DIGITAL INDUSTRIES SOFTWARE

Scenario-based verification and validation of self-driving vehicles: relevant safety metrics

Executive summary

The field of automated driving systems (ADS) has rapidly evolved over the past years. Press releases have continuously announced new technical achievements and partnerships among original equipment manufacturers (OEMs), startups and mobility service providers. Self-driving technology that operates in urban areas where customer demand is high, such as autonomous shuttles, robot-taxis and autonomous delivery are in the focus of new developments.

Some of the main challenges the companies developing these new innovative solutions face, are how to prove safety compliance, meet regulatory requirements and achieve user acceptance. In this white paper, we have partnered with IVEX to explain how to accelerate the safety validation of self-driving vehicles considering a scenario-based approach in combination with well-defined safety metrics. Furthermore, we highlight the integrated software (SW) toolchain from Siemens Digital Industries Software in combination with IVEX's Safety Analytics tool solution, which enables this acceleration.

Contents

How scenario-based validation is performed

Safety plays a central role in our daily lives and safety assurance has a long history in different industrial sectors (oil and gas, nuclear energy, manufacturing, etc.). However, safety assurance of complex systems operating in complex environments (for example, self-driving vehicles) is a challenging and continuously evolving domain.

Currently, for an ADS, safety is assured by safety design. This refers to how we define, design, develop and deploy systems and solutions.

Safety has different aspects, which are covered by different standards. Some of them are region-specific (for example, traffic rules) and some of them are more general and widely used by the automotive industry, such as the International Organization for Standardization (ISO) 26262, Safety of the Intended Functionality (SOTIF), ISO/DIS 21448, ISO 5083, ISO 22737, etc.

One of the objectives of SOTIF, is to validate the automated driving function in all relevant scenarios, especially in difficult conditions for sensors and algorithms. According to SOTIF, a scenario is a description of the temporal relationship between several scenes in a sequence, with goals and values in a specified situation that is influenced by actions and events. Every scenario starts with an initial scene. Actions and events, as well as goals and values, can be specified to characterize this temporal relationship in a scenario. In contrast to a scene, a scenario spans over a certain amount of time.[1]

SOTIF sends a clear message: the operational design domain (representing the driving conditions the autonomous vehicle will safely operate in) and the scenario space will be explored. Furthermore, during scenario-based safety validation, a statistical safety argument must be built.[2,3] In this sense, the scenarios can be directly controlled in:

- A physical environment (closed course proving ground)
- A virtual environment (simulation of pre-defined or randomly generated scenarios)
- A spontaneous manner during operation in the real world (open-road testing)

SOTIF classifies the scenario space in four areas:

1. Known and non-hazardous (safe) scenario
2. Known and hazardous (unsafe) scenario
3. Unknown and hazardous (unsafe) scenario
4. Unknown and non-hazardous (safe) scenario

The safety goal of SOTIF is to explore the scenario space and gradually increase the known safe scenario's area during the development with an iterative approach. This is done by discovering unknown and unsafe scenarios and moving them to known and safe scenarios. They will be transformed into known and non-hazardous scenarios with proper technical measures (figure 1).

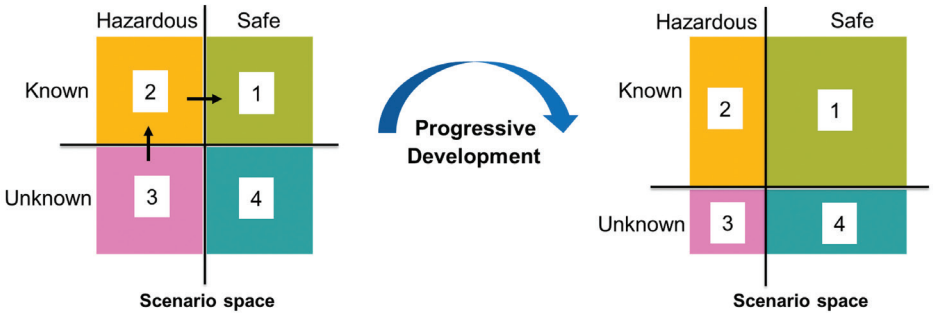

Figure 1. Scenario space evolution during progressive development – according to SOTIF.

