Autonomous Driving – How to Apply Safety Principles

Hendrik Schäbe, TÜV Rheinland InterTraffic GmbH, Köln, Germany

Hendrik Schäbe

Abstract: *We discuss safety principles of autonomous driving road vehicles. First, we provide a comparison between principles and experience of autonomous or automatic systems on rails and on the road. An automatic metro operates in a controlled and well-defined environment, passengers and third persons are separated from driving trains by fences, tunnels, etc. A road vehicle operates in a much more complex environment. Further, we discuss safety principles. The application of safety principles (e.g. fail-safe or safe-life) is used to design and implement a safe system that eventually fulfils the requirements of the functional safety standards. The different responsibility of human driver and technical driving system in different automation levels for autonomous driving vehicles require the application of safety principles. We consider, which safety principles have to be applied using general safety principles and analysing the relevant SAE level based on the experience from projects for the five levels of automated driving as defined by the SAE. Depending on the level of automation, the technical systems are implemented as fail-silent, fails-safe or as safe-life.*

Keywords: *safety architecture, autonomous driving, road vehicles*

For citation: *Schäbe H. Autonomous Driving – How to Apply Safety Principles. Dependability 2019; 3: 21-33 p. DOI: 10.21683/1729-2646-2019-19-3-21-33*

1. Introduction

Autonomous driving on the street [2] has become more and more popular and the first demonstrator systems are operational [4,10,21]. On the other hand, automatic metros and people movers are already successfully working for many years.

In this paper, we compare the different levels of automation as defined by UITP [23] and SAE [22] and their meaning for the system. In addition, manual fallback modes are considered.

For road vehicles, currently a large number of assistance system is available that are able to handle specific situations. This leads to the impression that these vehicles move autonomously.

In general, the situation for a road vehicle is much more complex than that of a train.

We describe differences regarding approval for automated metros, road vehicles and so called automated guided vehicles (AGV). Legal requirements for homologation of road vehicles according to the convention on road traffic are discussed and the implication for the system and the behavior of the driver.

Autonomous driving has become a very important subject of research and first pilot projects. In safety technology, the application of safety principles as e.g. fail-safe or safe-life is a very important tool to design and implement a safe system that eventually fulfils the requirements of the standards for functional safety. Safety principles have already been described and applied to guided transport systems, including system with immaterial guidance principles.

In earlier papers, safety principles have been described and later applied to guided driving.

In the present paper, we systematically consider which safety principles have to be applied for which SAE level of autonomously driving systems und we show how an autonomous system could be built. This is partially done with the help of general safety principles, partially by analysing the relevant SAE level based on the experience from several projects.

According to UN resolution [24] or SAE [22], autonomous driving on the road knows five different levels:

- 0 No automation
- 1 Driver assistance
- 2 Partial automation
- 3 Conditional automation
- 4 High automation
- 5 Full automation

For the levels 0-2, the driver is fully responsible for driving, starting from level 3 the automated driving equipment monitors the vehicle.

This different responsibility of human driver and technical driving system requires the application of safety principles. In the present paper, we systematically consider which safety principles have to be applied for which level und we show how such a system could be built.

This is partially done with the help of general safety principles, partially by analysing the relevant level.

We start with a very simple and abstract model of the system and show that there exist different possibilities to implement autonomous driving. An important result is that an arbiter needs to be installed that gives the human driver the possibility to override the decisions of the autonomous system to fulfil legal requirements.

For the five levels of automated driving as defined by the SAE [22], safety principles are derived. For the levels 0-2, the driver is fully responsible for driving, whereas starting from level 3, the automated driving equipment monitors the vehicle. To give the driver the possibility to intervene, means that this must be implemented according to the relevant safety integrity level and that the driver must have enough time to take over control. The latter strongly depends on the level of automation and the speed and the environment in which the vehicle moves.

Depending on the level of automation, the technical system are implemented as fail-silent or as safe-life. There are also exclusions, when the technical systems can be implanted as fail safe, when the vehicle always can be brought to a safe stop, e.g. when driving with low speed and on a controlled territory.

We consider the two main functions of guidance and braking / acceleration and their role for autonomous driving. Moreover, detection and reaction with regard to fixed and moving obstacles is discussed.

Two basic requirements for autonomous systems are that they need to be developed according to the relevant standards of functional safety fulfilling an ASIL (or SIL) level and that the capability of the autonomous driving system must at least on the same level as that of a human driver.

We note that Wachenfeld[26] has proposed a stochastic approach to show that an autonomous system fulfils a certain level of performance or safety. This, however, can only be seen additional evidence, the main evidence for a safe system is an appropriate safety architecture implemented according to the rules of functional safety, see ISO 26262 [18].

We sketch the current technical possibilities for automated driving and the existing technical solutions. Especially, we discuss the possibilities and restrictions of artificial intelligence. We briefly describe a roadmap of possible next steps.

2. The status with metros, people movers and road vehicles

2.1. Metros and people movers

In many cities in the meanwhile automated metros and automated people movers are working

Examples are

- On the New York City Subway, the BMT Canarsie Line.

• On the London Underground, the Central, Northern, Jubilee, and Victoria lines run with ATO.

• On the Nuremberg U-Bahn, existing U2 and new U3 lines converted to ATO.

• On the Barcelona Metro, the L9 (as the Europe's longest driverless line), L10 and L11 runs with ATO.

• The Rio Tinto Group has the iron ore railway driverless go-ahead.

• The Tren Urbano, has an Siemens ATC system that allows for fully automatic operation.

• The Vancouver SkyTrain.

• Frankfurt Airport Skyline.

• Copenhagen Metro.

• On the Milan Metro, the M1 Red Line runs with ATO.

On the Mass Rapid Transit (Singapore), all lines operating currently run with ATO since 1987.

