

AN OPTIMIZATION-BASED METHOD TO IDENTIFY RELEVANT SCENARIOS FOR TYPE APPROVAL OF AUTOMATED VEHICLES

Thomas Ponn

Frank Diermeyer

Technical University of Munich
Germany

Christian Gnant

TÜV SÜD Auto Service GmbH
Germany

Paper Number 19-0024

ABSTRACT

The objective of this paper is to propose a novel approach for an intelligent selection of relevant scenarios for the certification of automated vehicles. During this process, two main challenges occur. Firstly, since the number of possible traffic situations is unlimited, a selection of a manageable number of representative situations to be tested must be applied during the certification of automated vehicles. Secondly, nowadays a limited number of standardized test cases are used for the type approval of vehicles. This can lead to so-called gaming of tests, which means that the manufacturer optimizes the system's performance in the predefined test cases. A prominent example are the current discussions about the large differences between the emissions of vehicles in the driving cycle (e.g., WLTP) and in everyday use in road traffic. This paper addresses both stated challenges and exemplifies a method for the system-specific selection of test cases for the certification of automated vehicles, which are not known to the manufacturer in advance. Based on a system analysis and an objective driving behavior characterization, weak spots of the system under test are identified and connected to complex scenarios to be tested. This approach allows an economic and meaningful certification process for automated vehicles.

INTRODUCTION

In recent years, intensive work has been carried out on the implementation of highly automated and autonomous vehicles (Level 3 and higher according to the SAE classification [1]). Many manufacturers already have prototypes of these vehicles, which are increasingly being tested in real traffic, especially in the USA. A high level of safety is indispensable for social acceptance - especially due to the danger posed to uninvolved parties in the event of system faults.

It is difficult to prove the safety of automated vehicles in an economically feasible manner, due to the open parameter space. For the type approval of vehicles, in particular, only a very limited scope can be tested. Therefore, an intelligent and manufacturer-unknown selection of the scenarios to be tested is necessary in order to avoid the so-called gaming of tests and to prove a sufficient safety performance. In addition, regulations and laws for the certification of automated vehicles are still missing.

In general, due to the infinite number of possible traffic situations, there is a need for an economically feasible method of performing safety assessments on automated vehicles. A promising approach is scenario-based testing. Based on the assumption that a large part of traffic situations is irrelevant and uncritical, scenario-based testing is limited to meaningful events (scenarios). The framework for this approach is being developed, for example, in the German funding project PEGASUS [2]. The challenge of selecting and finding all relevant scenarios remains with this approach. Since only an extremely limited number of tests can be carried out during certification, it is particularly important for this application to conduct an intelligent selection of scenarios. This contribution, therefore, presents a novel approach for a system-specific selection of relevant scenarios for the certification of automated vehicles. Furthermore, with this approach, the manufacturer can no longer perform so-called gaming of tests.

The article is structured as follows: First, an overview of existing literature is given, and the research objective is derived. Subsequently, the procedure developed is described in detail in the METHODOLOGY section. An exemplary derivation of results using the method presented is explained in the RESULTS section. Then, the approach will be critically discussed, and its limitations demonstrated. The paper concludes with a summary and an outlook on future work.

LITERATURE REVIEW AND RESEARCH OBJECTIVE

This section addresses in detail the challenges already raised in the introduction regarding the certification of automated vehicles. Finally, the state of the art is critically evaluated, and the main research question of this paper is derived.

Unlimited number of possible traffic situations

Due to the infinite number of possible traffic situations, the safety assessment of automated vehicles can no longer be carried out economically in road tests [3]. With the scenario-based approach, the level of effort required can be reduced considerably if possible traffic situations are restricted to relevant events. Irrelevant situations, such as driving in a straight line with no action taken by drivers in surrounding traffic are left out. Nevertheless, the challenging task of finding all relevant scenarios remains with this approach. Before we examine existing methods for the selection of these scenarios in detail, important terms are defined.

Definition of vocabulary

Object and Event Detection and Response (OEDR): Different factors influence the safety of automated vehicles. These are, for example, the human machine interface (HMI) as well as the functional safety and cyber security of the vehicle. In our approach, we focus on Object and Event Detection and Response (OEDR) according to NHTSA [4, p. 7]. OEDR includes the detection of objects, their classification, the planning of a suitable response to the detected object and the execution of the planned action.

Operational Design Domain (ODD): According to SAE [1], the ODD is defined by the area for which the automated vehicle was developed. The ODD can be restricted, for example, by road classes (e.g., highway or city center) or environmental conditions (e.g., weather conditions).

System Under Test (SUT): The automated vehicle to be tested and certified is denoted as system under test.

Traffic Participant (TP): All kinds of movable objects within a traffic situation are called traffic participants. Among others, this includes pedestrians, cyclists, motorcycles, passenger cars and trucks.

Scenario: ULBRICH [5] defines a scenario as:

[...] the temporal development between several scenes in a sequence of scenes. Every scenario starts with an initial scene. Actions & events as well as goals & values may be specified to characterize this temporal development in a scenario. Other than a scene, a scenario spans a certain amount of time.

Where a scene is defined as:

[...] a snapshot of the environment including the scenery and dynamic elements, as well as all actors' and observers' self-representations, and the relationships among those entities. Only a scene representation in a simulated world can be all-encompassing (objective scene, ground truth). In the real world it is incomplete, incorrect, uncertain, and from one or several observers' points of view (subjective scene).

Furthermore, MENZEL [6] distinguishes between three types of scenario – functional, logical and concrete scenarios. Functional scenarios represent a linguistic description of the scenario on a semantic level. The information content of this description is low. For logical scenarios, the parameters required to describe the scenario, such as the initial speed of the SUT or the lane widths, as well as their ranges are included. Concrete scenarios have the most information content. Starting from a logical scenario, a specific value is defined for each parameter in the concrete scenarios and is thus unambiguous. It should also be noted that the term test case is used here as a synonym for a concrete scenario, although in [7] the test case also includes pass-fail criteria. In the present use case, pass-fail criteria can be regarded as prescribed by future regulations for the type approval of automated vehicles.

Relevant scenarios: All scenarios that contribute to the type approval of automated vehicles are considered relevant. Relevant scenarios can also be very simple, such as the beginning of a speed limit. This is relevant for certification because an automated vehicle must comply with existing traffic regulations. This type of scenario is taken into account in the method developed when driving behavior is characterized. A subset of the relevant scenarios are critical and complex scenarios (Figure 1). These two subsets are defined below.

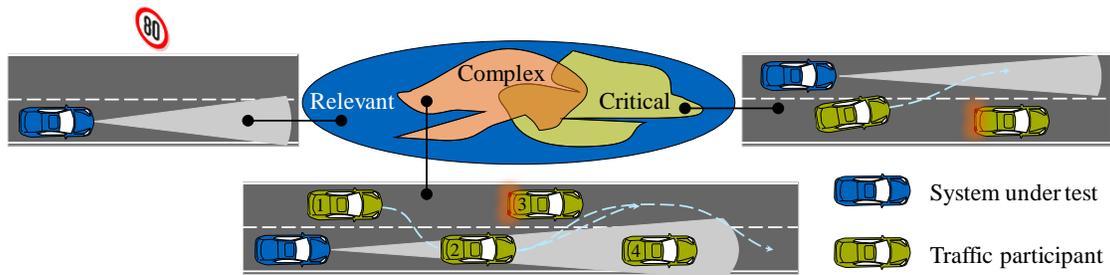


Figure 1. Definition of relevant, complex and critical scenarios

Complex scenarios: Complex scenarios are scenarios that present a challenge for the planning algorithm of the SUT. This is achieved by the presence and movement of other TPs. Complexity can thus be understood as how difficult it is for the planning algorithm to plan a safe trajectory under consideration of other TPs. Other influences such as the width of the road or the need to maneuver are not explicitly taken into account in this method.

Critical scenarios: Criticality is defined in this paper as the closeness to an accident. To measure criticality, indicators such as Time-To-Collision (TTC) [8] can be used. The smaller the value of this indicator, the more critical the scenario is. Critical scenarios can have two different causes. On the one hand, they can arise due to high differential speeds and small distances. This means that a logical scenario is not inherently critical. However, a critical concrete scenario can very easily be derived from any logical scenario if the distances between the objects and their velocities are adjusted accordingly. Thus, a cut-in situation is not automatically critical. But if a slow-moving TP changes into the SUT's lane at a short distance before it, it

is very critical. These types of critical scenarios can be defined without extensive analysis and will not, therefore, be considered further in this publication. On the other hand, critical situations can arise due to errors of the SUT. These can be errors and inaccuracies in perception and errors in the planning algorithm. The former is considered in the developed methodology in the analysis of the sensor setup, while the latter is addressed in the complexity of the traffic situations. It is assumed that increasing complexity increases the probability of an error in the planning algorithm. An example of a critical scenario due to an error in the SUT is a (complex) traffic situation in which the SUT incorrectly predicts the trajectory of a TP, resulting in an accident in the further progress of the scenario. Since this type of criticality is taken into account in the methodology by the analysis of the sensor setup used and also by the addition of complexity, the identification of critical scenarios is no longer explicitly discussed in the following.

