Introducing VeriCAV: Addressing the challenges of virtual testing of Automated Vehicles.

Thomas Levermore [1], Alan Peters [1]
1. Connected Places Catapult, United Kingdom

Abstract

Virtual testing in simulation is crucial to the safety assurance of highly automated vehicles, however significant challenges remain. Challenges relating to the automation of virtual testing in simulation, test generation and test analysis are described here, and potential routes to solutions proposed. Based on the VeriCAV project, a system architecture for virtual verification of highly automated vehicles is described with a focus on the interfaces between test framework elements. The framework includes elements for test generation and assessment, simulation, realistic actors and the automated driving system being tested.

Keywords:
Scenario based testing, virtual verification, automated vehicles

1. Introduction

It is widely accepted that highly automated vehicles (HAVs) will need to be verified and validated using a combination of virtual testing in simulation, closed-road testing and public road trials [1, 2, 3]. This combination is described by the UNECE working group on *Validation Method for Automated Driving* (VMAD) as the “multi pillar approach” [4].

Whilst necessary, physical testing is expensive and time-consuming, particularly when considering the extent of testing required to demonstrate sufficient safety. There is therefore the need to maximise use of virtual testing. Early testing of the Automated Driving Systems (ADSs) that control HAVs is already undertaken solely in simulation due to the cost of prototype systems and the consequences of failure in a physical test. The need for virtual testing to form a larger proportion of the system development lifecycle is increased as ADS software development is becoming increasingly iterative and changes to software can occur regularly with practices such as Continuous Integration [5] resulting in more frequent repeat testing. Virtual testing can identify design flaws earlier in the development cycle than physical testing.

The use of virtual testing in simulation to verify and validate HAVs can include a range of techniques, noted in [6] as:

- Software-in-the-loop (SIL), which includes some or all of the ADS software running with the remaining elements of the final HAV system modelled in the simulation
- Hardware-in-the-loop (HIL), which includes running the software on some or all of the target hardware and/or sensors
- Vehicle-in-the-loop (VIL), which includes the physical vehicle in the test, generally on a rolling road.

The use of virtual testing in simulation is proposed in the context of scenario-based testing [7, 8]. A scenario, as defined in [9] can be described by road information, stationary objects, movable objects and their movements and environment conditions. Testing in this way allows challenging, rare, recently observed and/or standardised sets of scenarios to be the focus of ADS testing without having to go through the extended periods of routine, unchallenging driving that exist in on-road testing.

It has been suggested that the amount of on-road testing required to demonstrate the safety of a HAV is in the order of hundreds of millions miles [10], if not more. This brute force approach is clearly not feasible on the road, but a combined approach that includes simulation makes extensive testing possible by exploiting parallelism [3]. Even with simulation, testing requires significant effort and resource and thus reducing this requirement is necessary. Equally, targeted scenario generation is necessary as opposed to a brute force approach that searches for *any* scenario that causes a failure, known as falsification [11].

The Verification of Connected and Autonomous Vehicles (VeriCAV) project addresses challenges that still exist in virtual testing to enable the combined testing approach to ultimately lead to verified and validated HAVs. These challenges are outlined in the following sections. The VeriCAV project focuses on the SIL technique and to a limited extent HIL but the concept does not preclude VIL testing. While the focus of the project is on HAVs, the approach is equally applicable for lower levels of automation and advanced driver assistance systems (ADAS).

Paper Contribution
The contributions of this publication are as follows:

- A summary of the challenges facing the virtual testing in simulation element of an ADS testing programme.
- A summary of scenario-based testing for ADSs.
- A framework for virtual testing of ADSs.
- A summary of outstanding research questions related to virtual testing of ADSs.

2. VeriCAV Project

VeriCAV[1] is a £3m, 2-year collaborative research project involving a partnership of four organisations: Connected Places Catapult, HORIBA MIRA (as industry lead), the University of Leeds, and Aimsun. VeriCAV is part funded by the Centre for Connected and Autonomous Vehicles via Innovate UK.
The VeriCAV project focuses on the use case of a central validation service to be used by different ADS developers, Original Equipment Manufacturers and approval authorities. Within the context of this use case the project addresses some of the key challenges associated with virtual testing of HAVs:

- The amount of human effort required in the creation, running and analysis of effective test cases prevents scaling testing to the necessary level.
- The realism of simulation actors' behaviour is limited.
- To establish an interface between a new ADS and a simulation tool is complex, often bespoke and not reusable.

The focus of the project is on the decision-making of the ADS but the approach is designed to be expanded to include other aspects of ADS functionality.

3. Six Key Simulation Challenges

The following sections of this paper will expand and address these challenges as outlined below:

1. How can scenarios be described in a way that:
    a. captures the complexity of diverse scenario sets,
    b. allows automated scenario creation,
    c. and can be unambiguously interpreted by simulation tools?
2. How can the onerous activity of manual scenario generation be automated and made more efficient and targeted?
3. How can sufficient validity and fidelity of the simulation be ensured? For example, the realism of simulation actors.
4. How can the use of the virtual testing system be maximised across a range of different ADSs?
5. How can virtual testing coverage be ensured in an efficient manner?
6. How can test success be assessed?

