Exploiting Learning and Scenario-based Specification Languages for the Verification and Validation of Highly Automated Driving

Werner Damm*
OFFIS
Oldenburg, Germany
werner.damm@offis.de

Roland Galbas
Robert Bosch GmbH, Chassis Systems Control
Abstatt, Germany
Roland.Galbas@de.bosch.com

ABSTRACT

We propose a series of methods based on learning key structural properties from traffic data-basis and on statistical model checking, ultimately leading to the construction of a scenario catalogue capturing requirements for controlling criticality for highly autonomous vehicles. We sketch underlying mathematical foundations which allow to derive formal confidence levels that vehicles tested by such a scenario catalogue will maintain the required control of criticality in real traffic matching the probability distributions of key parameters of data recorded in the reference data base employed for this process.

CCS CONCEPTS

• **Computing methodologies** → **Reasoning about belief and knowledge**; **Intelligent agents**; **Machine learning**; **Supervised learning by classification**; **Cost-sensitive learning**; **Modeling and simulation**; **Model verification and validation**; **Uncertainty quantification**; *Knowledge representation and reasoning*; *Nonmonotonic, default reasoning and belief revision*; *Spatial and physical reasoning*; *Planning under uncertainty*; *Partially-observable Markov decision processes*; *Rare-event simulation*; *Simulation by animation*; • **Social and professional topics** → *Codes of ethics*;

KEYWORDS

highly automated driving; requirement analysis; formal specification; learning; statistical model-checking; verification and validation

ACM Reference Format:
Werner Damm and Roland Galbas. 2018. Exploiting Learning and Scenario-based Specification Languages for the Verification and Validation of Highly Automated Driving. In *SEFAIAS'18: IEEE/ACM 1st International Workshop on Software Engineering for AI in Autonomous Systems , May 28, 2018, Gothenburg, Sweden.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3194085.3194086

*Corresponding author Werner Damm. E-mail address: werner.damm@offis.de

SEFAIAS'18, May 28, 2018, Gothenburg, Sweden

ACM ISBN 978-1-4503-5739-5/18/05…$15.00
https://doi.org/10.1145/3194085.3194086

1 INTRODUCTION

It is well-known (c.f. [20]) that traditional approaches for type homologation fail for highly autonomous vehicles due to the impossibility of covering sufficiently many kilometers in field testing to achieve a statistically valid basis for building safety cases, due to the extreme variability of environmental contexts and the resulting complexity in both the perception- and trajectory planning systems of highly autonomous vehicles. The approach followed by the German automotive industry and presented in this paper builds on scenario catalogues to capture for all perceivable traffic situations' requirements on such systems jointly ensuring global safety objectives. Test drives are to be replaced to a significant extent by placing the vehicle under test in test environments exposing the vehicle to traffic situations covering all scenarios of the scenario catalogue, and monitoring compliance of the vehicle's reaction to such scenarios in a virtual environment. Such test environments will allow testing separately the perception components (along all stages covering preprocessed sensor data, sensor fusion, object identification algorithms) and the trajectory planning component (which involves exploring possible future evolutions of the currently perceived traffic situation to decide on the planned maneuver). Projects already running and pushing this approach are the Pegasus project funded by the German Federal Ministry for Economic Affairs and Energy, involving all major German OEMs and Tier 1 companies, and the ENABLE-S3 project funded by the Joint Undertaking ECSEL, including both German and French automotive companies.[1]

Currently, new projects are under formation to build industry strength processes and test environment elaborating this approach.

There are several challenges which must be addressed to make this approach viable:

(1) Can we capture at design time the space of all possible traffic situations and environmental factors relevant for safe trajectories for autonomous vehicles?
   (a) Can we characterize the environmental conditions for all elements in the perception chain under which identification of objects can be guaranteed for a given desired confidence level?
   (b) Can we characterize the variability of dynamics of other participants to allow safe predictions of future evolution of traffic situations for a given confidence level?
   (c) Can we find generalized risk limits for certain traffic (environmental) situations which a vehicle has to cope with?

[1](european initiative to enable validation for highly automated safe and secure systems). Grant nr. 692455-2 Call H2020-ECSEL- 2015-2-IA-two-stage call (2016)

(d) Given the complexity space of real-world traffic situations and relevant environmental factors, how can one at all achieve sufficiently concise specifications to make construction of scenario catalogues capturing these viable?
(e) Given the ill-structuredness of the space of real-world traffic situations, how can we achieve completeness of scenario catalogues, i.e. demonstrate with high confidence that all relevant real-world situations have been captured?