For metros and people movers, a principle of separation has been applied: The automated trains are separated from all other traffic, running in the tunnels, open track is separated by fences, platform screen doors are used to separate the trains from passengers. This simplified the exploitation conditions significantly.

The automated train protection system (ATP) is used to prevent collision and derailment. This allows also manually operated trains to use the same network.

The normal safety requirement for the ATP is a safety integrity level SIL 4. Nevertheless, manually operated fallback modes exist. Partially stewards are present to assist the passengers, especially in case in case of evacuation.

For metros and people movers, the UITP [24] has established 5 levels of automation. That means, the picture is not black and white, knowing either manual or automated driving. Automation is a stepwise process. The following five levels are established, UITP [24]:

GoA 0 is on-sight train operation, similar to a tram running in street traffic.(No automation at all)

GoA 1 is manual train operation where a train driver controls starting and stopping, operation of doors and handling of emergencies or sudden diversions.

GoA 2 is semi-automatic train operation (STO) where starting and stopping is automated, but a driver operates the doors, drives the train if needed and handles emergencies. Many ATO systems are GoA 2.

GoA 3 is driverless train operation (DTO) where starting and stopping are automated but a train attendant operates the doors and drives the train in case of emergencies.

GoA 4 is unattended train operation (UTO) where starting and stopping, operation of doors and handling of emergencies are fully automated without any on-train staff.

As a conclusion, automatic metros and automatic people movers can be seen as established systems. However, one needs to note that they operate in a controlled and simplified environment.

2.2. Road vehicles

We need to distinguish two situations:

a) driving on an open road and

b) driving on private territory

Without going into details we must be aware of the fact that for driving on an open road, the Convention requires a driver to be always present which is implemented in the national law of almost all countries. For driving on private territory, the traffic law is not applicable – the car would be a moving machine. Nevertheless, also here, safety requirements have to be obeyed. This type of vehicles is known as Automated Guided Vehicles (AGVs) and is becoming more and more popular.

The general impression on how autonomous driving works is mainly dominated by vehicles as the Google[14] vehicle or the Tesla[9] and other systems that have shown up in the meanwhile. Simpler systems are those for automated parking, which is carried out using the mobile phone, the driver being outside. Studies for autonomous driving have been carried out with a driver on board for testing purposes or for demonstration. Automated Guided Vehicle on closed areas or transport systems in workshops are also applied. The latter systems are strictly speaking not road vehicles but moving machines.

As an example, just consider the Google vehicle [14]. This is a Smart-like vehicle with two seats and one can read that it drives autonomously, with no driver action being necessary.

Alas, an accident has been reported and Google said it bears "some responsibility" after the car struck the municipal bus in Mountain View, Google [14]. That means that the Google vehicle caused a crash. In that case, the car would be responsible, i.e. finally its manufacturer. However, also the driver and his responsibility need to be discussed.

Another example is a Tesla vehicle [9] that crashed into a trailer. The driver did not react since he relied on automated driving and died as a consequence of the crash. In fact, the technical driving system of the Tesla was not able to detect the trailer. Then the question arises on the responsibility for the accident. Surely, the automatic systems needed permanent supervision by the driver and the question arises whether the driver was sufficiently instructed. Also, it needs to be discussed whether the driver had the possibility to stop the vehicle or take over the steer. This includes reaction time as well as features of the technical systems.

By the SAE [22] and the UN [24] the following levels have been defined.

• 0 No automation
• 1 Driver assistance
• 2 Partial automation
• 3 Conditional automation
• 4 High automation
• 5 Full automation

Table 1. Overview of automation levels [22]

SAE level	Name	Narrative definition	Execution of Steering and Acceleration / Deceleration	Monitoring of Diving Equipment	Fallback Performance of Dynamic Driving Task	System capability (Driving Modes)
Human driver monitors the driving environment						
0	No automation	The full-time performance by the human driver of all aspects of the dynamic driving task, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
1	Driver assistance	The driving-mode specific execution by a driver assistance system of either steering or acceleration / deceleration using information about the driving environment and with expectation that the human driver performs all remaining aspects of the dynamic driving task	Human driver and system	Human driver	Human driver	Some driving modes
2	Partial automation	The driving mode-specific execution by one or more driver assistance systems of both steering and acceleration / deceleration using information about the driving environment and with the expectation that the human driver performs all remaining aspects of the dynamic driving task	System	Human driver	Human driver	Some driving modes
Automated driving system ("system") monitors the driving equipment						
3	Conditional automation	The driving mode-specific execution by an automated driving system of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene	System	System	Human driver	Some driving modes
4	High automation	The driving mode-specific execution by an automated driving system of all aspects of the dynamic driving task, even if the human driver does not respond appropriately to a request to intervene	System	System	System	Some driving modes
5	Full automation	The full-time performance by an automated driving system of all aspects of the dynamic driving task under all roadway and environmental conditions that can be managed by a human driver.	System	System	System	All driving modes

Detailed information on the levels is shown on the following table 1.

The currently present systems are mainly systems for assisted driving. The assistant helps in simple situations, however, the driver has always full responsibility. Examples are

- Distance assistant,
- Platooning,
- Lane assistant,
- Highway pilot for trucks.

A short glance on the approval systems shows the differences:

- Automated metros are assessed according to EN 50126 [6], EN 50128 [7], EN 50129 [8] and approved based on local laws on metros, that differ per country,
- Road vehicles are approved by a European approval based on ECE rules. In Germany this institution for approval is the KBA, in Netherlands this is the RDW,
- AGVs are not road vehicle and not a train, they are considered as automated machines and approval is according to Machine directive [19] and IEC 61508 [16].

A new law for homologation of road vehicles in Germany allows automated driving in specific cases – note that this is not assisted driving – but driver must be able to overrule the technical system.