In summary, it can be concluded that relevant scenarios may be very simple (e.g., speed limitation). In addition, critical scenarios are not automatically complex (e.g., accident involving an autonomous prototype in the USA [9]) and, on the other hand, not all complex scenarios are automatically critical (e.g., if the algorithm masters the scenario correctly). Nevertheless, all types of scenarios mentioned are relevant for the type approval of automated vehicles.

Five-layer model: To define the required parameters for the scenarios in a systematic manner, SCHULDT [10] introduces a four-layer model, which BAGSCHIK [11] extends to a five-layer model. This allows all relevant parameters for the following five layers to be defined:

- Road-level (L1)
- Traffic infrastructure (L2)
- Temporary manipulation of L1 and L2 (L3)
- Objects (L4)
- Environment (L5)

The five layers also contain continuous parameters. These include, for example, the speed of other objects. More specifically, this may be the speed of a traffic participant cutting in front of the SUT. Through the theoretically infinitely fine discretization of continuous parameters, an infinite number of concrete scenarios can be defined. In addition, each concrete value of a parameter can be combined with any other value of the remaining parameters, which corresponds to the so-called N-wise testing. Consequently, the scenario-based approach also requires a methodology that identifies relevant test cases. For this reason, an outline of existing procedures in the literature for selecting and reducing concrete scenarios is given below.

Scenario selection and reduction methods

Instead of combining all values of one parameter with every other parameter (N-wise testing), an intelligent selection of parameter combinations is chosen during the Design of Experiments (DoE). According to KUHN [12], DoE-approaches can be used for complex software systems, because only the combination of a few parameters is sufficient to cause a faulty behavior of the system, which is expressed by the failure triggering fault interaction (FTFI) number. At NASA, for example, the combination of only six parameters covers almost 100% of the errors occurring [12]. Further information on various methods of DoE, such as covering arrays, as well as the application in the field of automated driving, can be found in [13–16].

- + Good parameter space coverage
- The selection of important parameters is difficult in advance
- No selection of test cases based on relevance

SAATY [17] provides the basis for an approach to detect important parameters and their connection using the Analytic Hierarchy Process (AHP). XIA [18, 19] takes this as a basis and uses the AHP, in combination with an expert knowledge-based analysis of key influencing factors. The scenarios generated are evaluated using a complexity index.

- + Analytical method for the determination of relevant parameters
- + Creation of complex scenarios
- Requires expert knowledge
- Does not consider presumably simple scenarios that nevertheless lead to faulty behavior

An approach that relies entirely on expert knowledge is the creation of scenarios with the help of ontologies [11, 20, 21]. Ontologies are a formal representation of knowledge and its relations, which have their origin in the Semantic Web. Starting from the definition of knowledge, for which the five-layer model by BAGSCHIK [11] described above

can be used as a basis, scenes are automatically derived to safeguard the automated driving function. Scenarios can be created by sequencing individual scenes. Furthermore, knowledge-based approaches ensure that elements from the knowledge base can also be found in the scenes and scenarios created. For example, if a pedestrian is defined in the knowledge base, it can be ensured that there are also test scenarios that contain a pedestrian.

- + Elements defined in the knowledge base are also part of the test catalog
- Solely based on expert knowledge
- No evidence of the relevance of the defined scenarios for the proof of safety

Instead of using expert knowledge, test scenarios for automated driving can also be based on real and simulated traffic situations with a high level of criticality. All extracted scenarios can be stored in a database. This approach, with a database filled with scenarios as the central element of the validation procedure, is used in the German-funded project PEGASUS [22, 23]. In addition to the data sources already mentioned, in principle all possible sources of scenarios can be taken into account. For example, these can also be scenarios from a knowledge-based approach.

- + Inclusion of scenarios of various origins possible
- High storage requirements with nearly identical scenarios
- If the number of stored scenarios exceeds a manageable number, a method for selecting relevant scenarios is required again

Existing accident databases can also be used to select relevant scenarios [24]. Accident scenarios are reconstructed in simulation and examinations are performed to establish whether the accident could have been prevented or mitigated by the automated driving function to be tested. The prerequisite for this is detailed accident data that contains information about the pre-crash trajectories of the vehicles involved in the accident scenario. These currently only exist for driver assistance systems. Consequently, the significance of the safety level of automated vehicles based on these accidents is extremely limited. This can be improved by further varying the parameters, such as the ego speed.

- + Shows accident avoidance potential of the automated vehicle
- Detailed accident data required
- Provides only limited information on new risks and accidents introduced by the automated vehicle

Another method based on real driving data from human drivers is the accelerated evaluation of automated vehicles [25–28]. Based on the real driving data of a maneuver (e.g., cut-in), frequency distributions of the parameters involved (e.g., relative speed) are determined. These parameter distributions are adapted in such a way that more severe situations arise that can be used for the accelerated evaluation of the SUT. The importance sampling theory is used to ensure that the accelerated result is valid, and that the acceleration factor can be calculated. Thus, a factor of up to 10^5 can be achieved, which means that each simulated kilometer corresponds to a real distance of 10^5 kilometers.

- + Conversion from simulated to real life traffic kilometers
- + Straightforward comparability with human performance through, for example, kilometers per accident
- Frequency distributions of all relevant parameters necessary → time-consuming and cost-intensive data collection
- Number of necessary involved parameters unknown

The aim of [29–32] is to adapt existing logical scenarios in such a way that relevant, and as critical as possible, concrete scenarios can be created based on them. Starting from a baseline scenario, the trajectories of road users are adapted in such a way that the planning of a safe trajectory for the automated vehicle becomes particularly challenging. Starting with the start scene, the possible trajectories of the road users are predicted into the future. The Reachable Sets, for example, can be used here. All areas in which no other road users can be located in the future are considered safe. Minimizing these safe areas results in particularly important situations for the trajectory planning module of the SUT.

- + Also suitable for online evaluation of the selected ego trajectory
- Trajectory planning is more the focus of the tests than the overall system

While all previously explained procedures evaluate scenarios before they are executed, and optimize them with regard to criticality, critical scenarios can also be derived with the help of simulation executions [14, 33]. A concrete scenario is chosen as the starting point, executed in simulation and evaluated using a criticality metric. Subsequently, specific parameters of the scenario are varied and the change in criticality is evaluated. To maximize criticality during optimization, classical optimization methods [14] as well as machine learning approaches [33] can be used. Theoretically, this approach is also possible in real experiments, but due to the high number of experiments, this is not feasible.

- + Criticality optimization during test execution ensures that critical scenarios are discovered
- Numerous simulations of concrete scenarios required → Particularly time and cost-intensive when high-fidelity simulation models for vehicle dynamics, sensors and environment are used

One approach to reducing the number of relevant scenarios that can be used in parallel to the methods mentioned above is functional decomposition. Based on the decomposition of the human driving task into five layers by GRAAB [34], AMERSBACH [35, 36] adapted this division to fit automated vehicles, using a six layer model. This division can be seen as a further subdivision of the well-known sense-plan-act principle. The aim is to reduce the number of relevant scenarios by considering the individual functional levels separately. For example, to test layer three (situation understanding), all parameters that have no influence on this layer can be omitted.

- + Approach can be used in parallel with other scenario selection methods
- Tests at overall system level are still necessary
- Reduction of the number of scenarios in the overall assessment not yet conclusively clarified

As a final method for the reduction and selection of relevant scenarios for the safety verification of automated vehicles, the formal methods will be introduced [37–39]. The aim of this approach is a mathematical proof of the safety of the SUT. If this proof is successful, then the formal methods are the most effective reduction method, because all tests, whether in simulation or in real tests, become obsolete. This currently fails because of assumptions that must be made but that do not correspond to reality. For example, the authors of [39] assume that the automated vehicle can always precisely determine the current coefficient of friction. However, the exact online determination of the coefficient of friction has been an unsolved problem for years.

- + If formal proof is provided, no tests need to be carried out
- Assumptions must be made that do not accurately reflect reality
- Whether or not driving behavior according to the formal methods leads to unreasonably defensive behavior has not yet been conclusively determined
- It must be demonstrated that an automated vehicle is implemented according to formal methods

Gaming of tests during type approval

As already mentioned in the INTRODUCTION, type approval of automated vehicles is the central application of this publication. In this context, gaming of tests is referred to as a performance optimization towards the standardized test cases. For type approval, the regulations of the UNECE¹ are relevant for the European area. These regulations must be tested by a technical service and confirmed to be complied with so that a new vehicle model can be introduced onto the market in the countries of the contracting parties. To this day, the UNECE regulations for the type approval of vehicles have made significant effort to ensure comparability and reproducibility. For this reason, the test execution, environmental conditions and evaluation of the tests are defined precisely in the regulations. This offers vehicle manufacturers the advantage of knowing in advance which tests are to be carried out, and of optimizing the performance of their systems within these test cases. A reliable statement about the system's behavior in real traffic conditions is, therefore, only possible to a limited extent, as the problems in the emission tests starting in September 2015 [40] clearly revealed.