(2) Given the remaining likelihood of experiencing failures in perception and interpretation after deployment, how can we establish a process learning from field incidents and accidents leading to updates of the scenario catalogue avoiding reoccurrence of this incident in the field?
(3) How can we assure, that the interpretation of scenarios and thus interpretation of test results is unambiguous across all test platforms?

The following three sections elaborate on each of the challenges and describe possible lines of attack. Section 2 analyses the space of all real-world traffic situations and environmental factors, identifying reasons allowing to collapse infinitely many situations into a parametrized set of finite equivalence classes, thus addressing Challenge 1. Section 3 elaborates the approach proposed by the SafeTRANS Working Group on Highly Automated Systems [5] for learning from in-field incidents to address Challenge 2. Section 4 discusses, how an ongoing standardization approach for capturing scenarios can be extended into a formal scenario specification language with a well-defined reference semantics allowing automatic generation of monitors for classifying test results, thus addressing challenge 3. Section 5 discusses related work. We conclude this paper with an outlook on the future steps in taking the concepts presented in this paper into reality.

2 ANALYSING THE SPACE OF TRAFFIC SITUATIONS AND ENVIRONMENTAL FACTORS

This section is structured as follows. Subsection 2.1 identifies key observations allowing to ultimately reduce the space of all traffic situations and environmental factors to a set of finitely specifiable parametrized equivalence classes. Subsection 2.2 proposes learning based approaches to identify these equivalence classes with quantified confidence levels. Subsection 2.3 discusses how to learn requirements on reducing criticality.

2.1 Identifying structure

If there would be no underlying structural principles in the extreme complex space of driving situations and/or environmental conditions, humans would not be able to drive. Yet we are able, with limitations and varying skill levels, to learn to navigate in traffic, after a reasonable amount of training. Human drivers are able to identify structures, which allow them to assess complex situations and dynamics in seconds, even though miss-assessments occur. Our approach is based on the key assumption, that today's machine learning algorithms are sufficiently powerful to identify the structural principles allowing human drivers to master these extremely complex tasks, because the wealth of data available from in-field data, and the stimulation of such learning algorithms with billions of traffic situations and environmental conditions in virtual environments, coupled with advanced techniques for refining identified concepts from too coarse abstractions, provide enough ground truth information to identify all aspects of traffic situations and environmental conditions relevant for determining safe trajectories for autonomous driving.

Further evidence for the capability of automatically identifying structure comes from recent results in formal verification of complex systems. Research in counter-example guided verification of hybrid systems [1, 23, 30] has shown how abstraction predicates can be automatically learned to achieve a sufficiently precise finite characterization of those aspects of non-linear hybrid systems sufficient to establish their safety. Intuitively, this seemingly astonishing result, that a finite number of abstraction predicates is sufficient to characterize safety in spite the complexity of the uncountable state space underlying such systems, comes from arguments based on continuity and robustness of systems. Indeed, both safety and liveness properties of non-linear hybrid systems have been proven to be decidable under industrially used robustness assumptions [9, 12, 28]. In the intersection of research on learning and research on formal verification, it has been suggested in restricted contexts, that machine learning techniques can be used to learn the initial abstraction predicates in such verification approaches [2, 31].

Our overall approach can be seen as a generalization of this line of research to the much more complex class of mathematical systems required to formally capture the dynamics of evolving traffic situations in the space of environmental conditions, which essentially requires the expressive power of dealing with unbounded parallel compositions of (probabilistic and nondeterministic) hybrid automata with highly non-linear dynamics [18], which are subject to disturbances and failure events with unknown probability distributions. The predicate abstractions we expect to learn have to be able to characterize criticality of driving situations under different environmental conditions. Much as we as humans will orient our decisions e.g. to change a lane on an assessment of the gap available on the lane into which we want to change (c.f. [39–41], we expect such predicate abstractions to automatically cluster relative dynamics of the ego vehicle and vehicles on the adjacent lane into a finite set of classes. E.g., for a lane change maneuver, these will include predicates "lane sufficiently clear" – the distance and relative speed of vehicles behind us is such that it will not impact our planned maneuver -, "gap large enough to perform lane change", "gap allows lane change with good driving skills", "no lane change possible". Clearly these predicates must be complemented by predicates on prevailing weather and road conditions to determine whether a lane change maneuver can be performed safely, such as taking into account aquaplaning, icing, fog, etc.

In general, we need to learn *all* aspects relevant for judging the *criticality* of driving situations, as elaborated below. Let us refine our understanding of what this entails.