This is in line with Convention on Road Traffic [3], which says:

- article 8, 1: "Every moving vehicle or combination of vehicles shall have a driver",
- article 8, 3: „ Every driver shall possess the necessary physical and mental ability and be in a fit physical and mental condition to drive.",
- article 8, 5. "Every driver shall at all times be able to control his vehicle or to guide his animals."

Currently, these principles are implemented in the laws of the countries.

From this discussion we can conclude that experience and also safety principles from automatic metros cannot be directly used for road vehicles. First, the legal situation is different, second, there are differences regarding the applicable standards and third, the environment is different. An automatic metro is located in a controlled and well-defined environment that makes automatic driving possible. Passengers are separated from moving systems, e.g. by using platform screen doors that allow access only directly into the train. This does not hold in the general situation for a road vehicle.

3. General Safety principles and safety integrity levels

In this chapter we will briefly remember the main safety principles, see Gülker & Schäbe [15] and Gayen & Schäbe [11,12] and Gräfling & Schäbe [13] and give a short review on safety integrity levels.

Fail safe: If the system has a safe stopping state, i.e. a safe state in which it is not operational and this state is stable which can be reached fast enough, then the fail safe principle can be applied. It means that a system is brought into this sate if a failure occurs which cannot be tolerated. This principle can be implemented as inherent fail-safety, reactive fail-safety or composite fail-safety.

Safe life (fail operational): If the system does not have a safe stopping state which can reached fast enough, then the safety function has to be ensured. This is mainly done by using redundancies.

Fail silent: The fail silent principle is applied to a function the loss of which is tolerable since it is either an assistance function or the function is implemented in several instances. Then, failure of the function must be such that there is no repercussion on the safe functioning of the system. That means, that a fail-silent system must detect its failures and possible dangerous states and switch itself off without influencing other systems in a dangerous way.

Whenever a function might lead to harm, i.e. injury of fatalities to persons, material damage, damage to the environment, functional safety has to be applied. That means that the risk arising from a possible functional failure must be reduced to an acceptable level.

For this sake, safety integrity levels are defined. According to ISO 26262 [17] this can be QM, ASIL A to ASIL D with ASIL D being the most severe. IEC 61508 [16], which knows the safety integrity levels 1 to 4. is applicable for moving machines.

In practice this means, that for all driving functions and all driving sub-systems, the necessary safety level (ASIL or SIL) has to be determined using a risk analysis.

A safe life system is a system, in contrast to a fail-safe system, does not switch itself off in case of a failure, but where the safety function is ensured even in case of one (or sometimes several) failures.

The safety integrity levels (SIL / ASIL) are defined in standards for functional safety. IEC 61508 and EN 50129 define SIL 1 to SIL 4. ISO 26262 [17] defines the automotive SIL (ASIL) A to D.

The SIL / ASIL consists of two essential requirements:

- Maximum tolerable rate of dangerous failure which cannot be exceeded
- Measures against systematic failures (verification, traceability of requirement, specific techniques)

4. Abstract Model of the System

Lotz [18] proposed an architecture consisting of three levels: a navigational level, a manoeuvring level and a controller level. We will try to discuss a model that is as simple as possible.

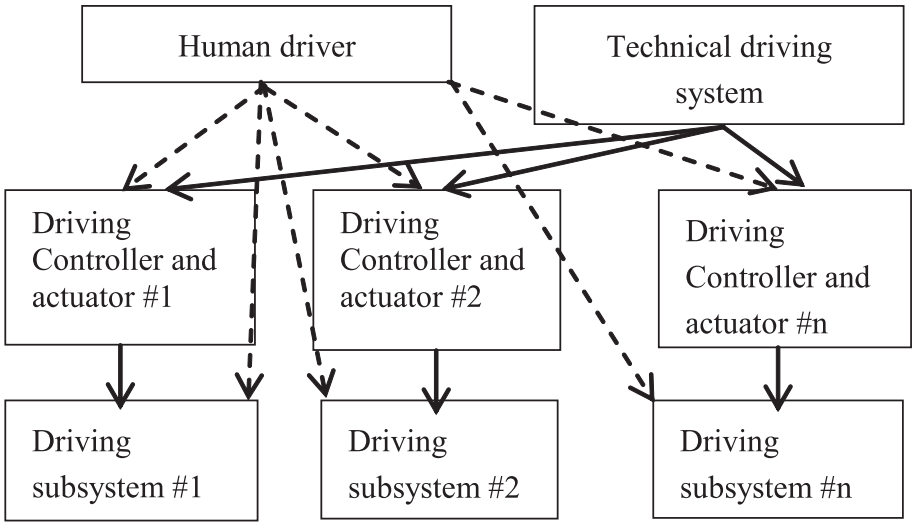

Figure 1. Scheme of a vehicle with automatic driving capabilities

For systems that drive automatically, partially automatically or autonomously, we will use the following very simple structure for the system. In fact, this system must be equipped not only with a human driver, but also with a technical driving system, that carries out the driving.

The vehicle consists of driving sub-systems as steering, braking, acceleration systems etc. in a very abstract manner. These sub-systems could be even very simple systems as pure mechanical steering system, pneumatic brake systems etc. The driving is carried out by the human driver using these driving sub-systems directly.

The manoeuvring and navigational level according to Lotz [18] have here been combined in one system (human driver / technical driving system).

If a vehicle shall be operated by a technical system which does the driving in place of the human driver or supports the human driver, then this system must have access to the driving sub-systems. This is possible only using a driving controller and actuator. That means that these types of systems must be present in the vehicle to allow for driving by a technical system.

Then, this allows also the human driver to access the driving sub-systems via the driving controllers.