Each of these challenges maps to a functional sub-system of the VeriCAV system architecture which is outlined in the following section for context.

4. Framework

The VeriCAV system is made up of several complex sub-systems developed by the consortium partners outlined in Figure 1. The VeriCAV system architecture includes core components:

- a Test Generator, to produce scenarios to test,

---

[1] https://vericav-project.co.uk/

- a Simulation Master, to interface a simulation tool to the rest of the system,
- a Test Oracle to analyse test results and direct the test generator.

Additionally, the VeriCAV system architecture includes components that can be exchanged:

- a Simulation Tool, in which the virtual scenario unfolds,
- a Smart Actor Controller, to provide realistic actor behaviour
- the System under Test, the ADS.

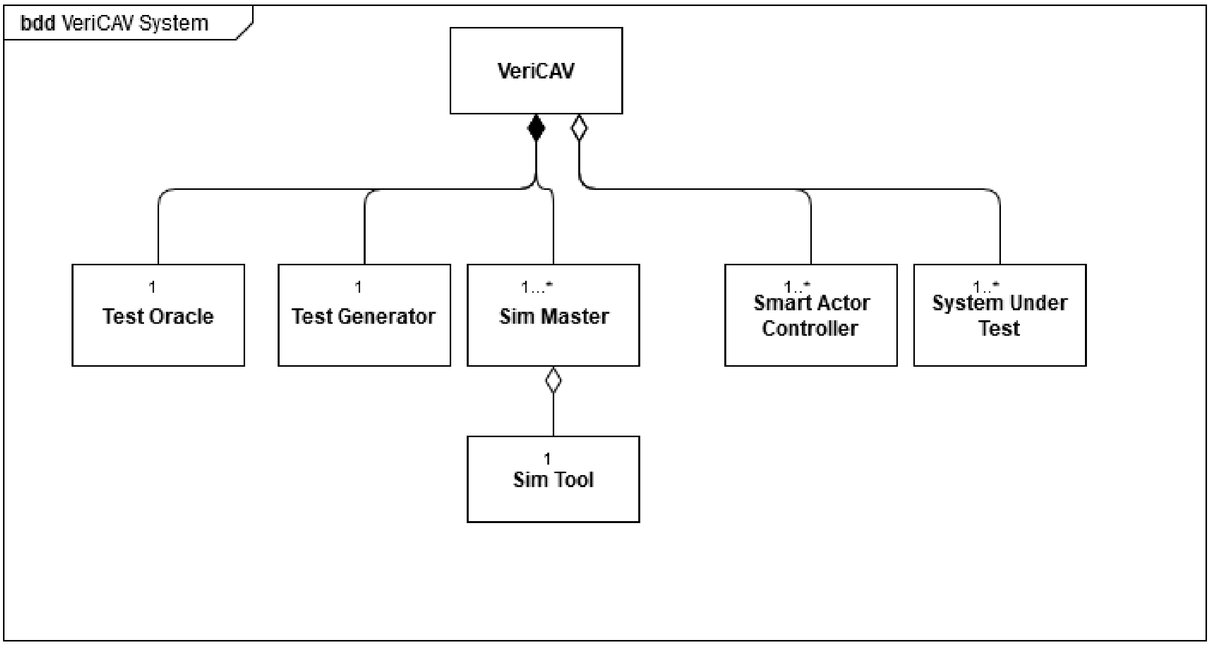

Figure 1 – Main components of the VeriCAV system expressed in SysML. The components that compose the core of the VeriCAV framework are shown on the left, with black diamond connector. The aggregated components (those with white diamonds) are modular components for which the framework is designed to be agnostic to their selection (i.e multiple simulation tools).

Each of the sub-systems in the VeriCAV will be described in subsequent sections. Firstly however, the common thread across all the sub-systems is the use of scenarios, so the following section will outline the concept of scenario-based testing and review the state of the art.

5. Scenario-based testing

The term scenario is in regular non-technical use and as such needs defining more precisely when used in the technical domain of scenario-based testing of ADSs. A definition of scenario in the context of ADS testing was proposed in [12] as:

> "*A scenario describes 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*"

Where a scene describes the environment, static and dynamic elements and actors at an instance in time. While these commonly used definitions, a more detailed technical definition from [13] has also become widely used. This definition divides a scenario into five layers. The first layer captures the basic road information such as geometry, while the second layer captures road infrastructure such as traffic signs and more abstract information such as road speed limit. The third layer captures temporary adjustments to the first two layers, such as road works which change lane configurations. The fourth layer captures objects in the scenario that are not part of the road infrastructure such as vehicles and pedestrians and these can be static or dynamic with interaction between objects described. The fifth layer captures environmental factors in the scenario such as weather and lighting conditions. This definition can be extended to include vehicle state such as load, connectivity related conditions such as vehicle-to-vehicle communication [14] and the wider scenario environment such as tunnels.

