robust enough.

IV. General Implementation Suggestion

In order to make a transition tolerant to any single manual mistake, there are a few different general ways to design the protocol. The redundant action from the manual driver can in general be either in time or in space, or a combination of these. By time redundancy, we mean here to request a sequence of actions where the second must follow in a certain time interval after the first one. Space redundancy is on the other hand when the manual driver is requested to apply several actions simultaneously. In both cases, the idea is that it can be argued that the set of actions is extremely unlikely to be performed by mistake.

A. Example HMI Protocol and Implementatios

As an example in this paper, we chose to describe a protocol based on manual time redundancy. This means that we always require two actions from the driver for any transition from MD to AD or from AD to MD. Furthermore, we say that the second action of the manual driver defines the transition, which means that there is no requirement on the manual driver to observe the resulting outcome correctly, more than knowing what she is doing herself. As long as the second action is fulfilled, the transition is deemed to have occurred.

In Figure 1, a general protocol is illustrated, where two coordinated actions are required from the manual driver. When implementing this it is important not to allow the driver to perform the second action, without having acknowledged the first one.

In this example, we chose the first action to be a press of a button and the second to be a change of lever position. This lever has exactly two possible positions: AD and MD. The vehicle is always started in MD, and the driver may change the mode after reaching the proper state in the transition protocol. We consider the lever to be locked at any other time. Furthermore, if the lever is not moved fast enough after getting acknowledge by the autopilot, it will be locked again requiring the protocol to start over again in order to perform a transition.

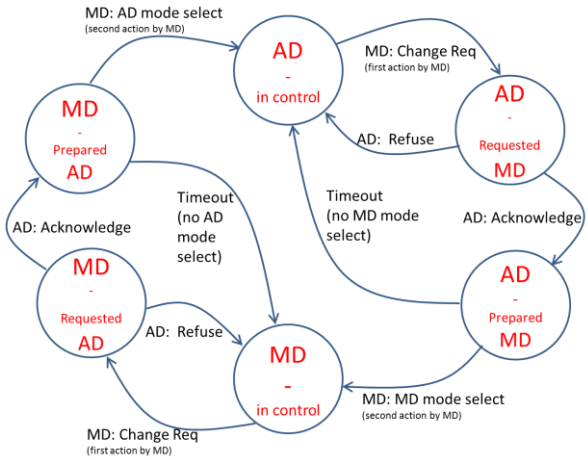

Figure 1. Example of a simple transition protocol.

This protocol is based on the assumption that it is always safe to keep the mode if nothing else is agreed. The current driver should always be able to continue to take care of the vehicle in a safe manner.

We can extend the protocol to cover the cases where the AD can suggest a transition, either by declaring that the AD is ready to take over from the MD, or by telling the MD that the AD performance is limited. Such a protocol is depicted in Figure 2.

To implement this protocol we suggest the following HMI components:

- Telltale light showing the AD view of preferred mode
- Pushbutton to for the MD to ask for mode change (first action)
- Telltale light showing whether the AD is prepared for a change as requested by the MD
- Lever for the MD to select mode (second action)

Any failure mode of these four HMI components then needs to be included in the safety analysis, and this in combination by any single mistake by the manual driver.

To summarize, a fault-free uninterrupted transition from the MD to the AD in this example follow the steps:

- The MD drives the vehicle (MD mode)
- The AD declares it is ready to take over by changing the preference telltale to AD available
- The MD asks to take over by pressing the pushbutton
- The AD acknowledges that it is prepared by indicating the readiness telltale and unlocking the lever
- The MD changes the lever to AD position
- The AD locks the lever, and continues to drive in AD mode

The transition from AD to MD is performed in a similar way, i.e., the MD may either independently, or suggested by the AD, start by asking for a mode change. The AD then acknowledges by indicating on the readiness telltale and unlocking the lever. Finally, the MD changes the lever to the MD position and starts to drive manually.