Hence, there are different possibilities to operate these subsystems.

a. The driver can directly access the driving sub-systems, e.g. the steering wheel is mechanically connected to the steered axle.

b. The driver accesses the driving sub-system via a controller and an actuator which operate the sub-system electronically. A typical example for such a system is an electric parking brake.

c. The technical driving system accesses the driving subsystem via a controller and actuator

Discussing figure 1 it becomes clear that arbitration between the commands of the human driver and the technical driving system must take place.

There are different levels on which arbitration can take place:

a) driving subsystems

In this case the force applied by the driving controller and actuator must be so small that the driver can always overrule without a problem. However, he would be either required to switch off the driving controller an actuator manually, or those system need to have an in-built function to detect the interference of the driver and switch themselves off.

b) driving controller and actuator

Here, the driving controller has two inputs with different priority. The high priority input is used by the driver, the low priority input by the technical driving system. The arbitration is done by the driving controller which detects overruling by the human driver and switches off the input from the technical driving system. Many controllers in modern cars (brake controller, steering controller etc.) have an additional input for assistance systems which just fulfils this requirement. This approach assumes that the human driver himself controls the vehicle via x-by-wire via the relevant controllers.

c) technical driving system

Arbitration is between the human driver and the technical driving system. If the human driver overrules the technical driving system the latter does not generate its own control signals but simply transfers the signals of the human driver to the driving controllers.

The choice on one of the approaches is a choice of the manufacturer of the vehicle. However, this choice influences the suppliers of the driving controllers and actuators. They need to implement different architectures in their controllers.

In case a) they need to detect intervention of the man driver and deactivate the actuator.

In case b) they need to have two inputs with different priority and need to carry out arbitration

In case c) only one input is necessary and no arbitration is necessary.

We see that x-by-wire is a necessary precondition for solutions b) and c).

We will guide ourselves by the requirements for a fully autonomous driving vehicle with a possibility for the human driver to take over responsibility at any time.

5. Level analysis

5.1. Levels 0 and 1

In this section we will analyse the levels (SAE [22]) of automatisation and draw conclusions for the safety architecture of a vehicle.

In levels 0-1 execution of steering and acceleration and deceleration is in the responsibility of the human driver, the driver is responsible for monitoring and the technical system is able to support some driving modes (level 1).

That means, the human driver is doing the driving and the technical driving system can only add some supporting functionality as warn the driver or react in cases, when he is not able to react (emergency brake assistant). This means that the technical driving system must be fail silent, i.e. upon failure of this system the driving behaviour of the vehicle must not be influenced or only influenced in such a manner that safe driving is still possible. The driver should be warned, if such an assistance system fails to work.

5.2. Level 2

In automation level 2, the system takes responsibility in some driving modes. The human driver monitors the technical driving system and he is the fall back solution. That means, that all technical systems are pure assistance systems and that

R1) The driver must have the technical possibility to interfere, i.e. to override the technical systems. That means, that each controller for acceleration, braking and steering that receives signals from the human driver and from the technical driving system must have a voter which always gives priority to the driver. In fact this means that an electronic control system needs to be present for these function that has an ASIL that coincides with the function, mainly this would be ASIL D. This control system then must have a priority input for the driver and another non-priority input for the technical driving system. The relevant driving controller must detect, when the driver wants to override the technical driving system and has to carry out the required reaction.

R2) The driver must have enough time to detect wrong or faulty behaviour of the technical driving system and react and be able to bring the vehicle back to a safe driving state. That means, that the controllers have to limit the influence of the technical driving system, e.g. limit the level of acceleration, deceleration and the steering angles or angular speed and angular acceleration and jerks so that the driver always has the time to react.

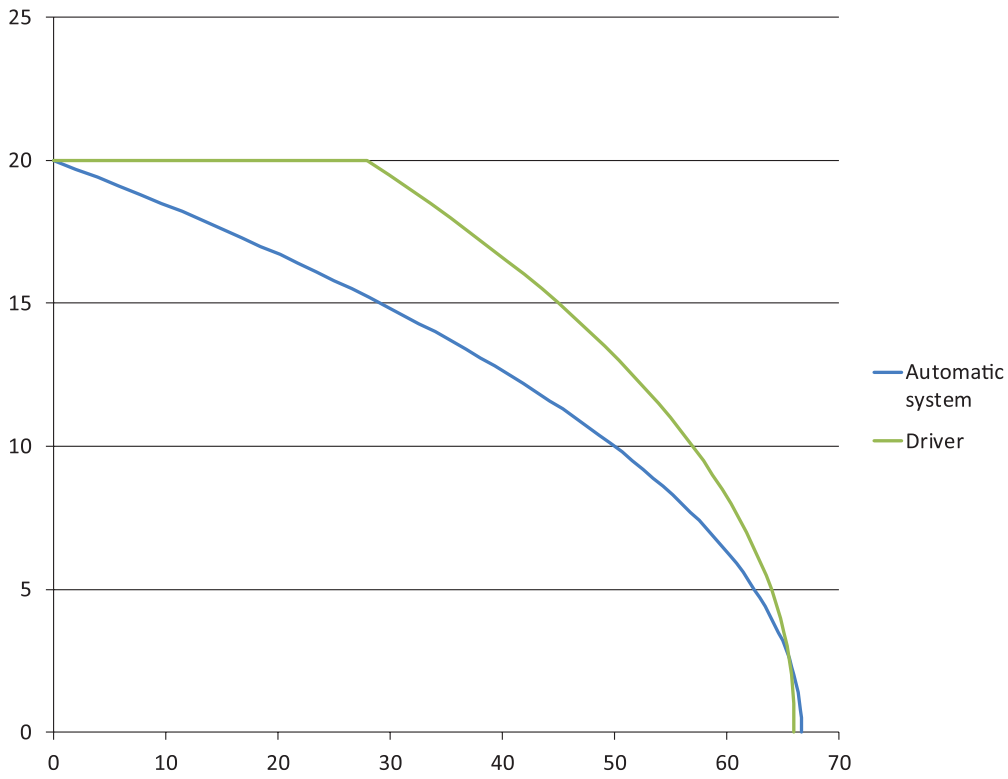

Figure 2. Example for a brake curve.