These layers give a structure to build a scenario conceptually but to actually get a simulation tool to execute the scenario as expected requires a machine-readable description language.

Formats

There are several **scenario description language (SDL)** formats that have been developed in recent years, primarily for virtual testing in simulation of ADSs. OpenSCENARIO is a well-supported format for describing the dynamic content of a scenario and works in conjunction with the OpenDRIVE format which describes road networks. Both of these formats use Extensible Markup Language (XML) and are now managed by the Association for Standardisation of Automation and Measuring Systems (ASAM) and open to the public. Under ASAM management both standards have been developed with significant industry backing and updated in 2020, while future development of these standards is on-going with objectives very much aligned with the scenario-based testing of HAVs.

There are a number of competing and complimentary formats for describing scenarios, some from commercial organisations such as the Measurable Scenario Description Language [15] from Foretellix Ltd and several from academia. These are summarised in Table 1.

Table 1

SDL	**Data format**	**License**	**Road Description**	**Simulator compatibility**
CommonRoad [16]	XML	3-clause BSD	Lanelets [17]	SUMO [18]
OpenSCENARIO	XML	Basic, non-exclusive and unlimited	OpenDRIVE*	Prescan [19], Vires VTD [20], CARLA [21].
SCENIC [22]	Probabilistic Programming Language	3-clause BSD	OpenDRIVE, OSM, *	GTAV [23], Webots [24], LGSVL [25]
Measurable Scenario Description Language	Programming Language	Apache2.0	OpenDRIVE*	CARLA
GeoScenario [26]	XML		Lanelets	Unreal Engine Plugin
Traffic Sequence Charts [27]	Visual / Formal Logic (conversion to OpenSCENARIO possible)		As per conversion applied	

* These SDLs are not tied to one particular road description, but the road descriptions listed are provided as options by the SDL authors.

Relevant projects that have publicised the SDL they are using include:

- PEGASUS [28] using OpenSCENARIO
- TNO Streetwise [29] using OpenSCENARIO
- Enable-S3 [30] includes using OpenSCENARIO and Traffic Sequence Charts
- MUSICC [31] using OpenSCENARIO with extensions.

There are a number of projects such as Simulation of Autonomous Vehicle Safety (SVA) [32], Open Autonomous Driving Accelerator (OpenADx) [33], SafetyPool [34] which have not publicised what SDL they are using. Additionally, there are several commercial services offering scenario-based ADS testing such as Metamoto [35], Cognata [36] and Alpha Drive [37] all of which can be cloud based services. Focused solely on validating path planning algorithms, Aimsun offer the scenario-based testing tool Aimsun Auto [38].

For the VeriCAV project, the OpenSCENARIO format is used due to the community support, the tool support, relative maturity of the standard and documentation and prior consortium experience.

Scenario Management

A machine-readable SDL allows simulation tools to execute scenarios without manual intervention as required to overcome Challenge 1.b. However, the identification of a scenario to test, the creation of this in the SDL and analysis of the results are all time-intensive activities to undertake manually. These functions in the VeriCAV project are undertaken automatically by the combination of a Test Generator and a Test Oracle.

6. Test Generator

The automated generation of test scenarios involves programmatically creating and combining the elements that make up a suitable scenario. The complexity in this process comes from the size of the parameter space in terms of both variance within individual parameters and number of parameters or dimensions of the parameter space – "the curse of dimensionality" [39]. Using a base set of scenarios and randomising scenario parameters within defined ranges, known as "fuzzing" [11, 7] to generate mutated scenarios introduces variation but can produce test scenarios that are physically unrealisable or extremely unlikely to occur in the real world. Such scenarios would be redundant, similarly for scenarios that are outside of the designated Operational Design Domain (ODD) of the ADS [40]. A caveat to this being that it is important to test the behaviour of the ADS at either side of the boundaries of the ODD. While avoiding manual creation of scenarios or significant use of manually created heuristics will minimise the human workload in the testing process it is crucial that the scenarios generated automatically are valid and insightful.

As the number of generated tests undertaken increases, information from the Test Oracle about the performance of the ADS in different situations can be used to produce scenarios specifically targeted to test the particular ADS. Adversarial testing such as this is commonly used for testing vision-based machine learning algorithms [41] but the principles can be applied to scenarios.

A challenge for the Test Generator is then to balance the exploitation of scenarios that are known to challenge ADSs and are valid and exploring novel scenarios that may provide useful knowledge of the particular ADS-under-test performance [42]. This area of research as encapsulated in Challenge 5 is being explored in the VeriCAV project by HORIBA MIRA. Once the Test Generator has produced a scenario to test, the Simulation Master is then responsible for executing it.

7. Simulation Master and Tool

Central to the VeriCAV framework is the Simulation Master which controls the simulation tool and interfaces to the ADS and Smart Actors.