These unknown and unsafe scenarios must be identified during scenario space exploration. This can be done by using simulations, which can help us find out edge and corner cases. A corner case is a scenario where two or more parameter values are each within the capabilities of the system, but together constitute a rare condition that challenges its capabilities. An edge case is a scenario where the extreme values or the presence of one or more parameters results in a condition that challenges the capabilities of the system.[1]

The methodology of scenario-based validation follows the steps depicted in figure 2. In the first phase, the desired scenarios (considering the operational design domain) must be extracted from driving data recorded in real-world, publicly accessible driving datasets or simulated with long runs. The extracted test scenario categories must fulfil the mandatory test requirements indicated by regulation authorities, specific driving situations that the manufacturer wants to test or must be discovered based on risk analysis and scenario-space exploration – for example, system-theoretic process analysis (STPA) and search algorithms. The scenario extraction objective is to generate the required scenario database from which test cases are generated and safety assessments are performed. Additionally, other kinds of scenario databases can be used to generate the test cases, such as accident databases or specifically recorded scenarios.

Once the scenario database is created, an analysis of the existing scenario must be performed to gain insight into the parametric scenario space covered and complement the database if needed.

When the database coverage is considered enough, these scenarios are extensively tested in the virtual world – model-in-the-loop (MIL), software-in-the-loop (SIL), hardware-in-the-loop (HIL) and driver-in-the-loop (DIL). If the driving functionality achieves the specific requirements, then field operational tests (FOT) in close circuits and real-world traffic situations are performed (figure 2). If new hazardous scenarios are discovered in these FOTs or the performance obtained does not meet the requirements, these specific scenarios can be extracted with the proper SW and re-simulated including more thorough variations. All the results are correlated and assessed using different safety metrics, such as the IVEX safety models and metrics, the Mobileye's responsibility sensitive safety (RSS), or NVIDIA's safety force field. If the results pass the certification and homologation criteria, then the certificates can be issued.

During each stage of the development, verification is performed in MIL, SIL and HIL. What do we need for reliable virtual verification and validation? We need at least two things: reliable SW tools and validated mathematical models. The next section will briefly present a possible SW toolchain, which can accelerate scenario-based validation.

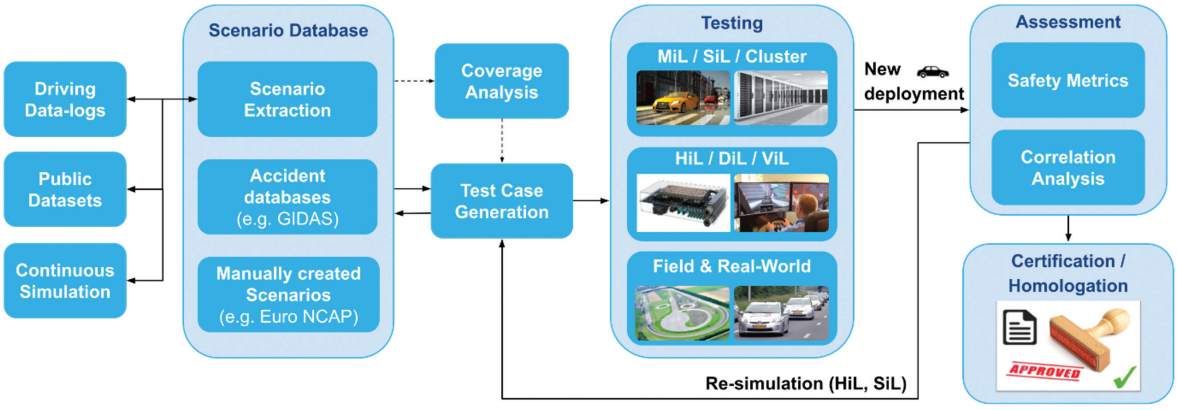

Figure 2. Scenario-based validation methodology.

Integrated software toolchain for scenario-based validation

Siemens offers an integrated SW environment for virtual verification and validation, which allows scenario imports from different databases, scenario space exploration and design optimization. This solution contains:

- Simcenter ™ Amesim™ software – for vehicle dynamics modeling and simulation[4]
- Simcenter Prescan™ software – for modeling the world (environment) and modeling of sensors[5]
- HEEDS™ software – for optimization and test automation, which allows the user to explore the design space and visualize the results[6]

We briefly discuss each toolkit we used and will later elaborate on the features of these toolkits, which were used in our experiments.