Moreover, the driver must be trained for this function or the controllers must be designed in such a manner that they give enough time for reaction for any driver.

Requirement R2 leads to the following requirements for automatic driving.

• Braking: braking by automatic systems must be with a smaller acceleration than the driver could apply, the difference in accelerations (vehicle, driver) must still allow for a reaction time of the driver (braking curves),

• Steering: the distance from dangerous objects (other vehicles, border of the lane etc.) must be large enough to allow for drivers reaction, together with a limit of the steering angle. This might lead to speed restrictions.

• Perhaps the driver needs special training.

Figure 2 shows an example of a brake curve. Speed (m/s) versus distance is shown. There are two curves, one for automatic braking (deceleration 3 m/s^2) versus braking by driver (5 m/s^2), where a reaction time of 1.3 s has been taken into account for the driver. The initial speed is 20 m/s.

In this example, the driver is still able to come to a standstill in time, if he detects that the automatic system fails to brake. Of course, the driver must react and be able to react with 1.3 s.

For steering, similar requirements must be taken into account: Driver must have necessary reaction time. This reaction time depends on the distance to shoulder or adjacent lane, the speed and the reaction of the system. The latter includes maximal angular velocities and accelerations with which the system might show a faulty reaction.

The current technical solutions are supported by the following existing equipment:

• Different controllers or safe computers are available that are qualified according to up to ASIL D / SIL 4,

• Sometimes even „intelligent sensors“ with a SIL available.

• Different, diverse sensors (no SIL), which are cross-validated by the safe computer. Examples of such sensors are cameras, lasers, radar, infrared, ultrasonic etc.

• Multiple, diverse actors; safety relays as electric actors, the use of proven mechanical systems is also possible.

5.3. Level 3

Level 3 differs from level 2 in just one point. The technical driving system is responsible for monitoring of the driving equipment. That means that the system must diagnose itself and the environment in order to decide whether it can go on with driving or whether the human driver must act as a fall back solution. The following questions are important

R3) A clear handshake must be defined between human driver and technical driving system. Either the technical driving system must go on with functioning until the human driver has accepted to take over control or

R4) A certain time of e.g. one second is foreseen for the human driver to take over control at any time, if the technical driving system asks him to do so.

In the first case, the technical driving must be safe life, in the second case the latency time for the human driver to take over must be ensured by technical systems – either by the safe life property or just by the driving situations and speed. Timing considerations can be found in Vogelpohl et al. [25].

5.4. Levels 4 and 5

Levels 4 and 5 are even more advanced. The difference between levels 4 and 5 is relatively small, since the distribution of responsibilities is the same, only in level 4 some driving modes are excluded, which allows the technical driving system to have limited capabilities. However, when this system is active, it must be able to take full responsibility.

As a consequence, the technical driving system must always ensure safe driving and would need to be safe life.

The relevant requirements are derived the so called GAMÈ principle, which can be found e.g. in EN 50126 [6] “All new guided transport systems must offer a level of risk globally at least as good as the one offered by any equivalent existing system“. Here we apply the phrase on guided transport system just to an autonomously driving vehicle. We compare the classical vehicle with a human driver with an autonomously driving vehicle. Then, there are two aspects to be considered:

a) Performance and

b) The technical system (vehicle and technical driving system) are sufficiently free from dangerous failures.

Both aspects are considered separately. For performance, the technical driving system must be at least as good as a human driver in the relevant driving situations, see Mazzega et al. [20]. If this cannot be ensured for all driving situations, the set of relevant driving situations must be limited and the human driver must handle the most complex ones.

The second, the safety aspect, can be handled as for any technical system by defining an appropriate safety level (ASIL or SIL). This leads to

R5) The performance of the technical driving system as reaction time, detection and handling of traffic situations etc. with an un-failed system must be at least as good as that of a human driver.

R6) The technical driving system must be developed according to a reasonable SIL / ASIL.

For level 4, a clear handshake must be defined how to pass over responsibility between technical driving system and human driver. Especially, the driving modes must be defined, where the technical driving system must not be used for reasons of e.g. insufficient performance. Handshake must be carried out either during standstill or the technical driving system must early enough inform the driver that it wants to pass control to

the driver and the driver must take responsibility. If the driver does not take over, the technical driving system must still have the possibility to stop the vehicles as long as it is in a driving mode, where automatic driving is allowed and possible.

If the driver passes responsibility to the technical driving system he must have responsibility until the technical driving system informs him that it has taken over responsibility.

When driving on an open road, the driver must be in full responsibility of the driving behavior of the vehicle, see the Convention [3]. Then, even if the technical driving system is able to perform up to SAE level 5 with the necessary safety integrity, the driver must have the possibility to intervene. So, the requirements under a) and b) mentioned for SAE level 2 hold if driving on an open road.

Autonomous driving, i.e. driving without intervention of a human driver is in fact only realized in SAE level 4 (partially) or 5 (completely). This holds even if the laws require a driver to be present.

6. Implementation of safety principles

6.1. Assistance systems

It is clear that for technical driving systems in levels up to 2 the systems must and can be fail silent and R1 and R2 must be fulfilled to ensure that the driver has the possibility to take over control.

6.2. Application of the fail-safe principle

First of all, we need to determine whether there exists a safe stopping state that can be reached sufficiently quick. Assume the velocity of the vehicle is limited to a value v, the braking deceleration is a and the reaction time t then the vehicle will stop within a distance of

$$s = v \cdot t + v/(2a).$$

Assuming that the steering has no limitation, stopping the vehicle will be a safe action if there is no obstacle within a distance of s from the outer boundaries of the. This area can be made even smaller taking into account that

- actual direction of steering and the (physical) limitations of changes of the steering angle and
- physical limitations for changing the driving direction.