As a first approximation, criticality of driving situations is a function of

- the *complexity* of traffic situations, as e.g. expressed in the number and detail of objects and environmental conditions which must be correctly classified,

- the relative *dynamics* of involved vehicles and relevant objects

Figure 1 (a) gives a two-dimensional representation of a criticality measure for driving situations as a function of criticality and relative dynamics, by showing qualitatively lines of equal criticality. Note that the space of evolutions to be analyzed for taking safe driving decisions grows exponentially in both dimensions. Figure 1 (b) shows corner cases based on first functions with probable market introduction. If we assume that criticality is described as the probability for an accident including its severity then we can show that criticality is specific for vehicles because it depends on

- the observability (perception),
- the predictability (cognition) of the environmental conditions including its future, and
- the controllability (motion) of the vehicle.

We thus have to refine our understanding about relevant aspects influencing criticality by these three additional categories. Note that these built on key intellectual properties of OEMs and Tier 1 suppliers, hence an additional challenge in setting up a generic V&V environment is to propose characterizations of observability, predictability, and controllability, which do not reveal IP but are still sufficiently precise to be acceptable for type homologation. Our approach to solving this challenge rests again on learning from field data: we assume a sufficiently rich set of real-world observations labeled with ground truth, such that learning methods can be used to 1) learn models about dynamics of all relevant classes of traffic participants (e.g. through parameter learning in probabilistic hybrid automata of vehicle dynamics), 2) learn models of controllability of the dynamics of the ego vehicle, and to 3) learn models about perception errors along the complete perception chain (from multiple sensors, sensor fusion to object identification algorithms). Regarding item 1), this requires us to reflect on the complete relevant environmental behavior in traffic. This entails that we cannot and do not expect other vehicles to be compliant to traffic rules. Instead, our models will reflect the typical levels of deviations seen in different countries, in different road conditions, including rare rowdy like behavior, with probabilities justified empirically from the data base of real-traffic scenarios. Regarding item 2) it is necessary to understand the limits of controllability (safe ego-behavior) within the relevant dynamic of traffic environment. Here we rely on data observed in the field regarding manufacturer independent measures of controllability of critical driving situations. Regarding item 3) it is obvious that misclassifications leading to ghost objects, or failure to identify relevant objects in traffic situations, drastically raise criticality, as do miss-predictions of the evolution of the current traffic situation, e.g. because of employing poor or (due to misclassification) even wrong models of dynamics of other traffic participants. For establishing an overall safety case, for items 1)-3), it is thus key to characterize the confidence level of object identification along all elements of the perception chain, assuming sensor systems are used within their specified range, as well as to characterize for each sensor type all environmental conditions under which no high confidence data can be generated from this particular sensor system. The remainder of this Section addresses the following research questions

- How can we ensure that the vehicle will control all typical and known traffic situations with a tolerable safety-behavior? (treated in Subsection 2.2)
- How can we ensure that the vehicle control typical accident situations better or at least equal to non-automated vehicles? (treated in Subsection 2.3)

2.2 Employing learning techniques for structure identification

We assume —as in the Pegasus Project[2]— a given suite of data-bases *DB* of real-traffic situations, where each element in the database represents a timed sequence of heterogeneous data showing the evolution of a given traffic situation over time. We expect to use data-bases which contain subsets of combination of the following type of data (c.f. [27])

(1) Data bases coming from in-field recording with test vehicles, including time synchronized data from all types of sensor data (in particular including video images), vehicle dynamics data, data relating to the perception of the environment, data relating to control of vehicle dynamics by car and/or driver. These data bases need to be labelled with an accuracy close to ground truth with an at least known error.
(2) Data bases coming from infra-structures in test fields, including video data and data on status of infrastructure control
(3) Accident data bases

The OpenScenario Initiative [33–35] is currently harmonizing a taxonomy of the categories of objects considered to be relevant for testing highly autonomous vehicles, such as types of traffic participants, types of environmental conditions, types of relations between these. This ongoing pre-standardization effort is thus already creating a key abstraction process, in that it stipulates, that only those objects which are appearing in the ontology are relevant for judging the criticality of driving situations. The strategies for identifying factors influencing criticality, while building on the current ontology, are expected to provide extensions with further artefacts.

We factor the description of our learning process in separate dimensions, and use in each dimension an iterative learning approach. In a real implementation, we expect these learning processes to be interwoven.