With regard to vehicle emissions, UNECE Regulation 83 Revision 5 [41] is relevant for vehicle approval. In addition, the UNECE Global Technical Regulation 15 [42], which applies to a larger number of contracting parties (e.g., including the USA), specifies the Worldwide Harmonized Light Vehicle Test Procedure (WLTP). These two regulations stipulate the exact procedure for carrying out the tests and the ambient conditions to be satisfied. On October 10th, 2017, the UNECE clarified in Amendment 5 to Regulation 83 Revision 5 that the Contracting Parties using this Regulation in combination with Global Technical Regulation 15 (WLTP) no longer have to accept type approval on the basis of these Regulations as an alternative to their national/regional laws. The UNECE is currently revising the regulations within the Working Party on Pollution and Energy (GRPE), so that they reflect the actual emissions in real traffic more accurately [43][44, p. 2] and the results thus correspond better to actual driving behavior. By 2020, new regulations shall be adopted on the basis of test procedures already developed by other organizations (e.g., the European Union) [43]. The European Union already introduced its own methods and approval regulations in September 2017, which determine Real Driving Emissions (RDE) in a test procedure under real driving conditions [45]. The 100-percent reproducibility is thereby limited at the expense of better transferability to real traffic events.

¹ <https://www.unece.org/info/ece-homepage.html>

Initial tendencies exist not only in the area of emissions, but also in assisted driving, to reduce the exact reproducibility for a higher significance under real traffic conditions. In assisted and, especially, automated driving, the behavior of the system in real driving conditions is very important because these are safety-relevant systems. Therefore, it is important to prioritize the prevention of gaming of tests in this area. UNECE Regulation 79 Revision 4 [46] makes a first step in this direction by defining, among other things, the test cases for the Lane Keeping Assist (LKA). However, not all concrete test scenarios are explicitly defined. The logical scenario is defined in which the vehicle approaches a curve, where the lane shall have a clearly visible lane marking line on both sides. The test is passed if the vehicle does not cross a marking line during the test. In addition, paragraph 3.2.1.3 of Annex 8 [46] states: "the vehicle manufacturer shall demonstrate to the satisfaction of the Technical Service that the requirements for the whole lateral acceleration and speed range are fulfilled." This means that the technical service can define and test any combination of lateral acceleration and speed as a concrete scenario. It is not only within UNECE that the definition of exclusively standardized tests is avoided. In its Federal Automated Vehicles Policy [47, p. 77], the NHTSA also clearly opposes the exclusive use of standardized tests.

While today the UNECE Regulation 79 specifies a variation of two parameters (lateral acceleration and speed) for the LKA, the number of parameters to be varied will continue to increase with higher degrees of automation. It is conceivable, for example, that legislation (e. g., UNECE Regulation) will only define logical scenarios (e. g., cut-in) for the type approval of automated vehicles. During the certification process of the automated vehicle, the manufacturer and technical service have to find all relevant concrete scenarios of the predefined logical scenario by varying all possible parameters from the five-layer model introduced in the previous section. This gives the technical service the chance to adapt the concrete test cases towards the system's individual weaknesses. Additionally, this prevents performance optimization towards standardized tests, the so-called gaming of tests.

Research objective

The last sections revealed that it is important but difficult to reduce the number of test cases for the safety assessment and certification of automated vehicles. To achieve this, a system-specific derivation of relevant test scenarios is advantageous. None of the approaches stated above explicitly include system-specific properties, and therefore this kind of approach is lacking from the current state of the art. In addition, the method developed is intended to prevent the so-called gaming of tests in the certification process, ensuring the safety of automated vehicles even under real traffic conditions.

The aim of this paper is to propose a novel method of including system-specific characteristics for an efficient and individual scenario selection during the certification process of automated vehicles conducted by a technical service.

METHODOLOGY

This section describes in detail the overall procedure used, from the required input, to the method to the generated output. An overview can be found in Figure 2.

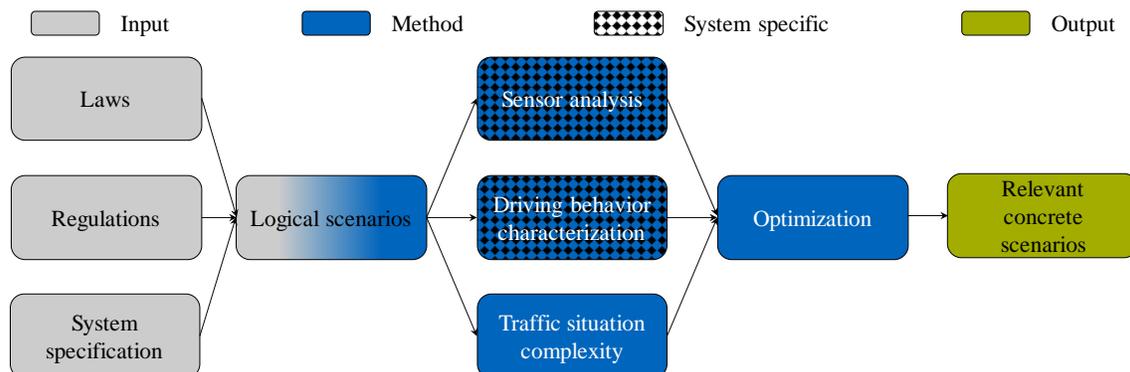


Figure 2. Overview of the procedure developed. Logical scenarios are assigned both to the input and to the method, because these are partly specified in regulations and in these cases, they serve as input.

Summary of presented approach

The starting point of the methodology developed is logical scenarios defined in laws and regulations. From each predefined logical scenario, an infinite number of concrete scenarios can be generated by a theoretically infinitely fine

discretization of the parameters. The aim of the method is to restrict the parameters of the logical scenarios in such a way that a technical service can efficiently identify the most relevant test cases for the certification of the automated vehicle. This results in two special requirements for the procedure. On the one hand, it should be system-specific so that different versions of various manufacturers can be addressed efficiently. On the other hand, it should be able to work without using the driving function to be tested as far as possible, because the technical service will not usually have access to the software. What will be available to the technical service is a system specification of the SUT. The information contained therein is used in the further course of the approach.

The parameters of the logical scenarios are determined by optimization, based on the three main elements of the concept. These elements are the analysis of the sensors, consideration of the driving behavior and integration of complexity. The first two elements are specially adapted to the identified weak spots of the SUT, and thus enable a system-specific definition of relevant scenarios. The predefined logical scenarios contain only a simple description of a scenario that is suitable for demonstrating certain basic skills. For example, in a cut-in scenario, only one vehicle is defined that performs a lane change into the lane of the SUT during the course of the scenario. In order to give the scenario a certain degree of difficulty, system-independent complexity is integrated in the scenario in the final step by defining further TPs. In the optimization, all three elements must be considered in parallel. The result will be relevant concrete scenarios that are adopted towards the SUT weaknesses, which are used by the technical service for the certification process. Below, each step of the method is explained in detail.

Input

Laws, regulations and a system specification are required as input for the approach developed and will be explained in more detail below.

Laws If an automated vehicle performs the driving task, the system also takes responsibility for the actions performed. Consequently, the actions of the automated vehicle must comply with the applicable laws. In Germany, for example, these may be the Road Traffic Act (in German: Straßenverkehrsgesetz (StVG)) or the Road Traffic Regulations (in German: Straßenverkehrs-Ordnung (StVO)), which define, for example, the minimum safety distance from the vehicle in front, or maximum speeds on certain types of road. The applicable laws can be explicitly tested in a scenario, or at least considered in the evaluation of the tests carried out.

Regulations Referring back to the LITERATURE REVIEW AND RESEARCH OBJECTIVE section, certain regulations must be complied with for type approval. For the European market, these are the UNECE regulations, which represent the central aspect in the considered use case of the approach developed. Like already discussed, the UNECE regulations no longer define all tests in detail, but provide a certain framework for the tests to be carried out, which may also implicitly mean logical scenarios. Furthermore, the system manufacturer and the technical service are assigned an increasingly higher responsibility, because they have to verify that the defined requirements are met over the entire operating range. This, in turn, gives the technical service the opportunity to define all scenarios that it considers relevant for the system to be tested as definite test cases during the certification process.

System specification In the future, the number of automated systems with different functional capabilities will continue to increase. For example, there will be a wide variety of systems for the highway domain. These will differ in terms of whether special situations, such as construction sites, ramps and exits etc., can be handled by the automated driving function. In order for the selection of the logical scenarios to be tested to be as effective as possible, a system specification with the definition of the system boundaries must be available to the technical service. The system boundaries can also be used to generate test cases where the system boundaries are exceeded. The SUT also has to achieve a safe state in these scenarios. In addition, the technical service must be informed of both the sensors used and their installation position. All in all, the type and level of detail of the information provided by the manufacturer in the system specification must be such that, on the one hand, the technical service has sufficient information available for the relevant scenarios to be selected reliably and, on the other hand, the manufacturer does not have to reveal too much manufacturer-specific data in order to protect internal know-how.