In such case, the technical driving system and the driving controllers could be a complete fail safe system, stopping the vehicle in case that a failure is detected.

Depending on the free space around, the vehicle speed is determined. Obviously, the less free space available, the slower the vehicle must drive. Driving controllers need to be developed and implemented according to an adequate SIL / ASIL, which depends on the speed of the system.

6.3. Application of the safe life principle

If the vehicle is intended to move faster, the technical driving system and the driving controllers must be safe life, at least as long as the vehicle is in motion.

Driving controllers need to be developed and implemented according to an adequate SIL. This is for the brake (ABS / ESP) mainly ASIL D, for the steering ASIL B…ASIL D, depending on the function of this controller. With such a choice most of the vehicles can perform with velocities up to 250 km/h.

The implementation using safety principles differs whether we are talking on a road vehicle or a moving machine. In the first case the environment cannot be assumed to be under control, in the second case this can be ensured since the technical driving system acts on private territory. In this latter case it is much easier to ensure enough free space.

From this consideration it becomes also clear, that not all functions must be always implemented with the highest SIL / ASIL. This depends very much on the speed and the environment. If speed is limited by physical or other means, then also a lower SIL or ASIL can be used. In any case this needs to be shown by the risk analysis that has to be accrued out based on ISO 26262 [17] or IEC 61508 [16].

The following functions are the main functions to be considered:

- Guidance

How to implement such a function including the steering is described in Bouwman, Schäbe & Vis [1]. Mostly the steering of the axles needs to be safe life and a safe computer has to be used in the technical driving system to determine the steering angles. Another important function is determination of the location, where differential GPS, maps together with ultrasonic sensors, radar or lasers or cameras or different types of marking placed physical on the lane of the vehicle can be used. The safe computer will determine the real location and compare this with the assumed location as a result of its steering activities and correct or stop he vehicle.

- Braking and acceleration

Assuming that the vehicle moves along the desired trajectory, the vehicle needs to start, move and stop. So the vehicle needs to react to these commands. It is important to limit the speed e.g. in curves or at narrow places and to be able to perform an emergency stop, if parts of the system fail. In order to perform this function, the system needs to know the location.

Solely with these two functions the vehicle would move without taking into account the environment. Any change in the environment could lead to a collision or the vehicle leaving its track.

- Reaction to unforeseen events (obstacle)

The vehicle must be able to detect obstacles. By an obstacle we denote any object that is in the (planned) or near the (planned) trajectory of the vehicle. We need

to distinguish fixed obstacles and moving obstacles. In the beginning we consider as only strategy of the vehicle to stop in front of the obstacle. Moving around the obstacle will be considered later together with moving obstacles

a) Stationery obstacle: To detect the technical driving system needs to have a blueprint of the environment and needs to compare the real environment with that blueprint and detect differences. This would require certain algorithms for detection and classification of objects. Note that "detection" and "blueprint" does not mean that the technical driving system uses optical means. It can be optical means, but also others or in combination.

In a first step the obstacle as such needs to be detected. This is possible only at a certain distance and takes a certain time. This performance of the system might limit the speed, since the vehicle must always come to a standstill in front of the object.

In a second step the technical driving system can classify the obstacles as small. Note that this classification can be present implicitly if the technical driving system will not detect obstacles of small size. Such a classification is always present due to limitations of the system.

If the obstacle is small enough and not tall, the vehicle might decide to go on with driving.

b) Moving obstacles: Moving obstacles must be traced and its motion must be predicted using the actual position and speed. It must also be taken into account whether the object can accelerate or decelerate or change its motion direction. The latter factors strongly depend on the nature of the object. E.g. a motorbike can reach other acceleration values as a pedestrian. In order to provide a good prediction, the technical driving system must cluster moving objects according to their capability of motion. Consequently, for each object of the different clusters future positions must be predicted and the technical driving system must define the motion of the vehicle in such a manner that collisions are avoided. This might lead to the decision to stop or to keep the present fixed position.

Depending on the performance of the clustering and prediction algorithms, the technical driving system would behave more or less conservatively. With better algorithms the technical driving system would stop less frequently. We remind that the performance of these algorithms together with the stopping process in case of doubts about the future trajectory of the obstacle must be as least as good as that of a human driver. This includes of course strategies to drive around an obstacle.

c) Stationery obstacles that could start moving are in fact a combination of cases a) and b) discussed above. This means, that the technical driving system must not only trace moving obstacles but must also be able to classify stationery obstacles and provide a judgement on whether they might move and with which velocity and in which direction. A most safe strategy would surely be to stop at a safe distance of any unknown object.

If a proper reaction of the vehicle cannot be ensured for all driving situations, the set of relevant driving situations must be limited and the human driver must handle the more complex ones. This would lead to an SAE level 4 situation. An example would be a strategy, where the technical driving system takes over control on a motorway and the human driver in the city.

7. Problems

In connection with autonomous driving some problems appear. We will, discuss only some of them and try to describe possible technical solutions.

a) Assume an autonomous vehicle cannot prevent an accident and needs to make a choice, e.g. between material damage, environmental damage and injury or – even worse – injuring or even killing either an older or younger person, another driver, the own passengers etc., see e.g. EK [5] (Ethic commission)

This type of discussions automatically comes up when the responsibility for driving is carried over from the human driver to a technical driving system. The ethic problem that is behind this discussion cannot be solved in this paragraph. It is obvious that a technical solution to this problem would require to distinguish between persons and objects or animals, to discriminate between different persons etc. This would require rather complex algorithms, if it is feasible at all.