Let's assume first *perfect observation*. Under this hypothetical assumption, all errors related to miss-perception are ruled out and tests derived will assess a vehicles capability to maintain a tolerated level of risks during all performed manoeuvres. In a first step, we analyze all scenarios of *DB* with a metric to measure risks. We initially focus our learning process on identifying all causes for accidents or near-accident situations. More precisely, we aim, for any given desired level of confidence *cl*, to identify all those real-world artefacts observable in the data-base such that the likelihood of having missed an artefact in the data-base contributing to an accident observed in the data base is less than *cl*. If we are able to identify such a set *H* (of hazards), we say that *H is DB complete for explaining accidents with confidence level cl*. Note that technically speaking elements of *H* are abstraction predicates: they cluster

[2] www.pegasusprojekt.de

Figure 1: Two-dimensional representation of criticality measure

those real-world phenomena such as "road is icy", "lateral distance too close", etc., which have been observed in *DB* in scenarios leading to accidents. Such predicates p are included in H if there is at least one accident scenario in *DB* leading to an accident A such that p is *DB*-necessary for A, i.e. all *DB* scenarios not containing a subset of p do not lead to an accident of type A. We propose an iterative learning process using counter-example guided abstraction refinement to derive such a set H. We initially start with already identified hazards H_0 in OpenScenario, and find counterexamples where these are not sufficient to explain some accident type A. We then learn from counterexamples new or refined predicates s.t. the extension of H_0 with these is *DB* necessary for A. For example, if H_0 would contain only the predicate "wet road surface", then for many dynamic situations this would not be sufficient for explaining accidents where a car crashes in a curve into a tree; only by refining this to identify that the level of water on the surface is causing aquaplaning will this predicate in the given dynamic situation be causally relevant to this accident. We iterate this process until we have found for all accident types A observed in *DB* all necessary hazards with confidence level *cl*. A rigorous formal justification of this process is given in a companion paper.

An orthogonal dimension refers to learning models of dynamics of traffic participants. We are assuming perfect observation, and can thus classify all traffic participants e.g. using the OpenScenario taxonomy. We also assume that scenarios in *DB* are labelled as being instances of a finite set of parametrized classes of *elementary traffic situations*, such that all scenarios in *DB* can be built by combining such elementary traffic situations (e.g. entering roundabout with X lanes, left turn on intersection of type Y, entering highway of type Z, ...). For each class of traffic participant *tpc* (such as pedestrians, bicycles, trucks, buses, police cars, normal cars, etc., as defined by OpenScenario) and each class of traffic situations *ts* we learn probabilistic hybrid automata $HA(tpc, ts)$ explaining the behaviors observed in *DB* by parameter fitting techniques (c.f. [13]). These learned models are key for creating virtual scenarios matching the characteristics of *DB* for simulation environments. Note that these models are not expected to conform to traffic rules nor to exclude rowdy behavior – they simply reflect distributions of behavior of traffic participants in real life.

Next, we learn scenarios for a test catalogue *SC*, testing whether vehicles are able to control criticality in the presence of hazards. Note that we are still assuming perfect perception; under this assumption, all predicates in H are observed by the autonomous vehicle, e.g. all types of traffic participants, and all types of traffic situations are "known" to the vehicle, and we want to test its capability of adequately assessing the risk in a given traffic situation with known traffic participants in the presence of a given set of hazards. To this end we generate scenarios on required behavior of vehicles under test requiring the vehicle to reduce risks in all combinations of presence of hazards, traffic situations, and traffic participants to an *acceptable tolerated level*, as discussed in Subsection 2.3, *assuming full control*. Again, we use an iterative learning approach using counter-example guided abstraction refinement until the derived set of scenarios is shown to be *DB-complete with confidence level cl*, i.e. the remaining probability to not control risk to the given level is less than *cl* for all combinations of hazards, traffic participants, and elementary traffic situations. In this iterative process, we rely on generating *virtual scenarios* using the dynamic models learned as described above, and what we call guided simulations [14, 26] driving environment models in simulations so as to increase risk levels, as a basis for a statistical model-checking argument [11, 16] regarding the confidence level with which a vehicle compliant to the test catalogue is able to reduce criticality with the given level of confidence.

The next learning processes need to relax the assumptions of *perfect observation* and *full control*.

To relax the assumption of perfect information, we propose the following method, initially excluding component (hardware) failures.

(1) We learn for each sensor type S a set of hazardous environmental conditions *Hs* s.t. *Hs* is *DB complete for explaining perception failures* due to S with a given confidence level *cl*, following the iterative learning paradigm used for hazards to criticality under the assumption of perfect observation, and learn probability distributions for their occurrence in *DB*. For this analysis, we need to be able to assess data from test vehicles, time synchronized with data from test-beds for automated driving, to identify discrepancies between the test-vehicle sensor's perception of reality, and ground truth provided from the test-beds infrastructure, and learn causal dependencies between such discrepancies and environmental conditions observed through test-bed infrastructures.

(2) We learn statistical models of the confidence level of the complete perception chain assuming absence of *Hs* for all sensors *S*, yielding a residual probability of perception failures even in the absence of component failures and environmental conditions hazardous to the employed sensor systems, again relying on the availability to detect discrepancies between the test vehicle's perception of reality and ground truth, but now considering only such traces, where no hazards from *Hs* occur.
(3) We then test the robustness of the vehicles risk control strategies by injecting hazardous environmental situations from *Hs* for all *S* in simulated environments using the learned probability distribution, injecting miss-classification failures as discussed under item 2 above, and using the probability distributions for component failures (assumed to be known).