HEEDS is a design exploration and optimization tool. In a design problem where an optimal configuration of participating design variables is desired, HEEDS uses its meta-heuristic approach of design space exploration to arrive at an optimal or near optimal set of designs while complying to user-specified constraints and objectives. Additionally, HEEDS incorporates a Design of Experiments (DOE) study to sample and explore a region of the design space for analyzing design parameters, and in some cases, their influence on responses (objective and constraints).

With this functionality, HEEDS enables the user to sample the scenario space in a random or pseudo-random way or with adaptive sampling to reduce the number of explored scenarios and maintain a good test coverage. Later, we provide more details regarding the procedure we use to generate test cases using HEEDS and automate the process of scenario-space exploration.

IVEX offers the Safety Analytics tool that can ingest different kinds of driving information from simulation and real-world tests.[7] It is a comprehensive solution that comprises different modules enabling scenario extraction and categorization, extensive validation of ADS/advanced driver assistance

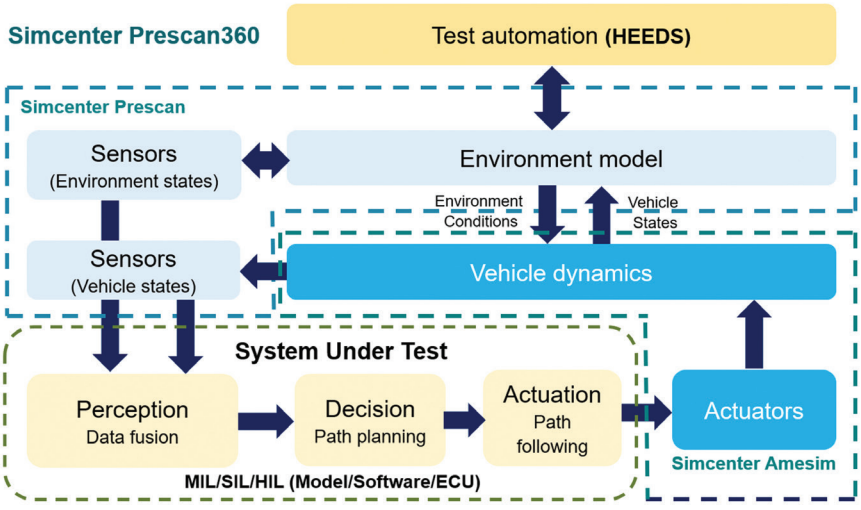

Figure 3. Integrated SW toolchain for scenario-based verification and validation from Siemens.

systems (ADAS) and detailed safety assessment of driving policies. IVEX's Safety Analytics tool includes:

- Scenario categorization – provides automatic scenario extraction and categorization from larger driving data, real-world recorded data-logs, long simulated runs or public datasets. The SW extracts and categorizes SOTIF known hazardous and known non-hazardous scenarios (for example, areas 1 and 2 in figure 1) and it enables the discovery of new unknown hazardous and non-hazardous scenarios (SOTIF areas 3 and 4) to be included in the scenario database
- Aggregated analysis – enables the selection of a large set of scenarios (for example, all the cut-ins in the database of scenarios) and provides an abstracted performance analysis according to several safety metrics. This way, the user will obtain grouped insights about the performance of the vehicle in a general manner
- Scenario inspection – enables the detailed inspection of the timestamps where a metric was triggered or a safety concern was raised. It provides insight into the problem, showing the full contextual information around the specific timestamp so the engineers can find the source of the error and work on a solution

The proper integration of these presented tools from Siemens and IVEX enable the engineers to seamlessly build strong scenario-based verification and validation cases. Simcenter Amesim, Simcenter Prescan and HEEDS are part of the Siemens Xcelerator portfolio, the comprehensive and integrated portfolio of software and services from Siemens. The next section presents a simple scenario use-case along with MIL scenario validation.

Scenario-based verification and validation – a simple use-case

In this section, a relatively simple urban-driving scenario is considered (figure 4). The use-case scenario is manually created using Simcenter Prescan, parametrized using HEEDS to obtain several variations and a safety assessment was performed using the IVEX safety models in its Safety Analytics tool.