The simplest solution to the problem is to apply the principle of driving on sight. That means the rule for the autonomous vehicle would be to drive only with such a velocity that it can stop before each obstacle that appears on the road. This requirement covers:

- Detection of any obstacle above a certain size,
- Prediction of movement of objects (which is the most complicated part),
- Reducing speed if necessary to come to a standstill before such an obstacle.

Based on such a "safety first" approach, later on objects of certain (small) size can be neglected to ensure performance and avoid the vehicle stopping in front of a leaf or a plastic bag.

b) Additional information

A vehicle might optionally use additional information provided by the infrastructure, which might lead to better performance regarding safety.

Let us consider the following example. The vehicle uses information from cameras mounted on the street and has the possibility to "look around the corner". Then, it could e.g. detect a suddenly appearing child running out of the house, what a human driver could not.

c) Safety targets

Since the target of autonomous driving behaviour would always be the performance of a human driver, the technical driving system would have to fulfil this important requirement. However, assume that autonomous systems will set a new target in the future – then the question

will arise: Does the driver have the right to switch the automatic system off and decrease the achieved level of safety? It would be somehow equivalent to a train driver switching off automatic train protection, e.g. to use some speed margins. This simple example shows that the way to autonomous driving would be a one-way street, with no return to manual driving at the end.

8. Possible next steps

Based on the current status one can imagine the following future steps for road vehicle.

• Safe guidance (lane keeping) could be implemented, e.g. using differential GPS together with good update service of precise maps. All work on the road and all temporarily blocked roads need to be present on these maps.

• Stopping before traffic lights enforced by a wireless transmission of information between traffic lights and vehicles. Nevertheless, the driver needs to watch out for violators, e.g. cyclists even if he has a green lights.

• Speed limit enforcement, e.g. the speed limit is transmitted in a wireless manner form a sign broadcasting the speed limit or the sign is read by a camera, alternatively a map is used as source.

• Handling of simple traffic situations as e.g. on motorways following the lane, without overtaking manoeuvres.

• Vehicles on separated areas and on separated road networks.

Further development might lead to a following scenario, which include:

• The road or lane might be separated by two fences forming a controlled environment and on this environment a vehicle can run automatically, with steering, braking, driving implemented according to ASIL D.

• Vehicles drive with very short distances using platooning.

• At certain places entry and exit to this network of roads is allowed. There, the driver takes over the automatic vehicle and drives it manually to the destination.

• The necessary information as maps, position, speed limits, communication with other automated vehicles would be implemented on the vehicle, rather than on the road.

• The infrastructure would be rather cheap, consisting of the road and fences. Comparing this with a railway, the infrastructure is more flexible, no signals, no switches, no ballast and sleepers are necessary.

In all these cases, the relevant technical systems would need to be safe life systems with a safety level up to ASIL D / SIL 4.

Regarding future development, also possible problems need to be considered, that an automatic or autonomous vehicle driving on the road need to face to become comparable with a human driver. First of all, such a system needs to distinguish objects as persons or animals from unmoving objects. Another example would be to distinguish vehicles on high wheel from bridges etc. Another problem is that sometimes intentions of a person or animal need to be guessed: does the person or the animal intend to cross the road and step on the road? A typical example would be a child with a ball standing on the sidewalk, having dropped the ball and this has moved on the street. There are a lot of such tasks would require intelligence and one would tend to use artificial intelligence for such a task.

Assume now that artificial intelligence should be implemented for autonomous driving. Then requirements for SIL 4 / ASIL D would need to be implemented in full rigor in the software and the hardware. On the other hand, the algorithms for artificial intelligence are voluminous and complex. If then e.g. traceability needs to be shown from a requirement as e.g. “The algorithm must distinguish human beings from other objects” one might imagine the complexity of such a task. This would only be one requirement. The entire complex of requirements to the software would have to take into account a lot of driving situations, in the environment etc. If the algorithm is a self learning algorithm, one needs to ensure that it has learned in a certain time enough and this must be proven in the light of the standards IEC 61508 [16] and / or ISO 26262 [17]. Another possibility would be to use a proven in use argument and accumulated $3\ 10^9$ hours in service, see IEC 61508 [16] part 7 annex D. With 600 hours of driving per year that would mean to have 5 000 000 vehicles driving an entire year under controlled circumstances, i.e. with trained drivers that can override the system and that would also register all events – or the vehicle has to do this. One can decrease the number of vehicles by increasing the number of driving hours per year, e.g. up to 6,000, which would mean driving in shifts. Nevertheless, still 500,000 vehicles would be necessary. In addition, each change of the software would require to repeat this approval process

The conclusions is that solutions for the safety relevant software must be simpler, without guessing intentions etc. in order to overcome these problems. Artificial intelligence would be good for assistance systems.

9. Conclusions

In this paper we have provided some considerations on automatic (or autonomous) driving for rail and road vehicles. It turns out that for road vehicles, the environment is much more complex than for rail vehicles. Therefore, the experience from e.g. automatic metros cannot be directly used.

In this paper we have presented some ideas on possible safety architecture for autonomous driving, deduced from known safety principles and from general requirements. We have analysed the SAE levels and the implication for the safety architecture per level.

Possible implementation principles have been described and specific problems of autonomous driving have been discussed. So, it is recommended to follow the design principles as described in chapter 6 for the implementation of autonomous driving systems. It is important to understand the safety architecture of the vehicle and to find out, whether it is a pure assistance system (fail-silent), whether the fail-safe principle is applied or the safe-life principle need to be applied. The guidance of this principles should be used for safety assessment of autonomously driving vehicles.

Most of the existing systems are either pure assistance systems or they are dedicated to simplified traffic situations

It has to be expected that the first safe solutions for autonomous driving would come for situations with a simplified environment, especially where the environment is controlled or even adapted to the task of autonomous driving. Here, a special solutions are AGV (automatic guided vehicles) that are just moving in an environment fully adapted to them, but not on an open road.