To relax the assumption of perfect control, we need to assess two aspects:

(1) Controllability is obviously impacted by component failures. This is a standard topic in safety analysis, and we assume that probability distributions of component failures are available.
(2) In SAE levels involving hand-over to drivers or higher levels of driver control, we need to learn statistical models about the controllability of given risk levels in given elementary traffic situations.

We thus follow the principle of separation of concerns for reducing the complexity of test-catalogues, in testing first the vehicles capability of risk mitigation under perfect perception and perfect control, and then testing for robustness in perception and control failures. A separate paper extending [11, 16] will provide a more rigorous mathematical analysis of the justification of this decomposition approach.

Note that risks coming from security threats demand additional analysis.

2.3 Building up a reference for automated vehicles

We now turn to refining the scenarios generated with techniques from Section 2.2, in making precise the required level of risk reduction. This section thus addresses the societally tolerated level of risk for autonomous driving, as defined by the finding of the ethics commission of the German Minister of Transportation [25]. From its rule 2, we should strive to achieve at least the same level of safety as human drivers. The risk measures employed in Section 2.2 must comply to its rule 7. The overall defensive style proposed in rule 5 must be reflected. Finally, rule 9 must be reflected again in risk measures. For convenience of the reader, the cited rules are shown in Annex A. In this paper, we focus on rule 2. As a reference about accidents statistics without highly autonomous driving we consider all entries from the GIDAS data base maintained by BAST with a minimum MAIS level greater than 1 (corresponding to light injuries). Note that these data were already taken into account for the steps in Subsection 2.2 in assessing the controllability of critical situations – in particular we can derive from these to what extend human drivers were able to control in particular classes of elementary traffic situations (such as driving through an intersections) risks factors (such as "child running out of a vehicle parked at intersection", "icy surface on intersection"), and scenarios generated in Subsection 2.2 would enforce automated vehicles to achieve the same level of controllability as observed in GIDAS, thus the same level of controllability as achieved by human drivers. We note that our ability to cluster such requirements into a finite set of requirement scenarios is inherently exploiting continuity arguments, in the following sense:

(1) As mentioned before, we expect to learn probabilistic hybrid automata as models for all aspects of the environment. Within one mode, dynamics follow probabilistic differential inequalities, hence represent continuous functions.
(2) All mode changes influencing criticality are expected to be identified in the risk analysis discussed in Subsection 2.2.
(3) The analysis of relaxing the observation on perfect information leads to requirements on confidence levels, with which such hazardous events are detected by the vehicle. We note that Car2X communication is expected to significantly increase our confidence levels in detecting such conditions.
(4) Since only finitely many mode-changes occur in a finite time window, and the continuous dynamics within one mode can be clustered into finitely many abstraction predicates (c.f. [1, 9], we are able to characterize the risk-control requirements in a finite set of scenarios.

It must be pointed out, though, that this analysis hinges greatly on our ability to learn such environmental models. Certainly, significant further research is required to achieve e.g. sufficiently precise pedestrian models (c.f. [10, 21, 22, 29, 36], to name but one critical research area. With respect to type homologation, the environmental models used for testing should be subject to standardization, and subject to a continuous improvement process such as discussed in Section 3.

In order to derive requirements improving controllability of hazardous situations beyond human capabilities, we consider the probability mass of accident classes in GIDAS of MAIS levels greater or equal 3, and refine scenarios generated in the process of Subsection 2.2 by tightening the requirements on the rate of reduction of criticality beyond human capabilities by selecting classes of accident scenarios of GIDAS with comparatively high probability masses. Note that this probability is determined by the (uncontrollable) probability of occurrence of such hazards, the relevance of the occurrence of this hazard in the given elementary traffic situation, and the remaining conditional probability of the ego car to control risk reduction in the presence of these hazards. The degree of tightening requirements on the controllability of the ego car is subject to a discourse with the relevant public authorities, and can be based on the already observed significant capabilities of accident reduction, such as e.g. for rear-end collisions or emergency breaking in urban environments demonstrated by the state of practice of recently introduced ADAS systems.

Altogether, the contributions of Section 2 thus lead to the following result:

(1) We have proposed an approach based on learning to derive scenario based requirement specifications for controlling criticality of autonomously driving vehicles, addressing Challenge 1 of the introduction.