Simulation Master

The Simulation Master is software developed by the VeriCAV project to allow the selection of a simulation tool and/or SUT to be decoupled from the rest of the system. There are a number of benefits to this and it is a key enabler to overcome challenges 3 and 4. To overcome challenge 3, it is expected that the fidelity of simulation tools will improve in the coming years and so the flexibility to upgrade this element of the system while minimising the disruption to the rest of the system is a key objective of the VeriCAV project. Both commercial off-the-shelf and open-source simulation tools are to be deployed in the VeriCAV framework to demonstrate this functionality. A similar need to interchange exists for the **ADS-under-test (ADSUT)** to overcome challenge 4. The open-source ADS, Baidu Apollo [43] is used as the primary ADSUT in the VeriCAV project with other ADSs being considered during interface design decisions.

The Simulation Master must ensure that there is sufficient temporal and spatial synchronicity across the distributed elements of the simulation: the ADSUT, the Smart Actors and the Simulation Tool. Ensuring temporal synchronicity in a distributed simulation usually involves a time management approach where advancing in time is controlled by the Simulation Master [44]. To test an ADS as it would normally operate, this time management is not possible. Publicly available information on ADS software architecture, including the Baidu Apollo Auto [43] and Autoware systems [45], shows they use a publish subscribe messaging system. This decouples the modules that make up the system, allowing communication to be maintained across distributed modules while being robust to timing variations (within limits).

As well as temporal synchronisation, another challenge for VeriCAV is spatial synchronisation. The simulation tool is responsible for providing positions of all the relevant objects in the virtual environment to both the ADS and the Smart Actor Controller. This is further complicated by different formats for describing position, with an ADS generally receiving its position in a global reference frame (latitude and longitude) from a positioning system involving GNSS whereas the Smart Actors are operating in a local reference frame. The format for defining position is also relevant for the base maps used by each of the ADS and the Smart Actors to ensure spatial synchronisation between reference frame and maps.

The use of an open, common, format of road layout such as OpenDRIVE would overcome some of these issues but a testing framework must allow for a variety of map formats that ADSs might use. Additionally, most ADSs would require more information than contained in an OpenDRIVE file. The Apollo Auto ADS, for instance, uses a bespoke High Definition Map format which, while based on the OpenDRIVE standard [46], does deviate from this, particularly in the representation of road geometry. Interfaces need to be clearly defined in geospatial terms and the accuracy of conversions needs to be taken in to account to ensure that all elements of the system are geospatially synchronised.

In line with Challenge 4, future adaptability of the Simulation Master to different ADSs, close alignment with developing standards in the area of sensor model interfacing such as the Open Simulation Interface (OSI) is actively pursued.

Simulation Tool

The role of the Simulation Tool is to allow the ADSUT to operate in a model of the real-world. While perception systems can be trained and tested on real sensor data [47, 48, 49] such as camera, lidar and radar data, a significant number of the ADS modules cannot be tested in this way. Data captured in the real-world cannot be de-coupled from the trajectory the data-capture vehicle followed nor the on-board sensor configuration, and therefore cannot be used to test an ADS that can deviate from this trajectory or differs in sensor configuration. This is in addition to the resources required to collect such data. Simulation can provide a solution to this by creating realistic environments to test the ADS in and maintain this environment regardless of the ADS or sensor position within the environment.

This requires the Simulation Tool to have a physics model for both vehicle dynamics and the environment, sensor models and visually and sensorially realistic environments. However, for VeriCAV the Simulation Tool is not the innovation, it is an enabler for the other functions to be demonstrated.

There are many commercial and open-source simulation tools for virtual ADS testing, [50]. VeriCAV is using the PreScan commercial simulation tool from Siemens to test the framework concept. This software is a mature product, well-used in the automotive industry and representative of the type of simulation tool ADS development teams would use in automotive manufacturing organisations. This simulation tool fulfils the requirements for physics, sensor and environment models however the VeriCAV project is specifically not focusing on sensor models in the ADS testing. This is to constrain the scope of the project as sensor modelling at a physical level is a complex task [51], and there is a lack of publicly available data to validate sensor models against the sensors deployed on target HAVs. The perception algorithms used by ADSs are tightly coupled to the sensor model they are trained with and so would require a representative sensor model to ensure valid virtual test results. Sensor models are an area of ongoing research that will continue beyond the timescale of the VeriCAV project. The scope, therefore, of the VeriCAV project covers testing of the decision making and path planning elements of the ADS.

A component of the Simulation Tool that remains critical for this testing is the vehicle dynamics model to ensure that ADS control commands are realistically enacted with correct vehicle dynamic behaviour. Modelling vehicle dynamics is an area of significant prior work [52], however the focus for VeriCAV is the correspondence of the vehicle dynamics model used by the ADS and the vehicle dynamics model used in the Simulation Tool. Trajectory control by an ADS generally requires a model of the vehicle dynamics to estimate the future vehicle trajectory based on actuator inputs, an example of this being Model Predictive Control (MPC) [53]. For real-world testing, the vehicle dynamics model would be calibrated to suit the real test vehicle, and a similar process is required to ensure the fidelity of the Simulation Tool vehicle dynamics model. Standardisation work in this area is ongoing [54].