References

[1] Bouwman, R., Schäbe, H., Vis, H. (2009), Application of safety principles for a guidance system in public transport, *ESREL 2009, Proceedings Reliability, Risk and Safety*, vol. 3, p. 2275-2278.

[2] Breitinger M. (2016), Kabinett erlaubt teilautomatisiertes Fahren, http://www.zeit.de/mobilitaet/2016-04/autonomes-fahren-gesetzentwurf-verkehrsrecht-alexander-dobrindt, published 13.4.2016, retrieved on 19.10.2017

[3] Convention (1973), *Convention on Road Traffic*, 8.11.1968, European Additional Treaty from 1.5.1071 and Protocol 1.3.1973.

[4] Daimler 2017 The Mercedes-Benz Future Bus The future of mobility, *https://www.daimler.com/innovation/autonomous-driving/future-bus.html*, retrieved on 19.10.2017

[5] EK 2017, *ETHIK-KOMMISSION AUTOMATISIERTES UND VERNETZTES FAHRE*N (Ethics Commission for automated and networked driving, in German), Bericht, Juni 2017, WWW.BMVDI.DE

[6] EN 50126 *Railway applications – The specification and demonstration of Reliability, Availability, Maintainability and Safety* (RAMS) (EN 50126), 1999

[7] EN 50128 Railway applications — Communication, signaling and processing systems — Software for railway control and protection systems, 2011, correction 2014.

[8] EN 50129 Railway applications – Communication, signalling and processing systems – Safety related electronic systems for signalling, 2003

[9] Focus (2016) Todesfall im selbstfahrenden E-AutoUS-Verkehrsaufsicht prüft Teslas "Autopilot", *http://www.focus.de/auto/elektroauto/todesfall-im-selbstfahrenden-auto-us-verkehrsaufsicht-prueft-teslas-autopilot_id_5687341.html*, 1.7.2016

[10] Frog 2017, Website, www.frog.nl, retrieved on 19.10.2017

[11] Gayen, J.-T., Schäbe, H. (Miss-) Konzeptionen von Sicherheitsprinzipien, *Signal und Draht*, 100 Nr. 7+8 (2008) pp. 11-18.

[12] Gayen, J.-T. , Schäbe, H. (Mis-) conceptions of safety principles, *ESREL 2008, Proceedings Safety, Reliability and Risk analysis*, vol. 2, pp. 1283-1291

[13] Gräfling, S., Schäbe, H., The agri-motive safety performance integrity level – or how do you call it?, *ESREL 2012 / PSAM 11*, paper 26 Fr2_1, 10 p..

[14] Google car (2016) Google self-driving car hits public bus near Mountain View headquarters *http://www.mercurynews.com/2016/02/29/google-self-driving-car-hits-public-bus-near-mountain-view-headquarters/*, retrieved on 19.10.2017.

[15] Gülker, J., Schäbe, H., 2006, Physical Principles of Safety, *Safety and Reliability for Managing Risk, Proc. of ESREL 2006*, pp. 1045-1050.

[16] IEC 61508 *Functional safety of electrical / electronic / programmable electronic safety-related systems*, 2010, parts 1-7,

[17] ISO 26262 Road vehicles — Functional safety, 2018, parts 1-10,

[18] Lotz, G.O. 2017, *Eine Referenzarchitektur für die assistierte und automatisierte Fahrzeugführung mit Fahrereinbindung*, Dissertation Technical University Darmstadt, 2017 (A reference architecture for assisted and automatic driving with driver intervention),

[19] Machine Directive (2006) DIRECTIVE 2006/42/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast)

[20] Mazzega, J., Köster, F., Lemmer, K.,. Form, T., *Absicherung hochautomatisierter Fahrfunktionen*, Automobiltechnische Zeitschrift, 118 (2016), no. 10, 48-52 (Safe Implementation of Highly automated Driving Functions)

[21] Nahverkehrspraxis (2017), Weltpremiere: Daimler Buses präsentiert autonom fahrenden Stadtbus", *http://www.nahverkehrs-praxis.de/news/nahverkehr-spraxis-top-news/article/weltpremiere-daimler-buses-praesentiert-autonom-fahrenden-stadtbus/*, retrieved on 19.10.2017

[22] SAE (2016) Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, *SAE J3016*, September 2016.

[23] UITP 2017, *International Association of Public Transport*. "A global bid for automation: UITP Observatory of Automated Metros confirms sustained growth rates for the coming years". *Belgium*, retrieved 19.10.2017

[24] UN (2017) *Economic Commission for Europe, Inland Transport Committee, World Forum for Harmonization of Vehicle Regulations, Consolidated Resolu-*

tion on the Construction of Vehicles, (R.E.3), Revision 6, 11.7.2017

[25] Vogelpohl, T., Vollrath, M., Kühn, M. Hummel, T. Gehlert, T., (2016), *Übergabe von hochautomatisiertem Fahren zu manueller Steuerung*, Forschungsbericht Nr. 39, Unfallforschung der Versicherer GDV, August 2016, ISBN 978-3-939163-67-1

[26] Short English Version:

[27] Vogelpohl, T., Vollrath, M. (2016) UDV (Unfallforschung der Versicherer) *Takeover times in highly automated driving Compact accident research*, Nr.57, 07/2016

[28] Wachenfeld, H. K. (2016), *How Stochastic can Help to Introduce Automated Driving*, Dissertation, Technical University Darmstadt, 19.10.2016

About the author

Hendrik Schäbe, Dr. rer. nat. habil., Chief Expert on Reliability, Operational Availability, Maintainability and Safety, TÜV Rheinland InterTraffic, Cologne, Germany, e-mail: schaebe@de.tuv.com

Received on: 18.03.2019