(2) We have indicated how societal standards such as recently proposed by the Ethics Commission of the German Ministry of Transportation can be taking into account when building a Scenario Catalogue used as basis for type homologation.
(3) We have explicated the dependencies of the generated scenario catalogue on the quality and representativeness of the data-bases used for learning processes. Clearly, the derived Scenario Catalogue will not be sufficient when used for type homologation in areas, where types and probability distributions of hazards or models on environment differ significantly from distributions observed in the reference data-base.
(4) On the other hand, as long as the actual driving situations experienced in field match these probability distribution, the proposed process allows to demonstrate for any given confidence level that automated vehicles will be compliant to the scenario catalogue.

The following two Sections now address Challenges 2 and 3 of the Introduction.

3 LEARNING IN THE FIELD

There is no perfect system – especially not in the beginning. Looking at the extremely complex environment of traffic situations, it is to be expected that after the homologation process of a highly automated vehicle new unpredictable, but risky situations may occur because of a certain environmental change. Uncertainties and risks in field have to improve the performance of *all* vehicles. This implies: not one vehicle learns for itself, but it contributes its data about such "miss situations" towards an aggregated data-base maintained by an independent public body. The data of many vehicles will be analyzed (e.g. in case of near misses), based upon this - new knowledge will be created and redistributed to adaptable vehicles – thus creating a process which could be called *"Community learning"*. Related to "Community learning" is the establishment of a *"Permanent safety procedure"* compared to today's one-time homologation process. The reason for this lies in the fact, that it has to be ensured, that safety-relevant changes are proceeded by a safety–proven process. Also related to "Community learning" is the establishment of a *"Permanent observation system"* at the vehicle site which provides e.g. the safety relevant vehicle-data for "Community learning" and enables to cope with the safety requirement for an independent instance and last but not least increase the safety performance of the vehicle (e.g. by health monitoring). Taken also into account that the increase of knowledge is growing exponentially it can be derived that

- Future mobility systems including highly autonomous vehicles will be adaptable for upgrades in the field in order to incorporate new knowledge and changes as e.g. safety updates resulting from the community learning process.
- The installation of an in-car observation system supporting the community learning and of mechanisms ensuring safe application of updates (permanent safety updates) is highly probable.

Assuming that these key issues lead to the described system changes, then each single OEM has to carry the complete liability risk for content and process of each release/upgrade. Therefore, it is more effective, that multiple OEMs and Tiers establish a **common standardized safety process** - see proposal [5] by the Safetrans organization – Germany), including standardized interfaces for a common observation and validation platform and a procedure for permanent field-observation.

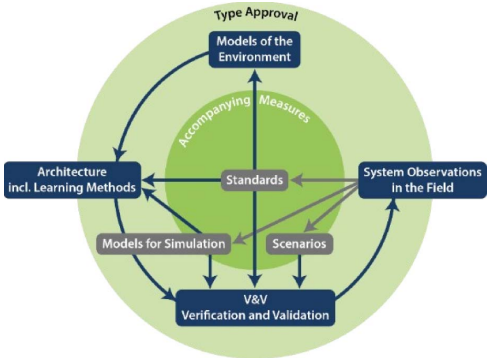

Figure 2: (cited from [5]) In field system observations contribute to community learning process.

4 FORMAL SPECIFICATIONS OF TRAFFIC SCENARIOS AND ENVIRONMENTAL CONDITIONS

A key challenge then rests in finding a representation of test-scenarios which are expressive enough to cope with the plethora of environmental conditions, traffic situations, and traffic participants, while yielding concise and **unambiguous** interpretations. In [6, 7, 16] we have been proposing the visual specification language of Traffic Sequence Charts to answer this challenge, which – much as its conceptual "father" Life Sequence Charts extends and gives formal semantics to Message Sequence Charts – extends the OpenScenario approach. Much as Message Sequence Charts ITU-T 2011 were lacking expressiveness and formal semantics, motivating the extension to Live Sequence Charts [4], the ongoing industrial pre-standardization effort for capturing scenarios, called OpenSCENARIO [33–35], falls significantly short in being able to support the methodology described in Subsection 2.2 and 2.3 to address the challenges in the introduction. OpenSCENARIO allows describing what we called existential charts, i.e. give examples of desired behaviors, rather than being able to specify requirements on all behaviors, such as in what we called universal LSCs. TSCs "inherit" from LSCs the concepts related to distinguishing between possible and mandatory behaviors, the concepts of pre-charts which is key for characterizing those situations from when on all behaviors must comply to universal charts, and cold and hot conditions for distinguishing case-distinctions from failures. Having a specification language for scenarios with a rigorously defined formal semantics is necessary for the following reasons. Recall the challenges of the introduction, slightly regrouped: (C1) Given the ill-structuredness of the space of real-world traffic situations, how can we achieve completeness of scenario catalogs, i.e. demonstrate with high confidence that all relevant real-world situations have been captured? (C2) Given the remaining likelihood of experiencing failures in perception and

interpretation after deployment, how can we establish a process learning from field incidents and accidents leading to updates of the scenario catalog avoiding reoccurrence of this incident in the field? (C3) Given the complexity space of real-world traffic situations, how can one at all achieve sufficiently concise specifications to make construction of scenario catalogs viable? (C4) How can we assure, that the interpretation of scenarios and thus interpretation of test results is unambiguous across all test platforms?