8. Smart Actors

While the visual and sensorial realism of simulators has increased dramatically in recent years, with realistic looking scene occupant models, these models can only be considered sufficiently realistic if their behaviour is realistic as well. Automated actors such as pedestrians, cyclists or other vehicles have historically been controlled by relatively simple single-aspect models such as vehicle following or by pre-defined trajectories. But as the trajectory of an ADS being tested is not known a-priori this presents challenges when the ADS interacts with other actors. These actors should behave and react in a realistic way, but a basic model cannot account for the range of possible behaviours of the ADS, which would be required to maintain realism. A 'Smart Actor' can be any actor in a scenario that demonstrates behaviour which is more complex and nuanced than simple trajectory following or random movement [7]. The goal is to have smart actors which demonstrate human-like behaviour but can also be constrained to a degree such that the test scenario maintains its integrity as expected by the Test Generator and Oracle.

Having human-like 'smart actors' is vital for improving the fidelity of the simulation but is complicated to achieve. If one considers all the variation in human appearance, body movements, subtle head and eye-direction cues (both singularly and how these change in relation to actor groupings and conditions) it is clear that much work is needed to capture such realism in the simulation. As VeriCAV is focused on decision making and path planning, the Smart Actor development is focused primarily on the behavioural realism rather than visual. Such development will include data-driven models such as those that use machine-learning with real-world driving data to train models [55] as well as cognitive [56] and game theoretical models [57, 58]. Consortium partner, University of Leeds, are building on their experience in this field and the VeriCAV project will explore assessment methods to validate the Smart Actor models, building on existing literature in this area [59].

To balance testing efficiency, realism, and validity of a given scenario, not all actors need the same level of complexity so non-scenario critical actors or background traffic can be implemented independently of the Smart Actors. An important design decision is whether Smart Actor models are to be deterministic or probabilistic. A benefit of a deterministic model is that the same scenario can be repeated, and the Smart Actor behaviour will be the same, which is useful for re-testing. However, a probabilistic model allows the variance in real human behaviour to be represented. Smart Actors that can exploit the benefits of both approaches are likely needed to fulfil the requirements of scenario-based testing.

9. Test Oracle

As the complexity and variability in test scenarios increases with automatic test generation and smart actors, the assessment of test success becomes more challenging, as noted in Challenge 6 and particularly the automation of this process. The VeriCAV consortium partner, HORIBA MIRA, are developing a test oracle to tackle this challenge.

The automated analysis of test results is achieved by implementing a range of objective metrics by which test scenario performance is judged. These objective metrics are fed by data provided by the Simulation Master at the conclusion of each test scenario. Initially this work will focus on implementing existing metrics in literature such as the Responsibility-Sensitivity Safety (RSS) model [60] in an automated manner and subsequently developing further metrics focused on different aspects of ADS performance [61, 62]. This will also expose some of the practical challenges of extracting the required information from a simulation and its automated processing in the Test Oracle. The Test Oracle output for a single scenario contributes to a body of evidence for an ADSUT's performance but can also be used along with prior scenario results to inform subsequent test generation.

Automating the analysis of single test scenarios is critical, however it is also critical to assess how each test scenario result contributes to the overall assessment of the ADSUT performance, a concept referred to as coverage. Running many test scenarios that only vary a small amount doesn't provide useful evidence of ADSUT performance away from the narrow subset of scenarios tested. But, as noted previously, the parameter space is simply too large to test every possible combination, therefore a reliable assessment of coverage is critical. Using existing literature in the area of coverage based testing [63] and capitalising on novel methods from other fields, the VeriCAV project aims contribute to this important field with a particular focus on directing coverage based on the observed performance of the ADSUT.

10. Conclusions

The VeriCAV project concept has been presented, along with the challenges the project aims to meet to enable efficient, rigorous automated testing of ADSs. The significant areas of research are identified and the VeriCAV approach to modular interfacing, test generation and assessment and smart actors has been outlined.

The pragmatic approach taken in the project acknowledges that significant developments in Simulation Tools and ADSs will occur within and beyond the lifetime of the project. However, in order to progress virtual verification as a pillar of a structured ADS testing programme, a framework of robust, modular interfaces needs to be in place to allow research and commercial challenges to be addressed. The framework presented enables this within a representative end-to-end system while also investigating solutions to the particular practical challenges of virtual verification.

Acknowledgements

This work is undertaken as part of the VeriCAV project which is part-funded Innovate UK through the Centre for Connected and Autonomous Vehicles (CCAV). The authors wish to acknowledge the valued technical input and collaborative approach of VeriCAV project partners HORIBA MIRA, the University of Leeds and Aimsun.