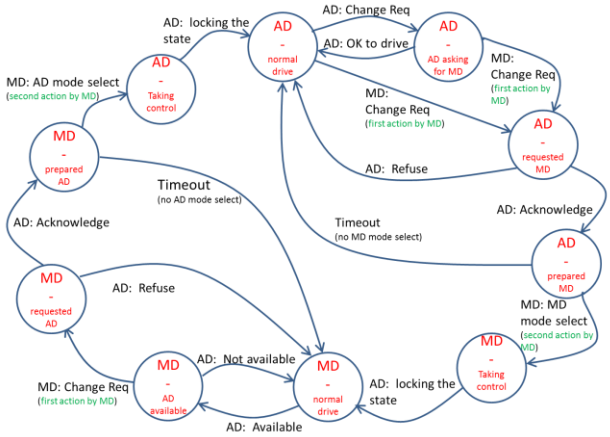

Figure 2. Example of an elaborated transition protocol.

B. *Safety Analysis*

In the following section, the above protocol and implementation is analyzed with respect to its sensitivity to any human mistake, vehicle component failure, or a combination of these. Hence, we walk through the detailed state diagram and investigate the possible failure consequences at any state.

When doing the safety analysis, we document the result in Table 1. The columns are:

- Protocol state
- HMI failure to investigate
- Possible driver mistake
- Consequence in words
- Consequence in terms of safe/unsafe

Each row in this table marked as unsafe in the last column needs to be protected by a corresponding safety requirement allocated to restrict this HMI failure. If all occurrences of an unsafe consequence are protected by appropriate safety requirements, the protocol implementation is deemed safe. In order for the safety argumentation to be valid, it is important that the table is shown to be complete. This includes an argumentation that all possible human mistakes are considered.

C. *Safety Assessments*

As concluded from the safety analysis in Table 1, there are three ways for the example protocol to fail in an unsafe way, caused by either of a manual mistake, a vehicle component failure, or a combination of these. The three failures that we need to avoid to maintain safety are:

- The AD cannot correctly sense the mode lever position, which may cause mode confusion.
- The AD cannot guarantee lock of the mode lever according to the protocol. This in combination with the MD moving the mode lever to AD mode, without noticing it, may cause mode confusion (or unfair transition if discovered by the MD).
- The AD cannot guarantee locking of the mode lever according to the protocol. This in combination with the MD changing lever position from MD to AD, without getting acknowledgment of a prepared AD, may cause unfair transition.

As we assume that the MD may make any single failure at any time, the way to argue for avoiding the above failures is to put the entire responsibility on the vehicle. This implies that we put two safety requirements on the HMI.

- ASIL D on restricting faulty lever sensor, i.e., the lever sensor needs to be always correct.
- ASIL D on restricting lever lock faulty unlocked (faulty locked consider as safe).

If we can guarantee that the HMI is implemented according to these two safety requirements we can claim that we make a safe transition even in the presence of an arbitrary single manual mistake. This handles both the mode confusion and the unfair transition aspects of a safe transition.