This workflow was achieved with the integrated toolchain shown in figure 5. Siemens' toolchain communicates with the cloud-based IVEX Safety Analytics tool via a restful application program interface (API). This simplified workflow does not cover scenario database creation and has two main stages. The first one is a preparation stage, which is mainly manual and defines the scenes, the host controller and the vehicle dynamics, and compiles the model. The second stage considers the scenario parameterization (generating different scenario variations), the test execution, results generation and data visualization and reporting.

In the context of exploring the scenario space, we implement the evaluation-only type of DOE study in HEEDS. The primary objective is to maximize the coverage of scenario space by uniformly sampling points across the scenario space. The scenario space consists of design variables, which we also refer to as scenario parameters that alter the environment and participating objects of the scenario under consideration. The design points sampled by HEEDS

Scene contains:
1. Road section in urban area
2. Host vehicle, with radar and adaptive cruise control (ACC) plus autonomous emergency brake (AEB) system
3. Two cars in front of host, one of which can cut-out to evade the other (slower) car
4. Car next to host vehicle, driving same speed
5. Three parked cars, one of which can reverse onto road
6. Two pedestrians that can cross the road, one of which is occluded
7. Two busses on a bust stop, one of which can cut-in to road

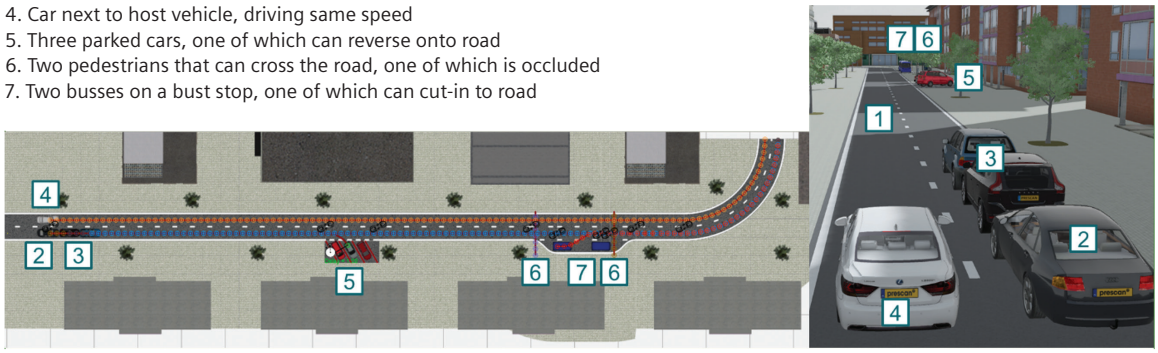

Figure 4. Scenario description using Simcenter Prescan – urban driving scenario.

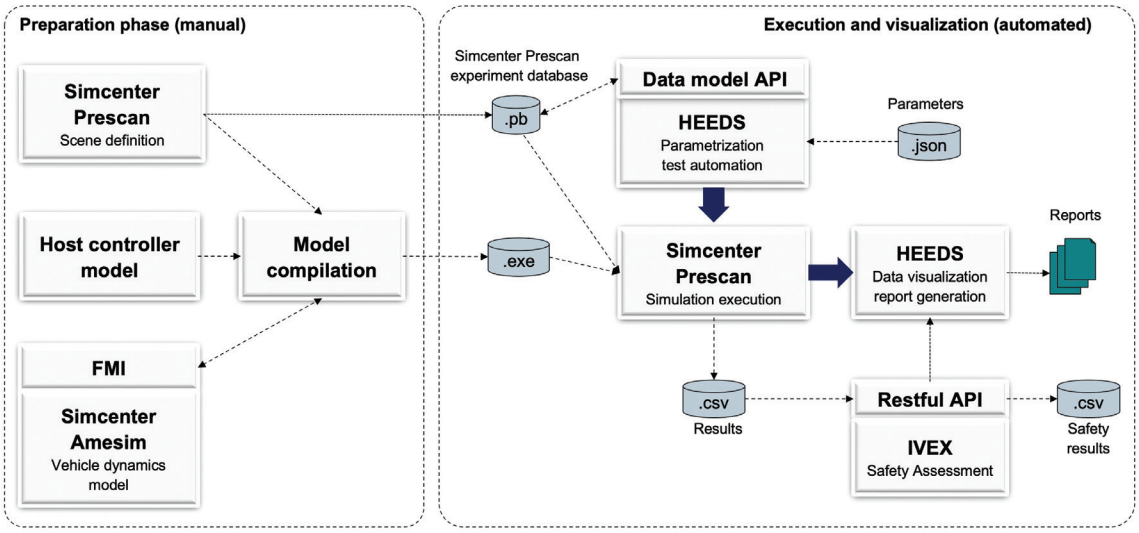

Figure 5. Siemens and IVEX integrated SW toolchain for scenario-based verification.

are then evaluated by running simulations using Simcenter Prescan and recording necessary responses.