Logical scenarios If logical scenarios are already defined in applicable regulations, these can be used as input for the developed approach. It may also be possible that further logical scenarios are required for the method, for example to test system functions or compliance with road traffic laws, which are required in the regulation but for which no logical scenarios are defined. An example could be that the UNECE specifies that speed limits must be observed by the system but does not specify a framework in which the tests are to be carried out. For the reasons described, the

logical scenarios in Figure 2 are assigned to both the input and the method. Requirements can also be specified in the regulation without a logical scenario being defined for the verification of these requirements. This may require the technical service to define further logical scenarios during the type approval process. It is also conceivable that a scenario is only described in linguistic terms and is therefore, by definition, a functional scenario. In this case, the corresponding logical scenario must also be defined by the technical service.

Output

The output of the presented approach is represented by relevant concrete scenarios for the type approval of automated vehicles. All scenarios required to prove compliance with the requirements (e.g., by UNECE) are considered relevant for testing. The unique selling point of the developed procedure is a system-specific adaptation of the relevant scenarios to accommodate possible weak points of the system to be certified.

Method

The following explanations of the methodology blocks shown in blue in Figure 2 outline how the output described here is generated from the input described above.

Sensor analysis The perception sensors of an automated vehicle have the objective of collecting information about the environment of the vehicle, which is an important part of the driving task. Only if sufficient information about the environment is available, a safe planning of appropriate actions of the automated vehicle is possible. Not only sufficient information needs to be available, but it must also correspond to reality. Currently, in the automotive sector, the perception sensors mainly used are Radar, Lidar, camera and ultrasonic. Each of these sensor types has its own advantages and drawbacks. In order to enable the driving task to be performed safely, the various sensor types are therefore combined, and the information obtained from them is fused. For the sake of simplicity, perception sensors are simply referred to as sensors in the following.

The resulting sensor costs have a major influence on the choice of the number and type of sensors installed, especially in mass production vehicles. Furthermore, there may be restrictions due to the package. Consequently, every vehicle manufacturer will use an individual sensor setup. The goal of this section is to formulate a method for a structured analysis of the sensor setup used and to identify its weaknesses. This is particularly relevant for technical services in type approval, because they have to efficiently test vehicles from different vehicle manufacturers with different sensor configurations. In addition to the ODD (e.g., motorway or urban area) of the vehicle under consideration, the type and number, as well as its installation position and pose, also play a critical role. In the following section, the most important influencing variables of the sensors are examined in more detail. All this data must be specified by the manufacturer in the system specification for each sensor.

- 1) **Field of view:** The field of view describes the visual area of a sensor and is influenced by the following properties:
 - a. **Range:** The different sensor types have widely-varying ranges. Ranges vary from a few meters with ultrasonic to several hundred meters with radar sensors. Different ranges also occur within the same sensor type, for example with camera sensors. While the range only has a secondary role in inner-city operation, it is a decisive factor in the use case of the highway pilot, due to the high speeds.
 - b. **Opening angles:** One can differentiate between the horizontal and vertical opening angle. Both the horizontal and vertical opening angles are particularly important for detecting the vehicle's immediate surroundings. Large opening angles can reduce or even eliminate blind spots between the mounting positions of the sensors, and improve the detection of low-lying objects. As the distance from the vehicle increases, the importance of the vertical opening angle decreases, because only a relatively narrow area of the environment is relevant in the vertical direction. Only with Lidar sensors, which usually have a small vertical opening angle, does this have to be taken into account at higher distances. For example, when braking strongly, a pitch angle can occur that shortens the range of the Lidar sensor.
 - c. **Mounting position:** The mounting position influences the field of view of the sensor. In general, high mounting positions are advantageous, because, in particular, objects at great distances can be detected better. In addition, sensors with high mounting positions are better protected against damage. These can be, for example, contamination by whirled-up dust and dirt, parking bumpers and similar.
 - d. **Orientation:** In addition to the mounting position, the orientation of the sensors is also relevant for the field of view. Sensors pointing slightly downwards can be used primarily to cover the immediate

surroundings, whereas sensors for the more distant surroundings tend to be oriented almost horizontally.

- 2) **Quality of data:** The quality of the recorded data is not always the same and varies depending on the following factors:
 - a. **Type:** As already mentioned, camera, Radar, Lidar and ultrasonic sensors are mainly used in the automotive industry. These sensor types have different strengths and weaknesses due to their physical principle.
 - b. **Performance attributes:** The performance attributes of the sensors used must be evaluated in detail. These attributes include, for example, the cycle times of the measurement data acquisition, which can provide information about the minimum reaction time of the system. In addition, sensor type-specific information, such as the transmitted power of the Radar or the number of pixels of the camera used, may be of interest. The latter, for example, provides information about the achievable depth of detail of the recorded sensor data.

Considering the interaction of all the factors shown, the sensor setup used can be examined. In the first step, the ideal sensor coverage of the system is investigated. This requires influencing factor 1), which is explained above. “Ideal” in this context means that all objects within the field of view of the sensor are correctly detected. This makes it possible to conclude not only whether and where the sensor setup shows blind spots, but also which areas are covered by several sensors or even several sensor types. From blind spots to multiple sensors to multiple sensor types, the probability of correct detection and classification of objects generally increases.

In the second step, phenomenological sensor models based on influencing factors 1) and 2) are used to study sensor coverage at greater distances from the SUT. These extend the ideal sensor models used previously by modeling individual phenomena, such as attenuation due to weather influences or the decrease of the signal received with increasing distance from the object. Phenomenological models, therefore, have a higher information content, but also require more computing resources. For completeness, reference is also made to physical sensor models that simulate the physical effects of the sensor and, thus, best represent reality. One example is so-called ray-tracing methods, which do, however, require enormous computing capacities and are therefore not practicable for the desired purpose. Phenomenological models, on the other hand, represent a good compromise between information content and required computing capacity.

With the phenomenological sensor models used, it is possible to investigate whether there are not only areas in the far field of the SUT that are not within the detection range of the sensors, but also areas in which the detection probability is poor. It is also possible to investigate whether any prevailing weather conditions are particularly critical for the sensor setup to be investigated. Road topology may also have an influence on the probability of an object being detected. Depending on the characteristics of the sensor setup under consideration, other curve radii, longitudinal slopes, or hilltops and valleys may pose special challenges for the SUT.

This investigation is particularly interesting for systems where the highway is part of the ODD. With the procedure described, environmental conditions, such as critical curve radii or weather conditions, can be identified, thus reducing the number of relevant scenarios. A further reduction of the number of scenarios is the adaptation of the trajectories of the other traffic participants. The trajectories can be optimized in such a way that the traffic participants approach the SUT in areas where the detection probability is as low as possible.

An extension of the sensor analysis is the inclusion of the sensor data processing, which allows the entire module of the perception to be considered. A detailed insight into the data processing software will not be available to the technical service, but it is still possible to include known state-of-the-art weaknesses in the adaptation of the tests. According to DIETMAYER [48], errors that can occur in the perception are assigned to the following three categories:

- 1) **State uncertainty:** Deviations between the measured state variables (such as position or speed) and those that are correct.
- 2) **Class uncertainty:** The classification of the detected object is incorrect. One example is the classification of a motorcycle as a cyclist.
- 3) **Existence uncertainty:** An existing object is not detected, or a non-existent object is incorrectly detected as an object (ghost object).

The technical service cannot accurately predict the occurrence of these errors, but the current state of the art can be used to identify situations in which the probability of one of the three error types occurring increases. For example, it is known that Radar sensors have problems in detecting stationary objects. This information can be taken into account when creating concrete scenarios.

In summary, the objective of analyzing the sensor setup used is for the technical service to adapt the given logical scenarios. This enables concrete test cases that address the identified weaknesses of the SUT's sensor setup to be selected efficiently for certification.

Objective characterization of driving behavior The aim of the objective characterization of the driving behavior of the automated vehicle is to identify systematic weak points in the driving behavior, so that the scenarios used for type approval can be adapted accordingly. This should be possible by means of a limited number of functionality-based tests carried out in advance in special scenarios, so-called characteristic situations. An example of this is the curve driving behavior of the system, which can be investigated beforehand. There may be systems that tend to drive on the inside of curves and others that have a more outside tendency. This information can be used, for example, to adapt cut-in situations in curves specifically to the side where the automated system is more likely to drive. If it drives on the inside of a turn, a vehicle cutting in from the inside direction is more relevant for this vehicle, because the lateral distance between the vehicles tends to be smaller.

A structured approach is required to ensure that all the necessary information can be obtained efficiently. Figure 3 gives an overview of the procedure used. Information sources are displayed in gray, the generated situation catalog in blue and the further utilization in green. Most of the information sources used are state-of-the-art and are already known, at least for manually controlled vehicles. The novelty of this approach is the transfer to automated vehicles and the further use of the information acquired for their safety assessment.