References

[1] M. Wood, P. Robbel, M. Maass, R. D. Tebbens, M. Mejis, M. Harb, J. Reach, K. Robinson, D. Wittmann, T. Sri-vastava, M. E. Bouzouraa, S. Liu, Y. Wang, C. Knobel, D. Boymanns, M. Lhning, B. Dehlink, D. Kaule, R. Krger, J. Frtunikj, F. Raisch, M. Gruber, J. Steck, J. Meja-Hernandez, S. Syguda, P. Blher, K. Klonecki, P. Schnarz, T. Wiltschko, S. Pukallus, K. Sedlaczek, N. Garbacik, D. Szmera, D. Li, A. Timmons, M. Bellotti, M. O’Brien, M. Sch¨ollhorn, U. Dannebaum, J. Weast, A. Tatourian, B. Dornieden, P. Schnetter, P. Themann, T. Weidner and P. Schlicht, “Safety first for Automated Driving,” 2019.

[2] L. Fraade-Blanar, M. Blumenthal, J. Anderson and N. Kalra, Measuring Automated Vehicle Safety: Forging a Framework, RAND Corporation, 2018.

[3] Waymo LLC, “On the road to fully self-driving,” 2017.

[4] United Nations Economic Commission for Europe World Forum for Harmonization of Vehicle Regulations, “New Assessment/Test Method for Automated Driving (NATM),” 2020.

[5] Uber Advanced Technologies Group, “A Principled Approach to Safety,” 2018.

[6] E. Thorn, S. Kimmel and M. Chaka, “A Framework for Automated Driving System Testable Cases and Scenarios,” 2018.

[7] Z. Saigol and A. Peters, “Verifying automated driving systems in simulation: framework and challenges,” in *25th ITS World Congress*, 2018.

[8] R. Hillman, “Test Methods for Interrogating Autonomous Vehicle Behaviour–Findings from the HumanDrive Project,” 2019.

[9] T. Menzel, G. Bagschik, L. Isensee, A. Schomburg and M. Maurer, “From Functional to Logical Scenarios: Detailing a Keyword-Based Scenario Description for Execution in a Simulation Environment,” 10 5 2019.

[10] N. Kalra and S. M. Paddock, “Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?,” *Transportation Research Part A: Policy and Practice,* vol. 94, p. 182–193, 12 2016.

[11] J. Norden, M. O'Kelly and A. Sinha, “Efficient Black-box Assessment of Autonomous Vehicle Safety,” 8 12 2019.

[12] 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.

[13] G. Bagschik, T. Menzel and M. Maurer, “Ontology based Scene Creation for the Development of Automated Vehicles,” 29 3 2017.

[14] M. Schiementz, K. Groh, S. Wagner and T. Kühbeck, “PEGASUS-Test Case Variation and

Execution," 2019. [Online]. Available: https://mediatum.ub.tum.de/doc/1506730/file.pdf.

[15] Foretellix Ltd., *Measurable Scenario Description Language Reference,* 2019.

[16] M. Althoff, M. Koschi and S. Manzinger, "CommonRoad: Composable benchmarks for motion planning on roads," in *2017 IEEE Intelligent Vehicles Symposium (IV)*, 2017.

[17] P. Bender, J. Ziegler and C. Stiller, "Lanelets: Efficient map representation for autonomous driving," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, 2014.

[18] M. Klischat, O. Dragoi, M. Eissa and M. Althoff, "Coupling SUMO with a Motion Planning Framework for Automated Vehicles," in *SUMO User Conference*, 2019.

[19] Siemens, "PreScan," [Online]. Available: https://tass.plm.automation.siemens.com/prescan. [Accessed 30 05 2020].

[20] VIRES, "Virtual Test Drive (VTD)," [Online]. Available: https://vires.com/vtd-vires-virtual-test-drive/. [Accessed 30 05 2020].

[21] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez and V. Koltun, "CARLA: An Open Urban Driving Simulator," 10 11 2017.

[22] D. J. Fremont, T. Dreossi, S. Ghosh, X. Yue, A. L. Sangiovanni-Vincentelli and S. A. Seshia, "Scenic: A Language for Scenario Specification and Scene Generation," 25 9 2018.

[23] Rockstar Games, "Grand Theft Auto V," 2015. [Online]. Available: https://www.rockstargames.com/games/info/V.

[24] O. Michel, "Cyberbotics Ltd. Webots™: Professional Mobile Robot Simulation," *International Journal of Advanced Robotic Systems,* vol. 1, p. 5, 3 2004.

[25] G. Rong, B. H. Shin, H. Tabatabaee, Q. Lu, S. Lemke, M. Možeiko, E. Boise, G. Uhm, M. Gerow, S. Mehta, E. Agafonov, T. H. Kim, E. Sterner, K. Ushiroda, M. Reyes, D. Zelenkovsky and S. Kim, "LGSVL Simulator: A High Fidelity Simulator for Autonomous Driving," 7 5 2020.

[26] R. Queiroz, T. Berger and K. Czarnecki, "GeoScenario: An Open DSL for Autonomous Driving Scenario Representation," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, 2019.

[27] W. Damm, S. Kemper, E. Möhlmann, T. Peikenkamp and A. Rakow, "Using Traffic Sequence Charts for the Development of HAVs," in *ERTS 2018*, Toulouse, 2018.

[28] H. Winner, K. Lemmer, T. Form and J. Mazzega, "PEGASUS—First Steps for the Safe Introduction of Automated Driving," in *Lecture Notes in Mobility*, Springer International Publishing, 2018, p. 185–195.