Next, we discuss the underlying methods available in HEEDS to create design sets for launching a DOE study. Design sets in HEEDS can be generated by importing pre-existing data, sweeping variable values and by using space-filling techniques.

Importing designs from pre-existing datasets and prior experiments/studies helps us leverage the past information for design space exploration. This enhances the aspect of collaborative workflow where datasets from multiple sources can be merged to conduct a DOE study.

Variable sweeps are useful in generating structured datasets by varying the user-specified set of variables in a specified range.

The space-filling approaches can be used in conjunction with pre-existing datasets. The use of pre-existing datasets encourage collaboration in dataset generation. The three space-filling approaches implemented in HEEDS are uniform variable coverage (UVC), uniform design space coverage (UDSC) and balanced variable and design space coverage.

Uniform variable coverage

The UVC approach uses an optimal Latin hypercube sampling approach to uniformly sample data across the design space. This approach is more useful when we do not have any pre-existing datasets and we desire to uniformly sample points across the design space (figure 6). Here, the blue datapoints are loaded from a pre-existing dataset and red datapoints are generated using the UVC. Irrespective of the prior distribution of blue points, the approach used in UVC uniformly scatters the points in each domain.

Therefore, in case the pre-existing dataset is non-uniform, the global distribution of datapoints (pre-existing plus UVC) will also be non-uniform.

Uniform design space coverage

As the name suggests, the UDSC approach tends to create a uniformly scattered distribution in the entire design space. Unlike the UVC approach, the UDSC approach generates points while taking the pre-existing set of points (supplied or generated so far) into consideration. In the UDSC approach, each new design point that is farthest from the existing set of design points based on the Euclidean distance metric, is added. This process is repeated one design point at a time until total number of desired points are generated. The resulting global distribution (pre-existing plus UDSC) is more uniform across the entire design space (figure 7).

Here, the UDSC approach biased the creation of design points (red) in the central cavity region of the distribution of pre-existing datapoints (blue). Since each new generated point is farthest from the existing set of points, the UDSC approach tends to generate more points near the boundaries of the design space. This becomes more evident for higher dimensional design spaces.

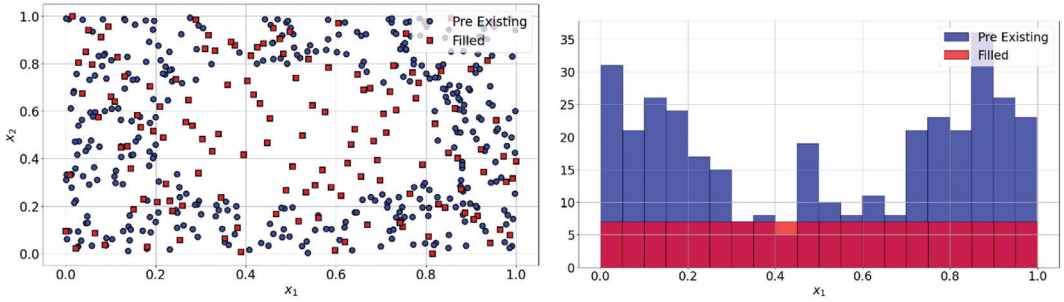

Figure 6. Design set generated using UVC.