 ISBN: 978-1-61208-492-3

TABLE I. SAFETY ANALYSIS OF TRANSITION PROTOCOL

Protocol state	HMI failure	Driver mistake	Consequence	Safe/ Unsafe
MD - normal drive	Fault in lever lock	No	MD driver not trying to touch lever. Stay in MD.	Safe
MD - normal drive	Fault in lever lock	Driver changes lever position without asking for change first.	Unfair transition.	Unsafe
MD - normal drive	Fault in preference telltale	Any mistake or correct behaviour	MD cannot change locked lever. Stay in MD- normal drive.	Safe
MD - AD available	Fault in lever lock	No	MD driver not trying to touch lever. Stay in MD.	Safe
MD - AD available	Fault in lever lock	Driver changes lever position without asking for change first.	Unfair transition.	Unsafe
MD - AD available	Fault in preference telltale	No	Stay in MD	Safe
MD - AD available	Fault preference telltale	Driver ignores lack of availability	Transition sequence fulfilled. Change to AD.	Safe
MD - requested AD	Fault in pushbutton	Any mistake or correct behavior	No Acknowledge by AD. Lever still locked. Stay in MD.	Safe
MD - prepared AD	Fault in prepared telltale	Driver correct: Driver stops transition sequence	Time-out in protocol. Stay in MD.	Safe
MD - prepared AD	Fault in prepared telltale	Driver incorrect: Driver ignores lack of ack.	Transition sequence fulfilled. Change to AD	Safe
MD - prepared AD	Fault in lever lock	Driver correct: Driver tries but cannot fulfil transition sequence.	Time-out in protocol. Stay in MD.	Safe
MD - prepared AD	Fault in lever lock	Driver incorrect:Driver doesn't continue transition sequence.	Time-out in protocol. Stay in MD.	Safe
AD – taking control	Fault in lever sensor	Any mistake or correct behavior	Mode confusion	Unsafe
AD – normal drive	Fault in lever lock	No	MD driver not trying to touch lever. Stay in MD.	Safe
AD – normal drive	Fault in lever lock	Driver changes lever position to MD without asking for change first, and without noticing what is happening.	Mode confusion. (Unfair transition, if realized later).	Unsafe
AD – normal drive	Fault in preference telltale	No	MD acts as in normal AD mode. Stay in AD or ask for transition.	Safe
AD – normal drive	Fault in preference telltale	Driver tries to changes lever position but it is locked in AD position.	Stay in AD.	Safe
AD – asking for MD	Fault in lever lock	No	MD not touching lever without asking for change first. Stay in AD.	Safe
AD – asking for MD	Fault in lever lock	Driver changes lever position by mistake without noticing it in the first place, and without asking for change first.	Mode confusion (Unfair transition, if realized later).	Unsafe
AD – asking for MD	Fault in preference telltale	Any mistake or correct behavior	MD can request MD mode or stay in AD mode.	Safe
AD – requested MD	Fault in pushbutton	Any mistake or correct behavior	No Acknowledge by AD. Lever still locked. Stay in AD.	Safe
AD – prepared MD	Fault in prepared telltale	No	Driver stops transition sequence. Time-out in protocol. Stay in AD.	Safe
AD – prepared MD	Fault in prepared telltale	Driver ignores lack of ack.	Transition sequence fulfilled. Change to MD	Safe
MD – taking control	Fault in lever lock	Any mistake or correct behavior	Driver tries but cannot fulfil transition sequence. Time-out in protocol. Stay in AD.	Safe
MD – taking control	Fault in lever sensor	Any mistake or correct behavior	Mode confusion	Unsafe

If ASIL D sensors and/or ASIL D locks are considered either unavailable or very expensive, we may consider redundancy implementation techniques. Instead of one sensor always telling the correct lever position with ASIL D attribute, we may consider three (sic!) sensors each with ASIL B. If at least two of the three are correct, we can stay safe. This means that we need to restrict that two of the three are failing. This shall be guaranteed with a total ASIL D, which we distribute as ASIL B on each sensor. Similarly, using ASIL A sensors would require seven times redundancy. If four out of seven are working we consider it as safe. This means that we need to restrict that four of the sensors are failing. This shall be guaranteed with a total ASIL D, which we distribute as ASIL A on each sensor.

Instead of one lever lock always guaranteeing that the lever is never faulty unlocked, we may consider two locks each with ASIL B. We consider faulty locked as a safe state. If at least one of two locks can guarantee freedom from faulty unlocked, we can stay safe. This means that we need to restrict that both of the two locks are faulty unlocked. This shall be guaranteed with a total ASIL D, which we distribute as ASIL B on each lock. Similarly, using locks guaranteeing absence of faulty unlocked with ASIL A would require quadruple redundancy. If only one of the locks is avoiding faulty unlocked, we consider it as safe. This means that we need to restrict that all the four locks are faulty unlocked. This shall be guaranteed with a total ASIL D, which we distribute as ASIL A on each lock.

 ISBN: 978-1-61208-492-3

V. CONCLUSION

When introducing an autopilot which in some driving situations takes full responsibility to drive the vehicle, it becomes crucial to ensure safe transitions between the manual and the automated driver. The existence of dual driving modes brings two new sources of risk, namely unfair transitions and mode confusion.

We propose to define a safe transition as a transition where either a manual mistake or an E/E failure, or combination of these, leads to an unfair transition or mode confusion. Furthermore, we demonstrate on a system example how to allocate safety requirements on system elements to ensure safe transitions.