[29] H. Elrofai, J. P. Paardekooper, E. d. Gelder, S. Kalisvaart and O. Op den Camp, "StreetWise: scenario-based safety validation of connected automated driving," 2018.

[30] A. Leitner, D. Watzenig and J. Ibanez-Guzman, Eds., Validation and Verification of Automated Systems, Springer International Publishing, 2020.

[31] Connected Places Catapult, "Multi User Scenario Catalogue for Connected Autonomous Vehicles," 2020. [Online]. Available: https://cp.catapult.org.uk/case-studies/musicc/.

[32] SystemX Institute of Research and Technology, "Simulation of Autonomous Vehicle Safety," [Online]. Available: https://www.irt-systemx.fr/en/projets/sva/.

[33] OpenADx Working Group, "OpenADx," [Online]. Available: https://openadx.eclipse.org/. [Accessed 01 06 2020].

[34] Deepen AI, World Economic Forum, McKinsey & Company, "Safety Pool," 2020. [Online]. Available: https://www.safetypool.ai/. [Accessed 01 06 2020].

[35] Metamoto, 2016. [Online]. Available: https://www.metamoto.com/.

[36] Cognata, 2016. [Online]. Available: https://www.cognata.com/.

[37] Alpha Drive, 2017. [Online]. Available: https://alphadrive.ai/.

[38] Aimsun, "Aimsun Auto: Simulate a Driverless Future," 2020. [Online]. Available: https://www.aimsun.com/aimsun-auto/.

[39] R. Bellman, Dynamic Programming, Princeton University Press, 1957.

[40] S. Feng, Y. Feng, C. Yu, Y. Zhang and H. X. Liu, “Testing Scenario Library Generation for Connected and Automated Vehicles, Part I: Methodology,” 9 5 2019.

[41] C. E. Tuncali, G. Fainekos, H. Ito and J. Kapinski, “Simulation-based Adversarial Test Generation for Autonomous Vehicles with Machine Learning Components,” 18 4 2018.

[42] Y. Abeysirigoonawardena, F. Shkurti and G. Dudek, “Generating Adversarial Driving Scenarios in High-Fidelity Simulators,” in *2019 International Conference on Robotics and Automation (ICRA)*, 2019.

[43] Baidu, “Baidu Apollo,” [Online]. Available: https://apollo.auto/. [Accessed 01 06 2020].

[44] R. M. Fujimoto, “Time Management in The High Level Architecture,” *SIMULATION,* vol. 71, p. 388–400, 12 1998.

[45] S. Kato, E. Takeuchi, Y. Ishiguro, Y. Ninomiya, K. Takeda and T. Hamada, “An Open Approach to Autonomous Vehicles,” *IEEE Micro,* vol. 35, p. 60–68, 11 2015.

[46] C. W. Gran, “HD-Maps in Autonomous Driving,” 2019.

[47] A. Geiger, P. Lenz and R. Urtasun, “Are we ready for autonomous driving? The KITTI vision benchmark suite,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

[48] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan and O. Beijbom, “nuScenes: A multimodal dataset for autonomous driving,” 26 3 2019.

[49] P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, V. Vasudevan, W. Han, J. Ngiam, H. Zhao, A. Timofeev, S. Ettinger, M. Krivokon, A. Gao, A. Joshi, S. Zhao, S. Cheng, Y. Zhang, J. Shlens, Z. Chen and D. Anguelov, “Scalability in Perception for Autonomous Driving: Waymo Open Dataset,” 10 12 2019.

[50] Y. Kang, H. Yin and C. Berger, “Test Your Self-Driving Algorithm: An Overview of Publicly Available Driving Datasets and Virtual Testing Environments,” *IEEE Transactions on Intelligent Vehicles,* vol. 4, p. 171–185, 6 2019.

[51] P. Rosenberger, M. Holder, S. Huch, H. Winner, T. Fleck, M. R. Zofka, J. M. Zollner, T. D\textquotesinglehondt and B. Wassermann, “Benchmarking and Functional Decomposition of Automotive Lidar Sensor Models,” in *2019 IEEE Intelligent Vehicles Symposium (IV)*, 2019.

[52] T. D. Gillespie, Fundamentals of vehicle dynamics, Warrendale, PA: Society of Automotive Engineers, 1992.

[53] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, M. Sokolsky, G. Stanek, D. Stavens, A. Teichman, M. Werling and S. Thrun, “Towards fully autonomous driving: Systems and algorithms,” in *2011 IEEE Intelligent Vehicles Symposium (IV)*, 2011.

[54] S. Schnelle, K. Salaani, S. Rao, F. Barickman and D. Elsasser, “Review of Simulation Frameworks and Standards Related to Driving Scenarios,” 2019.

[55] F. Behbahani, K. Shiarlis, X. Chen, V. Kurin, S. Kasewa, C. Stirbu, J. Gomes, S. Paul, F. A. Oliehoek, J. Messias and S. Whiteson, “Learning From Demonstration in the Wild,” in *2019 International Conference on Robotics and Automation (ICRA)*, 2019.