All these challenges can only be addressed if using a language for capturing scenarios, which is intuitively easy to understand, and, most prominently, which is equipped with a formal (declarative) semantics. As discussed in Section 2, Challenge C1 will be addressed by generalizing from data bases of observed traffic flows. A minimal requirement for checking for completeness is thus the need to formally define, whether a particular observed traffic behavior is already covered or not by the current scenario catalog, thus requiring the definition of a formal satisfaction relation. Moreover, as experienced in the play-out approach for Live Sequence Charts (LSCs) [4, 17] a formal semantics provides a basis for playing out the current scenario catalog, thus generating traffic flows which in an expert can judge for unrealistic or missing real-life traffic flows. As described in Section 3, Challenge C2 requires a formal semantics to identify the gaps between the space of possible worlds described in the scenario catalogue, and the concrete in-field incident or accident. Specifically, forthcoming regulations will require autonomously driving cars to record all those perceived environmental artefacts relevant to trajectory planning as well as the car's trajectory control for a sufficiently long time-period. A formal semantics allows to check the failed scenario(s), offering a basis for refining the scenario specifications to cope with the observed failure in perception or interpretation of the real world. Challenge C3 demands the use of a declarative specification language, where one single scenario specification stands for a possibly extremely large set of real world traffic situations, defined unambiguously through the satisfaction relation. Also, declarative specification languages allow for separation of concerns, such as focusing on particular kinds of critical situations in isolation, knowing that the car can only pass the test if all scenarios are passed. Finally, Challenge C4 can be addressed by automatically synthesizing monitors for compliance testing, using the standardized formal semantics.

5 RELATED WORK

See [32] for a general survey of the state of the art for V&V for ADAS and Automated Driving. This paper provides a summary of discussion and ongoing work in a number of activities involving both institutions of the authors. It builds on the findings of the Pegasus Project in using scenario catalogues for type homologation [19, 24, 27, 37], the SafeTRANS Working Group on Highly Automated Systems [5], a series of bilateral workshops between Bosch and OFFIS, with participants listed under Acknowledgement, and discussions of the Workshop on a future project for Verification and Validation of autonomous vehicles at Renningen September 2016. A scenario driven approach is also part of the strategy of the Enables3 Project funded by the Joint Undertaking ECSEL. Academic Research has proposed a number of different approaches for formal synthesis of controllers for autonomously driving vehicles [3, 8, 15, 38] which fail to address Challenge 1 of the Introduction. Statistical Model Checking has been used as a key tool to address the scalability challenge such as in [11, 16]. The OpenScenario Iniative [33–35] has proposed a baseline for capturing Scenario Catalogs. The formal scenario specification approach of this paper extends these results much as Live Sequence Charts [4] extend the then industry standard language of Message Sequence Charts ITU-T 2011.

6 CONCLUSION

The presented blueprint is a result of discussions within a wide network of automotive safety experts and may serve as a guideline for a systematic approach. This joint network has to follow up in order to establish a common safety assessment involving all relevant stakeholder as societal bodies. Only based on common commitments elaborated within e.g. public funded projects it will be possible to deploy highly automated driving in terms of mitigation of liability risks and societal acceptance. In a next step this process has to be deployed on a European level. One of the biggest challenges will be to accelerate the harmonization process in order to enable early market deployment.

ACKNOWLEDGMENTS

We thank Eckhard Böde, Martin Fränzle, Sebastian Gerwinn, Thomas Peikenkamp from OFFIS, and Peter Heidl, Knoop Michael; Becker-Asano Christian; Oehlerking Jens; Lubitz Florian; Heinkel Hans-Martin; Ahle Elmar; Kueperkoch Stefan; Mueller Martin; Classen Henning; Ahbe Dora; Boukricha Hana; Tiemann Nils; Pauli Bernhard, Poddey from Bosch for joint development of the concepts presented in this paper in a series of workshops.

A EXCERPTS OF THE REPORT OF THE ETHICS COMMISSION OF THE GERMAN MINISTERY OF TRANSPORTATION AND INFRASTRUCTURE

Rule 2: The protection of individuals takes precedence over all other utilitarian considerations. The objective is to reduce the level of harm until it is completely prevented. The licensing of automated systems is not justifiable unless it promises to produce at least a diminution in harm compared with human driving, in other words a positive balance of risks.