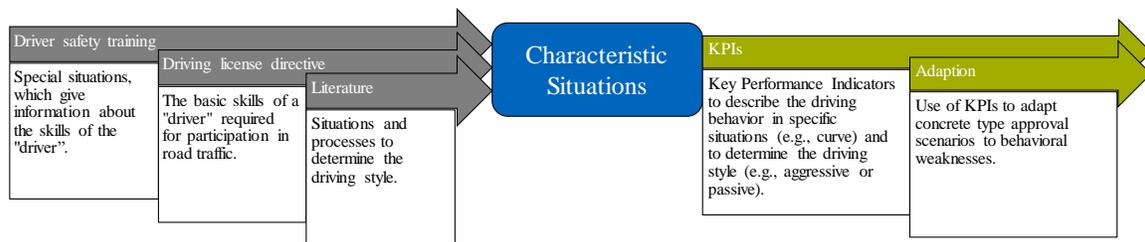


Figure 3. Developed approach for an objective characterization of SUT's driving behavior

In driver safety training courses, the driving skills of human drivers are tested in particularly difficult situations, some of which do not appear or occur extremely rarely in real road traffic. Nevertheless, some tests of driving safety trainings can be used to make a statement about the "capabilities" of automated vehicles. For example, the behavior of the automated vehicle at low coefficients of friction can be investigated.

In Germany, the Driving License Directive defines the minimum requirements that a person must meet in order to be allowed to actively drive a vehicle in road traffic. Not all requirements (e.g., correct adjustment of the side mirrors before driving) can be transferred to an automated vehicle. However, the practical driving test does include some requirements for behavior that can also be applied or transferred to an automated vehicle. These are applicable if the driving examiner drives the automated vehicle in real road traffic and can, at some effort, be transferred if they are carried out in simulation and the requirements of the driving examiners have to be objectified and automated. In addition, the theoretical exam contains a category of questions on how to behave in certain sample situations. These are currently being shifted into theory because the probability of these situations occurring is relatively low. When automated vehicles are tested, these types of tests can be transferred to simulation, and the correct behavior of the vehicles can be tested. From this, it can be concluded that, in addition to the accident-free handling of the type approval tests, the "functional" requirements of the Driving License Directive must also be considered and continuously checked. These types of scenario represent relevant scenarios that do not necessarily have to be critical or complex.

Existing literature can be used to determine the driving style. Studies have already been carried out to determine the driving style of both human drivers [49] and for automated vehicles. Comparisons have also been made between manual and automated driving, such as in [50]. Abnormal behavior (whether aggressive or defensive) can lead to

increased risk in the interaction between human drivers and automated vehicles. Hazards occur when the behavior of an automated vehicle does not match the response expected from human drivers, which may be due to inadequate communication between the driver and the automated vehicle. In order to identify this type of risk and integrate it into the further course of the type approval process, it is advisable to use the Systems-Theoretic Process Analysis (STPA) [51] risk analysis method. This method is particularly suitable because it examines the interaction between complex systems as the cause of errors. The individual systems themselves can operate error-free. Applied to our use case, this means that neither the automated vehicle nor the human driver behaves incorrectly, but their interaction nevertheless leads to risks.

Combining the knowledge from the three information sources shown, a catalog with characteristic situations can be created. If these are carried out in a real test or simulation, key performance indicators can be evaluated that reflect the driving style and weaknesses in driving behavior. This information can then be used for the further type approval process by adapting all future test scenarios to suit the weaknesses identified in the vehicle's driving behavior. For this part of the method, the driving function is required for a limited number of tests. If this is not available, the overall method is still applicable - however, the efficiency in selecting the test cases decreases.

Complexity The aim of this section is to design the concrete scenarios for the type approval of automated vehicles as complex as possible. For a better understanding of the distinction between complex and critical scenarios, we refer once again to the 'Definition of vocabulary' subsection. In future regulations (e.g., UNECE), logical scenarios will be defined that confirm that functional requirements for the system to be tested are fulfilled correctly. One example of this could be the possible logical scenario "object in the SUT's lane", in which the automated vehicle must react appropriately to an object in its lane. In [52], the functional requirement for this logical scenario is that the vehicle must be able to avoid the object by means of braking, steering or a combination of both. Analogous to the currently applicable UNECE R79, which requires proof that the Lane Keeping Assist can be transferred to general situations, this will also be required here. Adding complexity to traffic situations represents an essential component in the transfer to general situations in which an object is located in one's own lane.

The procedure described here should, therefore, be used to introduce a general system-independent difficulty to the given, simple logical scenarios. This is accomplished by adding complexity to the given logical scenario. In general, complexity can be caused by components of all five layers, according to BAGSCHIK [11]. For example, a change in the sign of the lateral slope of the roadway (roadway twisting), which is assigned to layer one of the five-layer model, can represent a particular complexity for the lateral guidance of the vehicle. Consequently, all five layers must be examined separately. This paper limits itself to layer four (objects). Objects are used here as an overarching term for obstacles and other road users of all kinds. Objects can, therefore, be stationary or movable. Due to the limitation to layer four, complexity is understood in this publication as the difficulty faced by the planning algorithm in planning a safe trajectory resulting from the movement or presence of objects.

According to BACH [53], no abstract definition of complex situations for automated vehicles has yet been determined. One approach is offered by SCHAUB [54], who has defined eight criteria for complex situations in which people have difficulty making decisions. In his work, SCHULDT [55] examined how these criteria can be transferred to the complexity of traffic situations for automated vehicles in theory, and concludes that they are also adaptable to this application. The complexity criteria defined by SCHAUB [54] and confirmed by SCHULDT [55] for automated vehicles are as follows:

- Number of elements
- Number of states per element
- Interdependency
- Self-dynamics
- Intransparency
- Multiple conflicting goals
- Openness of the target situation
- Novelty

For a detailed description of the meaning of each criterion in relation to complex traffic situations for automated vehicles, see Chapter 2 in SCHULDT [55]. It is not possible to use these characteristics directly to create complex scenarios, because they only exist in verbal form so far. In addition, the characteristics are used to evaluate existing scenarios and not to generate new ones. A further aspect that impedes direct use is the fact that the evaluation has so

far only been carried out subjectively. This is not sufficient for the procedure developed here, because an automated optimization of the scenarios is conducted in the next step.

In order to apply the criteria for describing complexity that exists in literature, four main points need to be analyzed. The procedure is summarized in Figure 4. First of all, it is necessary to ascertain whether further criteria are required to generate the scenarios carried out here. This can be accomplished by defining a list of requirements and comparing the extent to which the existing criteria meet all of them. Secondly, an evaluation must be performed to determine whether the characteristics have a certain upper limit. The number of the involved elements (objects) is theoretically not limited, but due to the physical allocation only a limited number of objects are important for executing the driving task. The number of objects to be considered must be examined in order to apply the presented method efficiently. Thirdly, the criteria that were previously only defined verbally must be described mathematically in order to be usable for the subsequent optimization. Finally, the objectified attributes must be validated using simulations and real driving data. Since complex scenarios do not necessarily have to lead to a critical outcome of the scenario, special key performance indicators are necessary for validation. A parameter that could be used for the validation is, for example, whether the automated vehicle changes its decision about the actions to be performed during the scenario. If the vehicle starts to dodge an object on the left and then decides to dodge it on the right, this is an indication of a complex scenario, even if the outcome of the scenario is not critical.

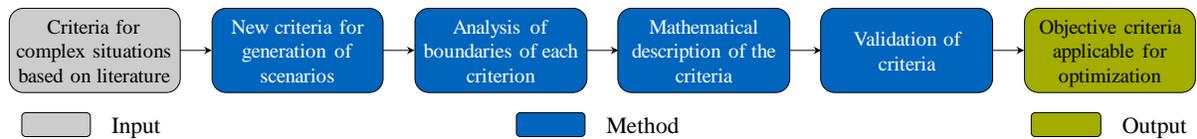


Figure 4. Approach for the definition of an objective complexity metric.

Optimization In the last step of the developed methodology, the individual components previously described are integrated. For this step, the driving function to be tested does not have to be available. The starting point is an arbitrary logical scenario, which is prescribed in the regulation under consideration. The aim is to adapt or select parameters of the logical scenarios that BAGSCHIK [11] defines in a five-layer model in such a way that relevant concrete scenarios are derived from the given logical scenario. Each of the three methods described in detail above contributes to the determination of specific parameter values within the five layers, which is described below and summarized in Table 1. The temporary manipulation of L1 and L2 (L3) represents unusual traffic situations, such as a changed lane routing marked by pylons within a temporary construction site. These special situations are not the focus of this work and will not, therefore, be considered further.

Table 1. Considered influence of the individual aspects on the determination of the parameters of the five-layer model according to BAGSCHIK [11].