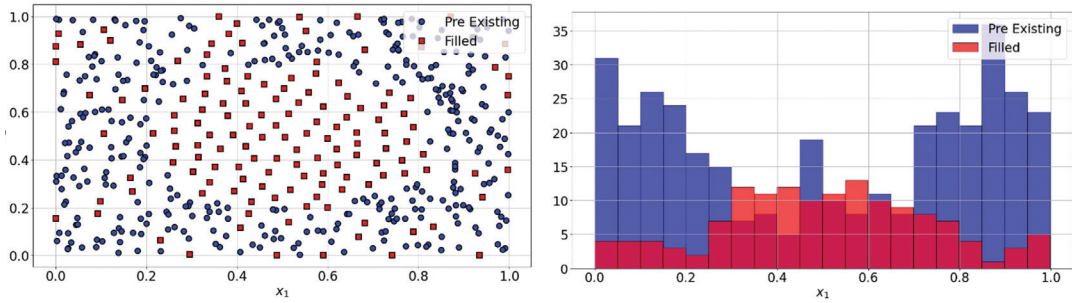

Figure 7. Design set generated using the UDSC method.

Balanced variable and design space coverage
This approach is a modification to the UDSC approach. Here, a new datapoint is generated by maximizing its Euclidian distance from the existing set of datapoints while avoiding the creation near the boundary. Thus, this approach is particularly useful when excessive boundary points are not desired. Resulting plots for a 2D case are shown in figure 8.

In figures 7 and 8, we see that less number of boundary points are created using the balanced variable and design space coverage approach than the UDSC approach. This difference is further highlighted for higher dimensional design spaces.

In summary, if no prior dataset is supplied then the UVC method can be used, and either of the other two approaches can be invoked. Additionally, the UDSC approach should be preferred if more datapoints near the boundary of the design space are desired. In our experiments, we have used the balanced variable and design space coverage approach to sample and explore the scenario space. Besides creating design sets, a DOE with adaptive sampling can also be used. Adaptive sampling is an active learning technique that combines various learning strategies to choose the sample points. It is an iterative sampling approach that can be tuned during the learning process. This way, a balance is created between learning about the global and local characteristics that can provide more insight into the more critical regions of the design space.

Currently, the adaptive sampling strategy in HEEDS balances three different learning algorithms in its search. Firstly, global spread sampling gets a more uniform coverage of the design space by choosing samples from holes in the design space. Secondly, accuracy improvement sampling uses five internal surrogate models to identify regions with higher error estimates and samples in those regions to reduce the fit error. Lastly, non-linearity sampling identifies regions of high curvature using the best of the five surrogate models and improves local fitting. Starting with a small optimal Latin hypercube, the search cycles use the three algorithms with varying weights. Using this strategy, a better understanding of the critical scenarios can be achieved.

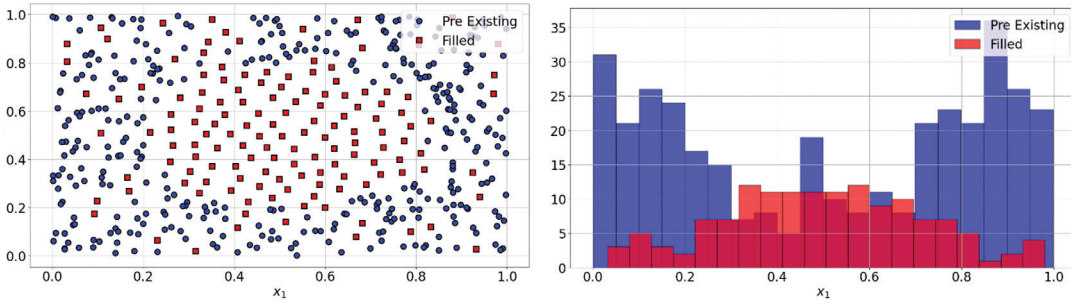

Figure 8. Design set generated using the balanced variable and design space coverage method.

Safety assessment using IVEX's Safety Analytics tool

IVEX's Safety Analytics tool enables scenario safety assessment based on safety models. These models can evaluate the quality of the sensor data, analyze the perception performance, assess the driving policy safety as the behavior of the vehicle, etc.

For the use-case under study, IVEX applies different behavior-safety models to measure the safety of the vehicle-under-test. Specifically, one of the safety models used on this use-case assessment is based on assumptions of a set of kinematics parameters, such as the maximum acceleration, the maximum deceleration or the response time of the ego and other road users. The metrics used in this specific model determine how safe the trajectory of the vehicle is, given a set of assumptions on the worst-case behavior of other road-users. Instead of just checking whether any collision happened or would happen on the chosen trajectory, the model assesses the safety of the vehicle by checking if it violates any safety constraints.