Results from this example show that it is sufficient to allocate safety requirements on the sensor and lock of a single lever to ensure safe transitions. No safety requirements are needed on visual feedback to the driver, e.g., displays. We remark that the example implementation by no means is a unique solution to the safe transitions problem.

ACKNOWLEDGMENT

This research has been supported by the Swedish government agency for innovation systems (VINNOVA) in the FUSE project (ref 2013-02650).

REFERENCES

[1] ISO, "International Standard 26262 Road vehicles -- Functional safety", November 2011.

[2] National Highway Traffic Safety Administration, "Preliminary Statement of Policy Concerning Automated Vehicles", http://www.eenews.net/assets/2016/01/14/document_pm_01.pdf, retrieved: June 2016.

[3] C. Gold, D. Damböck, K. Bengler, and L. Lorenz, "Partially Automated Driving as a Fallback Level of High Automation," 6. Tagung Fahrerassistenzsysteme. Der Weg zum Autom. Fahren., 2013.

[4] M. H. Martens and A. P. Van Den Beukel, "The road to automated driving: Dual mode and human factors considerations," IEEE Conf. Intell. Transp. Syst. Proceedings (ITSC) , 2013, pp. 2262–2267.

[5] F. Naujoks, C. Mai, and A. Neukum, "The effect of urgency of take-over requests during highly automated driving under distraction conditions," Adv. Hum. Asp. Transp. Part I, vol. 7, July 2014, p. 431.

[6] National Highway Traffic Safety Administration, "Human Factors Evaluation of Level 2 And Level 3 Automated Driving Concepts Past Research, State of Automation Technology, and Emerging System Concepts", http://www.nhtsa.gov/DOT/NHTSA/NVS/Crash%20Avoidance/Technical%20Publications/2014/812043_HF-EvaluationLevel2andLevel3AutomatedDrivingConceptsV2.pdf, retrieved: June 2016.

[7] Volvo Cars, "THE SELF-DRIVING CAR IN ACTION – DRIVE ME", http://www.volvocars.com/intl/about/our-innovation-brands/intellisafe/intellisafe-autopilot/drive-me, retrieved: June 2016.

[8] S. Brandenburg and E. Skottke, "Switching from manual to automated driving and reverse: Are drivers behaving more risky after highly automated driving?," IEEE 17th Int. Conf. Intell. Transp. Syst. (ITSC), 2014, pp. 2978–2983.

[9] R. Johansson, C Bergenhem, and H. Sivencrona, "Challenges of Functional Safety in ADAS and Autonomous Functions", SAE World Congress, Detroit, April 2014.

[10] R. Sukthankar, "Situation Awareness for Tactical Driving", Ph.D. thesis, Robotics Institute, Carnegie Mellon University, USA, January 1997.

[11] T. X. P. Diem and M. Pasquier, "From Operational to Tactical Driving: A Hybrid Learning Approach for Autonomous Vehicles", 2008 10th Intl. Conf. on control, Automation, Robotics and Vision, Hanoi, Vietnam, December 2008.

[12] T. Inagaki, "Design of human–machine interactions in light of domain-dependence of human-centered automation", Cognition, Technology & Work, Volume 8, Issue 3, 2006, pp 161-167.

[13] A. Marinik, R. Bishop, V. Fitchett, J. F. Morgan, T. E. Trimble, M. Blanco. "Human factors evaluation of level 2 and level 3 automated driving concepts: Concepts of operation." (Report No. DOT HS 812 044). Washington, DC: National Highway Traffic Safety Administration., July 2014.

[14] H. Orlady, R. Barnes, "A Methodology for Evaluating the Operational Suitability of Air Transport Flight Deck System Enhancements", SAE Technical Paper # 975642, 1997.

[15] J. Cunningham "Break the monotony." Professional Engineering, 20(20), 33-33. 2007.

[16] D.B. Kaber, L. J. Prinzel, "Adaptive and adaptable automation design: A critical review of the literature and recommendations for future research." (NASA/TM-2006-214504), September 2006.

 ISBN: 978-1-61208-492-3