	L1	L2	L3	L4	L5
Sensor analysis	✓	✓	✗	✓	✓
Driving behavior	(✓)	(✓)	✗	✓	(✓)
Complexity	✗	✗	✗	✓	✗

Influence: ✓ high (✓) weak ✗ none ✗ neglected

Sensor analysis: The sensor analysis can be used to determine parameters from all four considered layers of the five-layer model. For optimization, the relationship between the individual layers must be taken into account. If, for example, obstructions are taken into account, the position of objects (L4) can have an influence on infrastructure elements such as traffic signs (L2). In addition, connections also exist between complexity and sensor analysis. Sensor analysis identifies areas in which the probability of detection is low. In the subsequent optimization of the complexity, the trajectories of the objects are calculated in such a way that the scenario is as complex as possible, and the potential conflict partner is located, as far as is possible, in the previously-identified areas with weak sensor coverage. A potential conflict partner is the object that forces the SUT to act. Thus, an object in its own lane can be the potential conflict partner forcing the SUT to avoid or brake.

Driving behavior: The influence of the identified driving behavior on the optimization of the parameters is evaluated as weak for most layers. Thus, in Road-level (L1), the analysis of driving behavior cannot determine which exact curve radius is of particular importance for the system, but it can conclude whether the system tends to drive through curves more on the inside or outside. This information can then be used to adjust the position of the potential conflict partner from layer four (objects) within curves.

Complexity: The parameters of the fourth layer (objects) can be determined by systematically introducing complexity. As mentioned before, other layers can also contribute to increasing complexity, which is neglected in this paper. The definition of complexity is defined here as the complexity of the motion and presence of objects. Here, the connection to both other aspects is given by calculating the trajectories of the objects in such a way that they address not only the complexity but also the sensor deficiencies, as well as the identified driving behavior. While the two previous steps are mainly limited to the trajectory of the conflict partner, the optimization of complexity also focuses on defining further objects and their optimal trajectory.

All in all, this results in a two-stage procedure in which the parameters of layer one, two and five are first optimized by sensor analysis and consideration of driving behavior. In addition, the trajectory of the potential conflict partner (L4) is determined. In the second step, further objects are defined by considering complexity and their trajectories are optimized. Therefore, two consecutive optimizations are carried out, each with a suitable algorithm. Especially in the second optimization step, a multi-objective optimization (Pareto optimization) can be of significant importance due to the competition of multiple factors of complexity.

The definite cost function for the first optimization step will vary from system to system, because these two aspects have the very objective of finding system-specific weaknesses. Since the result influences the second optimization step, these results will also be dependent on the SUT. In addition, when optimizing the parameters and, in particular, when optimizing the trajectories of the objects, it must be ensured that certain constraints are met. For example, there must be enough space for the SUT that it is physically able to cope with the scenario without causing an accident, meaning that so-called dilemma situations are not considered. In addition, the trajectories of the objects or the other TP must be physically possible. Depending on the type of TP, approximations such as a simple point mass model, the circle of forces or a single-track model can be used.

The optimization method presented results in more than one relevant concrete scenario from the given logical scenario, due to the applied multi-objective optimization. To execute these scenarios efficiently, simulation is suitable, in which the required parameters - in contrast to test site tests - can also be set without major expense.

RESULTS

This chapter shows an exemplary elaboration of the concept using a driving function designed for the ODD highway. The individual blocks of the approach shown in Figure 2 and the resulting concrete scenario for the example function are explained.

The input of the exemplary results shown here is a fictive system description including the used sensors and a draft of a regulation² for the certification of automated vehicles. Within this fictitious regularity, the logical scenario "avoidance of a stationary obstacle" is defined, which represents the starting point for the individual steps of the method. The purpose of the predefined logical scenario is to show that the SUT has the principal functionality needed to avoid a stationary obstacle. A schematic sequence of the scenario in its simplest form is shown in Figure 5 on the left-hand side. The SUT drives in the right lane of a two-lane highway and detects a stationary object with its sensors. Detection takes place at an early stage due to the unrestricted view. Since the adjacent lane is not occupied, the SUT can change lanes and drive past the stationary object with sufficient side clearance. The successful handling of this scenario can be understood as the fulfilment of the functional requirements mentioned above. If, analogous to the existing UNECE R79, the technical service, in cooperation with the manufacturer, is required to prove that this capability can be transferred to general situations (here: general evasive situations of a stationary obstacle), the methodology developed can be used to define system-relevant concrete scenarios that are not previously known to the manufacturer.

² Proposal in UNECE 'Working Party on Automated/Autonomous and Connected Vehicles (GRVA)' informal working group 'UN Task Force on Automated Vehicle Testing (AutoVeh)' [52]

The following paragraph explains how the optimized relevant scenario (right-hand side of Figure 5) can be identified systematically using the method presented. In the optimized scenario, the SUT's view of the stationary object is restricted by a curve on the one hand, and by another TP on the other. Since the object was detected at a late stage, it is no longer possible to brake to a standstill in front of the object. The SUT must switch to the left lane to prevent an accident. This lane is blocked by other vehicles moving at slightly higher speeds than the SUT. The SUT must brake to avoid colliding with the object, but cannot brake too much, in order to avoid a risk to TP 3 when it changes lanes into the gap between TP 2 and TP 3.

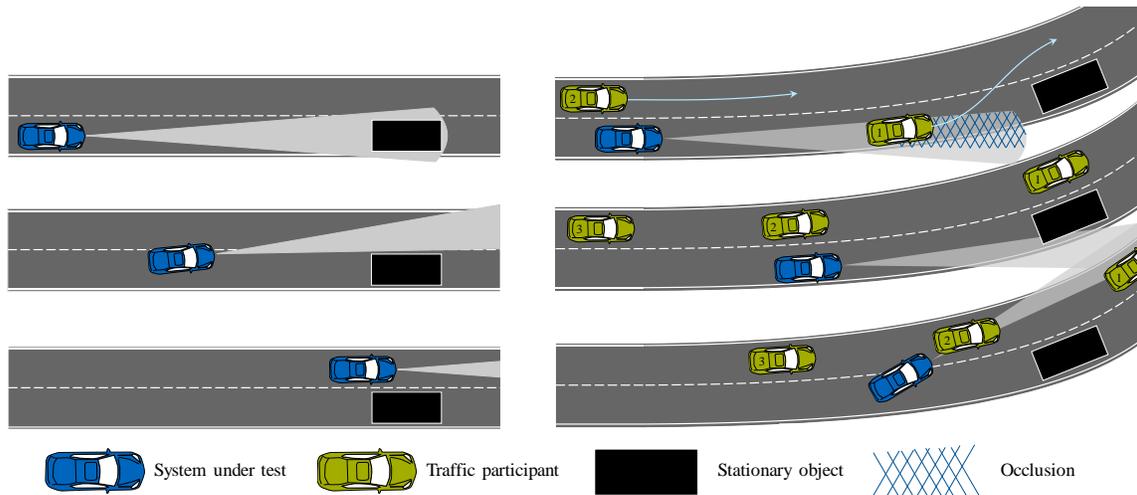


Figure 5. Logical scenario "avoidance of a stationary obstacle" in its basic version (to the left) and in an optimized version (to the right).

Sensor analysis: Using the system specification, disadvantageous Road-level (L1) parameters for the sensor setup used can be determined. In the example shown here, this is a curve with a certain radius, which means that the sensors used can hardly see the course of the road. In addition, the sensor analysis defines another TP (L4), which increases the occlusion of the stationary object. The distance between the SUT and the TP is defined on the basis of the opening angles of the sensors used, so that optimum occlusion of the stationary object is achieved during cornering. In addition, the environmental conditions (L5) can be selected in the scenario (no environmental conditions are shown in Figure 5, with the result that the operational weak points of the sensors used are taken into account. In a system that mainly uses cameras, this can be direct sunlight from the front.

Driving behavior: Characteristic tests carried out in advance show that the SUT tends to use the outer side of curves for orientation. Therefore, in the scenario under consideration, a left-hand curve is chosen because the lateral distance of the SUT to a safe position is then maximized. Instead of testing both left-hand and right-hand curves, this method allows a well-founded selection of this parameter.

Complexity: In the last step, system-independent complexity is added to the scenario on the basis of the criteria of SCHAUB [54] and SCHULDT [55]. For example, the number of elements is increased (definition of further TPs). The trajectories of this TP are adapted in such a way that they increase the number of actions required by the SUT to achieve a safe state, for example by representing restrictions on action. In the specific example, in order to successfully pass this scenario, the SUT has to brake in a controlled manner and reeve into a gap between two vehicles traveling at different speeds, because braking to a complete standstill is no longer possible without colliding with the stationary object. As already explained in the subsection 'Definition of vocabulary', the speed of the TP and the starting speed of the SUT are also increased, and distances between the TPs are reduced in order to increase criticality. In addition, a connection to the sensor analysis is made so that the other TPs stay as long as possible in areas with low sensor coverage or even in blind spots that are not visible to any sensor.

From the technical service's point of view, this scenario is relevant for the SUT to be tested and should, therefore, be taken into account when the SUT is certified. This can support the required proof of transferability of the capabilities to general situations in which an object must be evaded in its own lane. At the same time, it is difficult for the

manufacturer to prepare for the tests, due to the individual adaptation of the scenarios, and thus so-called gaming of tests is prevented, and the safety level achieved during certification can be transferred more effectively to real driving conditions.