IVEX's safety model continuously calculates the required action for a safety performance. In this case, the vehicle action under consideration is the maximum deceleration needed. So, even when other road users perform their worst-case maneuver (for example, a sudden brake or lane change), the ego vehicle would still be able to decelerate safely. Some set of safety constraints limit the required action.

For example, if the required deceleration to avoid a possible collision is large and the vehicle does not perform any action to avoid the situation, a safety violation is triggered.

Table 1 shows the mean-time to a first violation. This is the mean duration in seconds from the first timestamp of a trajectory to the timestamp where the first safety violation occurs according to IVEX's safety model.

A short mean-time to first violation does not mean the vehicle is in danger or poses a danger to other road users. The mean-time to first violation shows how much time the ego vehicle has to react, in the case that the other road users perform its worst-case maneuver. In table 1, we can see that the vehicle has 0.75 seconds on average, to react to worst-case behavior of other cars.

In scenarios involving a cut-in from a passenger car, the duration to the first safety violation is shorter than in scenarios where the cut-ins are performed by bus. This could mean that the driving policy of the ego vehicle is more cautious to cut-ins from big vehicles, such as buses, than to a car. We also see that in cut-out situations, the time-to-violation is very short, which means the behavior of the system could be not appropriate in cut-out cases.

Scenario	Mean-time to first violation [s]
All scenarios	0.72
Cut-in – bus	1.01
Cut-in – car	0.75
Cut out	0.48
Pedestrian	0.99
Pedestrian occluded	1.07

Table 1. IVEX safety metric – mean-time to first violation.

Conclusion and remarks

In this white paper, Siemens and IVEX briefly mention the relevant standards applicable during the development and deployment of an ADS such as ISO 26262, ISO/DIS 21448, ISO 5083 and ISO 22737. Furthermore, the scenario-based validation methodology has been described and the integrated SW toolchain from Siemens and IVEX, which can support accelerated development, has been presented. Finally, details about the main features of the SW tools such as Simcenter Prescan, Simcenter Amesim, HEEDS and IVEX's Safety Analytics tool have been provided, considering a simple urban-driving scenario as a use-case for scenario-based validation. The integrated SW toolchain, the presented methodology and workflow fully supports scenario-based validation of an ADS and contributes to the safety-case argumentation based on statistical evidence, according to ISO 5083 as well as ISO/DIS 21448 and ISO/TR 4804.

References

1. ISO/DIS 21448: Road vehicles – Safety of the intended functionality, edition 2021.
2. ISO/TR 4804: Road vehicles - Safety and cybersecurity for automated driving systems - Design, verification and validation, first edition, 2020-12.
3. ISO/DIS 22737: Intelligent transport systems – Low-speed automated driving (LSAD) systems for predefined routes – Performance requirements, system requirements and performance test procedures, edition 2020.
4. Simcenter Amesim: https://www.plm.automation.siemens.com/global/en/products/simcenter/simcenter-amesim.html, accessed in November 2021.
5. Simcenter Prescan: https://www.plm.automation.siemens.com/global/en/products/simcenter/prescan.html, accessed in November 2021.
6. Simcenter HEEDS: https://www.plm.automation.siemens.com/global/en/products/simcenter/simcenter-heeds.html accessed in November 2021.
7. IVEX Safety Analytics Tool: https://www.youtube.com/watch?v=F0domhEWNP0&t=22s accessed in November 2021.
8. IVEX Safety co-pilot: https://www.ivex.ai/products/safety-co-pilot accessed in November 2021.

Siemens Digital Industries Software

Americas: 1 800 498 5351

EMEA: 00 800 70002222

Asia-Pacific: 001 800 03061910

For additional numbers, click here.

Siemens Digital Industries Software helps organizations of all sizes digitally transform using software, hardware and services from the Siemens Xcelerator business platform. Siemens' software and the comprehensive digital twin enable companies to optimize their design, engineering and manufacturing processes to turn today's ideas into the sustainable products of the future. From chips to entire systems, from product to process, across all industries, Siemens Digital Industries Software – Accelerating transformation.

About IVEX

IVEX is an international team working on extremely challenging and interesting technical problems that have a huge societal impact. IVEX's unique IP and team experience in behavior planning, functional safety and vehicle engineering enables the development of safe autonomous vehicles. https://www.ivex.ai/

siemens.com/software

84312-D6 3/22 A