[56] C. C. Macadam, “Understanding and Modeling the Human Driver,” *Vehicle System Dynamics,* vol. 40, p. 101–134, 1 2003.

[57] F. Camara, R. Romano, G. Markkula, R. Madigan, N. Merat and C. Fox, “Empirical game theory of pedestrian interaction for autonomous vehicles,” in *Measuring Behavior 2018: 11th International Conference on Methods and Techniques in Behavioral Research*, 2018.

[58] K. Kang and H. A. Rakha, “A Repeated Game Freeway Lane Changing Model,” *Sensors,* vol. 20, p. 1554, 3 2020.

[59] J. Bernhard, K. Esterle, P. Hart and T. Kessler, “BARK: Open Behavior Benchmarking in Multi-Agent Environments,” 5 3 2020.

[60] S. Shalev-Shwartz, S. Shammah and A. Shashua, “On a Formal Model of Safe and Scalable Self-driving Cars,” 21 8 2017.

[61] R. Myers and Z. Saigol, “Pass-Fail Criteria for Scenario-Based Testing of Automated Driving Systems,” 19 5 2020.

[62] B. Klamann, M. Lippert, C. Amersbach and H. Winner, “Defining Pass-/Fail-Criteria for Particular Tests of Automated Driving Functions,” in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019.

[63] Z. Tahir and R. Alexander, "Coverage based testing for V&V and Safety Assurance of Self-driving Autonomous Vehicle : A Systematic Literature Review," in *The Second IEEE International Conference On Artificial Intelligence Testing*, 2020.

[64] W. Damm, E. Möhlmann and A. Rakow, “A Scenario Discovery Process Based on Traffic Sequence Charts,” in *Validation and Verification of Automated Systems*, Springer International Publishing, 2019, p. 61–73.

[65] T. Ponn, D. Fratzke, C. Gnandt and M. Lienkamp, “Towards Certification of Autonomous Driving: Systematic Test Case Generation for a Comprehensive but Economically-Feasible Assessment of Lane Keeping Assist Algorithms,” in *Proceedings of the 5th International Conference on Vehicle Technology and Intelligent Transport Systems*, 2019.

[66] F. Gao, J. Duan, Y. He and Z. Wang, “A Test Scenario Automatic Generation Strategy for Intelligent Driving Systems,” *Mathematical Problems in Engineering,* vol. 2019, p. 1–10, 1 2019.

[67] G. Chance, A. Ghobrial, S. Lemaignan, T. Pipe and K. Eder, “An Agency-Directed Approach to Test Generation for Simulation-based Autonomous Vehicle Verification,” 11 12 2019.

[68] P. Hintjens, ZeroMQ: messaging for many applications, " O'Reilly Media, Inc.", 2013.

[69] S. Kato, S. Tokunaga, Y. Maruyama, S. Maeda, M. Hirabayashi, Y. Kitsukawa, A. Monrroy, T. Ando, Y. Fujii and T. Azumi, “Autoware on Board: Enabling Autonomous Vehicles with Embedded Systems,” in *2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS)*, 2018.

[70] M. Montemerlo, J. Becker, S. Bhat, H. Dahlkamp, D. Dolgov, S. Ettinger, D. Haehnel, T. Hilden, G. Hoffmann, B. Huhnke, D. Johnston, S. Klumpp, D. Langer, A. Levandowski, J. Levinson, J. Marcil, D. Orenstein, J. Paefgen, I. Penny, A. Petrovskaya, M. Pflueger, G. Stanek, D. Stavens, A. Vogt and S. Thrun, “Junior: The Stanford Entry in the Urban Challenge,” *J. Field Robot.,* vol. 25, p. 569–597, 9 2008.

[71] C. R. Baker and J. M. Dolan, “Traffic interaction in the urban challenge: Putting boss on its best behavior,” in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008.

[72] T. Menzel, G. Bagschik and M. Maurer, “Scenarios for Development, Test and Validation of Automated Vehicles,” 5 1 2018.

[73] D. J. Fremont, E. Kim, Y. V. Pant, S. A. Seshia, A. Acharya, X. Bruso, P. Wells, S. Lemke, Q. Lu and S. Mehta, “Formal Scenario-Based Testing of Autonomous Vehicles: From Simulation to the Real World,” 17 3 2020.

[74] C. Amersbach and H. Winner, “Defining Required and Feasible Test Coverage for Scenario-Based Validation of Highly Automated Vehicles,” in *22nd IEEE Intelligent Transportation Systems Conference (ITSC) 2019*, 2019.

[75] C. Amersbach and H. Winner, “Functional decomposition—A contribution to overcome the parameter space explosion during validation of highly automated driving,” *Traffic Injury Prevention,* vol. 20, p. S52–S57, 6 2019.

[76] P. Koopman and M. Wagner, “Challenges in Autonomous Vehicle Testing and Validation,” *SAE International Journal of Transportation Safety,* vol. 4, p. 15–24, 4 2016.