For simplicity's sake, this example does not discuss all parameters that can be determined by the optimization. For example, further parameters of the road geometry can be adapted to the weak points of the SUT sensors. Even if not all parameters are explained, the example shows very clearly which results are achieved with the developed method.

DISCUSSION AND LIMITATIONS

This section critically discusses the developed approach. This concept has been specially developed for the type approval process of automated vehicles. As indicated in the section LITERATURE REVIEW AND RESEARCH OBJECTIVE, the focus is on Object and Event Detection and Response (OEDR), and therefore not all tests required for type approval are addressed. One of the basic assumptions of this paper is the specification of logical scenarios for type approval by legislation. This assumption is justified because UNECE already has existing regulations for ADAS according to this principle, and UNECE working groups already have initial proposals [52] for automated vehicles that also work with the definition of logical scenarios. In addition, large research projects, such as PEGASUS (highway) [2] and CETRAN (urban) [56], also work with the definition of logical scenarios.

With the method presented, it is not possible to determine all parameters of the five layers model precisely, because it is not entirely possible to determine the relevance of a parameter for a scenario. For instance, even after the methodology has been applied, it is unclear whether the type and shape of the central lane marking in the scenario in Figure 5 has an effect on the outcome of the scenario. Consequently, a parameter variation must still be performed in simulation. However, the extent of this variation is significantly reduced by the parameters that have already been defined. Despite constantly increasing computing power, it is important to keep the number of necessary tests in simulation low, because high-precision simulation models with considerable computing costs must be used for a meaningful simulation-based test of the overall system. The presented method thus has an advantage compared to the exclusive criticality optimization through simulation execution, because no additional prior knowledge is included in the latter. Using this prior knowledge, edge cases can be identified more efficiently with the novel method presented. In addition, the procedure presented here can largely be carried out without the actual driving function, which is of great importance for a technical service. The actual driving function is only necessary for a limited number of tests in order to determine the driving behavior. As stated in the subsection 'Objective characterization of driving behavior', if the driving function is not available, the overall method is still applicable - however, the efficiency with which test cases are selected decreases.

As with all methods for selecting and reducing scenarios described in the subsection 'Scenario selection and reduction methods', the method developed does not cover the entire parameter space. Therefore, there is no guarantee that all errors of the SUT will be detected. Finding as many faults as possible is of great importance for the system manufacturer in terms of product liability. The situation is different for a technical service when it comes to system certification. The main purpose of the type approval is to test relevant scenarios under given framework conditions by regulations. The technical service must ensure, with minimum effort, that no relevant error cases remain undetected and that the SUT conforms to the applicable regulations, which is achieved with the developed approach.

It is also possible to combine the method with existing methodologies for scenario selection. For example, it can be combined with the use of critical accident scenarios from a database. An accident stored in the database can be considered as a logical scenario to which the developed methodology can be applied. Thus, it is possible to show whether the SUT can prevent existing accidents even under disadvantageous conditions.

CONCLUSION AND FUTURE WORK

This paper presents a novel and advanced method for defining relevant test cases for the future type approval of automated vehicles. The method is specifically adapted to the requirements of a technical service performing type approval. Based on regulations currently under development, such as UNECE and other laws to be complied with, scenarios relevant to the system under test are identified. As with existing UNECE regulations (e.g., for the Lane Keeping Assist), it can be assumed that not all tests are specified in detail in the regulations, but only in the form of logical scenarios. This enables the technical service to carry out scenarios that are relevant from its perspective, thus preventing so-called gaming of tests, and at the same time perform an efficient evaluation of the vehicle to be tested.

The methodology presented is essentially based on three pillars: analysis of the sensor setup used, inclusion of driving behavior and consideration of complex traffic situations. In the first two steps, system-specific weaknesses of the system to be tested are identified. In the third step, the logical scenario given by the regulation is extended to include a complex traffic situation in order to challenge the planning algorithm of the vehicle. After a final optimization, in which the three mentioned sub-methods are combined, the concrete scenarios relevant for type approval are obtained.

The methodology for the analysis of the sensors is partly available and will be continuously improved in further work. A comprehensive analysis is currently being carried out for the characterization of driving behavior to determine what information on driving behavior can be included in the future type approval process for automated vehicles. In the future, the methodology for integrating complexity into the tests to be performed will be developed. To achieve this, linguistic definitions of complexity from existing literature will be analyzed and transformed into a mathematical form. Subsequently, a validation of the complexity will be carried out by means of simulations and real driving data. Finally, the complexity can be included in the optimization. If all three components described are available, the optimization can be carried out in a final work and the total method can be applied to a real system as well as being validated.

ACKNOWLEDGEMENT AND CONTRIBUTIONS

The work for this paper was funded and supported by TÜV SÜD Auto Service GmbH. Thomas Ponn (corresponding author) initiated and developed the main parts of this paper. Christian Gnanndt provided essential scientific background information regarding the safety assessment of Advanced Driver Assistance Systems and Automated Driving. Frank Diermeyer contributed to the conception of the research project and revised the paper critically for important intellectual content. He gave final approval of the version to be published and agrees to all aspects of the work. As a guarantor, he accepts responsibility for the overall integrity of the paper. The author would also like to thank Alexander Schwab for the critical and productive discussions on how to characterize the driving behavior of automated vehicles objectively.

REFERENCES

- [1] *Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems*, J3016_201609, 2016.
- [2] Deutsches Zentrum für Luft- und Raumfahrt e. V. (DLR), *PEGASUS-Projekt*. [Online] Available: <http://www.pegasus-projekt.info/de/>. Accessed on: Feb. 18 2019.
- [3] W. Wachenfeld and H. Winner, “The Release of Autonomous Vehicles,” in *Autonomous Driving: Technical, Legal and Social Aspects*, M. Maurer, J. C. Gerdes, B. Lenz, and H. Winner, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 425–449.
- [4] NHTSA, *Automated Driving Systems 2.0: A Vision for Safety*. September 2017. [Online] Available: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf. Accessed on: Feb. 18 2019.
- [5] S. Ulbrich, T. Menzel, A. Reschka, F. Schuldt, and M. Maurer, “Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving,” in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 982–988.
- [6] T. Menzel, G. Bagschik, and M. Maurer, “Scenarios for Development, Test and Validation of Automated Vehicles,” in *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018, pp. 1821–1827.
- [7] *Road vehicles — Safety of the intended functionality*, ISO/PAS 21448, 2019.
- [8] J. C. Hayward, “Near-Miss Determination Through Use of a Scale of Danger,” *Highway Research Record*, no. 384, 1972.
- [9] National Transportation Safety Board, *Preliminary Report Released for Crash Involving Pedestrian, Uber Technologies, Inc., Test Vehicle*. [Online] Available: <https://www.nts.gov/news/press-releases/pages/nr20180524.aspx>. Accessed on: Feb. 18 2019.
- [10] F. Schuldt, F. Saust, B. Lichte, M. Maurer, and S. Scholz, “Effiziente systematische Testgenerierung für Fahrerassistenzsysteme in virtuellen Umgebungen,” https://publikationsserver.tu-braunschweig.de/receive/dbbs_mods_00052570.
- [11] G. Bagschik, T. Menzel, and M. Maurer, “Ontology based Scene Creation for the Development of Automated Vehicles,” in *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018, pp. 1813–1820.
- [12] D. R. Kuhn, D. R. Wallace, and A. M. Gallo, “Software fault interactions and implications for software testing,” *IEEE Transactions on Software Engineering*, vol. 30, no. 6, pp. 418–421, 2004.

- [13] B. S. Ahmed and K. Z. Zamli, "A review of covering arrays and their application to software testing," *Journal of Computer Science*, vol. 7, no. 9, p. 1375, 2011.
- [14] M. Tatar, "Chasing Critical Situations in Large Parameter Spaces," in *Autonomous Vehicle Test & Development Symposium*, Stuttgart, 2018.
- [15] F. Schuldt, A. Reschka, and M. Maurer, "A Method for an Efficient, Systematic Test Case Generation for Advanced Driver Assistance Systems in Virtual Environments," in *Automotive Systems Engineering II*, H. Winner, G. Prokop, and M. Maurer, Eds., Cham: Springer International Publishing, 2018, pp. 147–175.
- [16] M. Grindal, J. Offutt, and S. F. Adler, "Combination testing strategies: a survey," *Software Testing, Verification and Reliability*, vol. 15, no. 3, pp. 167–199, 2005.
- [17] T. L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, 1st ed. New York: McGraw-Hill, 1980.
- [18] Q. Xia, J. Duan, F. Gao, T. Chen, and C. Yang, "Automatic Generation Method of Test Scenario for ADAS Based on Complexity," in *SAE Technical Paper*, 2017.
- [19] Q. Xia, J. Duan, F. Gao, Q. Hu, and Y. He, "Test Scenario Design for Intelligent Driving System Ensuring Coverage and Effectiveness," *International Journal of Automotive Technology*, vol. 19, no. 4, pp. 751–758, 2018.
- [20] W. Chen and L. Kloul, "An Ontology-based Approach to Generate the Advanced Driver Assistance Use Cases of Highway Traffic," in *10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 2018.
- [21] M. Huelsen, *Knowledge-Based Driver Assistance Systems*. Wiesbaden: Springer Fachmedien Wiesbaden, 2014.
- [22] A. Pütz, A. Zlocki, J. Küfen, J. Bock, and L. Eckstein, "Database Approach for the Sign-Off Process of Highly Automated Vehicles," in *25th International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Traffic Safety Administration*, 2017.
- [23] A. Pütz, A. Zlocki, J. Bock, and L. Eckstein, "System validation of highly automated vehicles with a database of relevant traffic scenarios," in *12th ITS European Congress*, Strasbourg, 2017.
- [24] L. Wang, F. Fahrenkrog, T. Vogt, O. Jung, and R. Kates, "Prospective safety assessment of highly automated driving functions using stochastic traffic simulation," in *25th International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Traffic Safety Administration, Detroit*, 2017.
- [25] D. Zhao, X. Huang, H. Peng, H. Lam, and D. J. LeBlanc, "Accelerated Evaluation of Automated Vehicles in Car-Following Maneuvers," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 3, pp. 733–744, 2018.
- [26] Z. Huang, D. Zhao, H. Lam, D. J. LeBlanc, and H. Peng, "Evaluation of Automated Vehicles in the Frontal Cut-in Scenario - an Enhanced Approach using Piecewise Mixture Models," *CoRR*, vol. abs/1610.09450, 2016.
- [27] Z. Huang, H. Lam, and D. Zhao, "Towards affordable on-track testing for autonomous vehicle — A Kriging-based statistical approach," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 1–6.
- [28] Z. Huang, H. Lam, and D. Zhao, "Sequential experimentation to efficiently test automated vehicles," in *2017 Winter Simulation Conference (WSC)*, 2017, pp. 3078–3089.
- [29] M. Althoff and J. M. Dolan, "Online Verification of Automated Road Vehicles Using Reachability Analysis," *IEEE Transactions on Robotics*, vol. 30, no. 4, pp. 903–918, 2014.
- [30] M. Althoff and S. Lutz, "Automatic Generation of Safety-Critical Test Scenarios for Collision Avoidance of Road Vehicles," in *Proc. of the IEEE Intelligent Vehicles Symposium*, 2018, pp. 1326–1333.
- [31] S. Manzinger, M. Leibold, and M. Althoff, "Kooperative Bewegungsplanung autonomer Fahrzeuge unter Verwendung von Manöver-Templates," in *AAET-Automatisiertes und vernetztes Fahren*, 2017.
- [32] F. Gruber and M. Althoff, "Anytime Safety Verification of Autonomous Vehicles," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 1708–1714.
- [33] M. Koren, S. Alsaif, R. Lee, and M. J. Kochenderfer, "Adaptive Stress Testing for Autonomous Vehicles," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018, pp. 1–7.
- [34] B. Graab, E. Donner, U. Chiellino, and M. Hoppe, "Analyse von Verkehrsunfällen hinsichtlich unterschiedlicher Fahrerpopulationen und daraus ableitbarer Ergebnisse für die Entwicklung adaptiver Fahrerassistenzsysteme," in *3. Tagung Aktive Sicherheit durch Fahrerassistenz*, 2008.
- [35] C. Amersbach and H. Winner, "Functional Decomposition: An Approach to Reduce the Approval Effort for Highly Automated Driving," in *8. Tagung Fahrerassistenz*, 2017.

- [36] C. Amersbach and H. Winner, “Funktionale Dekomposition - Ein Beitrag zur Überwindung der Parameterraumexplosion bei der Validation von höher automatisiertem Fahren,” in *12. Workshop Fahrerassistenzsysteme*, Walting, 2018.
- [37] D. Heß, J. Oehlerking, M. Woehrlé, and J. S. Cubillo, “The UnCoVerCPS Verification Approach to Automated Driving,” in *20th International Forum on Advanced Microsystems for Automotive Applications*, 2016.
- [38] M. Kamali, L. A. Dennis, O. McAree, M. Fisher, and S. M. Veres, “Formal verification of autonomous vehicle platooning,” *Science of Computer Programming*, vol. 148, pp. 88–106, <http://www.sciencedirect.com/science/article/pii/S0167642317301168>, 2017.
- [39] S. Shalev-Shwartz, S. Shammah, and A. Shashua, “On a Formal Model of Safe and Scalable Self-driving Cars,” *CoRR*, vol. abs/1708.06374, <http://arxiv.org/abs/1708.06374>, 2017.
- [40] European Commission, *New and improved car emissions tests become mandatory on 1 September*. [Online] Available: http://europa.eu/rapid/press-release_IP-17-2822_en.htm. Accessed on: Feb. 19 2019.
- [41] UNECE, *Regulation No. 83 Revision 5: Uniform provisions concerning the approval of vehicles with regard to the emission of pollutants according to engine fuel requirements*. [Online] Available: <https://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/R083r5e.pdf>. Accessed on: Feb. 19 2019.
- [42] UNECE, *Global Technical Regulation No. 15: Worldwide harmonized Light vehicles Test Procedure*. [Online] Available: https://www.unece.org/trans/main/wp29/wp29wgs/wp29gen/wp29glob_registry.html. Accessed on: Feb. 19 2019.
- [43] UNECE, *UNECE works on global methodology to measure on-road car emissions*. [Online] Available: <http://www.unece.org/?id=48872>. Accessed on: Feb. 19 2019.
- [44] Working Party on Pollution and Energy, *Proposal for Amendment 5 to global technical regulation No.15 (Worldwide harmonized Light vehicles Test procedures (WLTP))*. [Online] Available: <https://www.unece.org/fileadmin/DAM/trans/doc/2019/wp29grpe/ECE-TRANS-WP29-GRPE-2019-02e.pdf>. Accessed on: Feb. 19 2019.
- [45] European Commission, *Emissions in the automotive sector: Emissions from light-duty vehicles*. [Online] Available: https://ec.europa.eu/growth/sectors/automotive/environment-protection/emissions_en. Accessed on: Feb. 19 2019.
- [46] UNECE, *Regulation No. 79 Revision 4: Uniform provisions concerning the approval of vehicles with regard to steering equipment*. [Online] Available: <https://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/2018/R079r4e.pdf>. Accessed on: Feb. 19 2019.
- [47] NHTSA, *Federal Automated Vehicles Policy: Accelerating the Next Revolution In Roadway Safety*. September 2016. [Online] Available: <https://www.transportation.gov/sites/dot.gov/files/docs/AV%20policy%20guidance%20PDF.pdf>. Accessed on: Feb. 18 2019.
- [48] K. Dietmayer, “Predicting of Machine Perception for Automated Driving,” in *Autonomous Driving: Technical, Legal and Social Aspects*, M. Maurer, J. C. Gerdes, B. Lenz, and H. Winner, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 407–424.
- [49] W. Han, W. Wang, X. Li, and J. Xi, “Statistical-based approach for driving style recognition using Bayesian probability with kernel density estimation,” *IET Intelligent Transport Systems*, vol. 13, no. 1, pp. 22–30, 2019.
- [50] Björn Filzek, “Abstandsverhalten auf Autobahnen – Fahrer und ACC im Vergleich,” Dissertation, Technische Universität Darmstadt, Darmstadt, 2002.
- [51] N. G. Leveson and J. P. Thomas, *STPA Handbook*. [Online] Available: http://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf. Accessed on: Feb. 19 2019.
- [52] UNECE, *Guideline for Intelligent & Connected Vehicle Autonomous Driving Functionality Test*. [Online] Available: <https://wiki.unece.org/download/attachments/63309804/SG1-03-11%20China.docx?api=v2>. Accessed on: Feb. 19 2019.
- [53] J. Bach, S. Otten, and E. Sax, “Model based scenario specification for development and test of automated driving functions,” in *2016 IEEE Intelligent Vehicles Symposium (IV)*, 2016, pp. 1149–1155.
- [54] H. Schaub, “Exception error: über Fehler und deren Ursachen beim Handeln in Unbestimmtheit und Komplexität,” *gdi impuls*, vol. 4, no. 96, pp. 1–12, 1996.
- [55] F. Schuldt, “Ein Beitrag für den methodischen Test von automatisierten Fahrfunktionen mit Hilfe von virtuellen Umgebungen,” Dissertation, Technische Universität Braunschweig, Braunschweig, 2017.
- [56] Nanyang Technological University, *NTU, LTA and JTC unveil Singapore’s first autonomous vehicle test centre*. [Online] Available: <http://media.ntu.edu.sg/NewsReleases/Pages/newsdetail.aspx?news=39308c90-536c-4c3a-be6d-b9c07041a442>. Accessed on: Feb. 19 2019.