Rule 5: Automated and connected technology should prevent accidents wherever this is practically possible. Based on the state of the art, the technology must be designed in such a way that critical situations do not arise in the first place. These include dilemma situations, in other words a situation in which an automated vehicle has to "decide" which of two evils, between which there can be no trade-off, it necessarily has to perform. In this context, the entire spectrum of technological options – for instance from limiting the scope of application to controllable traffic environments, vehicle sensors and braking performance, signals for persons at risk, right up to preventing hazards by means of "intelligent" road infrastructure – should be used and continuously evolved. The significant enhancement of road safety is the objective of development and regulation, starting with the design and programming of the vehicles such that they drive in a defensive and anticipatory manner, posing as little risk as possible to vulnerable road users.

Rule 7: In hazardous situations that prove to be unavoidable, despite all technological precautions being taken, the protection of human life enjoys top priority in a balancing of legally protected interests. Thus, within the constraints of what is technologically feasible, the systems must be

programmed to accept damage to animals or property in a conflict if this means that personal injury can be prevented.

Rule 9: In the event of unavoidable accident situations, any distinction based on personal features (age, gender, physical or mental constitution) is strictly prohibited. It is also prohibited to offset victims against one another. General programming to reduce the number of personal injuries may be justifiable. Those parties involved in the generation of mobility risks must not sacrifice non-involved parties.

REFERENCES

[1] S. Bogomolov, A. Donzé, G. Frehse, R. Grosu, T. T. Johnson, H. Ladan, A. Podelski, and M. Wehrle. 2013. Abstraction-Based Guided Search for Hybrid Systems. In *Model Checking Software: 20th International Symposium, SPIN 2013, Stony Brook, NY, USA, July 8-9, 2013. Proceedings*. Springer, 117–134.

[2] L. Bortolussi and G. Sanguinetti. 2013. Learning and Designing Stochastic Processes from Logical Constraints. In *Quantitative Evaluation of Systems: 10th International Conference, QEST 2013, Buenos Aires, Argentina, August 27-30, 2013. Proceedings*. Springer, 89–105.

[3] S. Coogan and M. Arcak. 2014. Freeway traffic control from linear temporal logic specifications. In *2014 ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*. 36–47.

[4] W. Damm and D. Harel. 2001. LSCs: Breathing Life into Message Sequence Charts. *Formal Methods in System Design* 19, 1 (2001), 45–80.

[5] W. Damm and P. Heidel. 2017. Recommendations of the SafeTRANS Working Group on Highly Autonomous Systems. www.safetrans-de.org/en/Latest-reports/management-summary-for-highly-automated-systems/192. (2017).

[6] W. Damm, S. Kemper, E. Möhlmann, T. Peikenkamp, and A. Rakow. 2018. Traffic Sequence Charts - A Visual Language for Capturing Traffic Scenarios. In *Embedded Real Time Software and Systems - ERTS2018.* submitted.

[7] W. Damm, E. Möhlmann, T. Peikenkamp, and A. Rakow. 2017. A Formal Semantics for Traffic Sequence Charts. In *Festschrift in honor of Edmund A. Lee.*

[8] W. Damm, H.J. Peter, J.H. Rakow, and B. Westphal. 2013. Can we build it: formal synthesis of control strategies for cooperative driver assistance systems. *Mathematical Structures in Computer Science* 23, 4 (2013), 676–725.

[9] W. Damm, G. Pinto, and S. Ratschan. 2007. Guaranteed Termination in the Verification of LTL Properties of Non-linear Robust Discrete Time Hybrid Systems. *Int. J. Found. Comput. Sci.* 18, 1 (2007), 63–86.

[10] P. Dollar, C. Wojek, B. Schiele, and P. Perona. 2012. Pedestrian Detection: An Evaluation of the State of the Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, 4 (April 2012), 743–761.

[11] C. Ellen, S. Gerwinn, and M. Fränzle. 2015. Statistical model checking for stochastic hybrid systems involving nondeterminism over continuous domains. *International Journal on Software Tools for Technology Transfer* 17, 4 (01 Aug 2015), 485–504.

[12] M. Fränzle. 1999. Analysis of Hybrid Systems: An Ounce of Realism Can Save an Infinity of States. In *Computer Science Logic: 13th International Workshop, CSL'99 8th Annual Conference of the EACSL Madrid, Spain, 1999 Proceedings*. 126–139.

[13] M. Fränzle, S. Gerwinn, P. Kröger, A. Abate, and J.-P. Katoen. 2015. Multi-objective Parameter Synthesis in Probabilistic Hybrid Systems. In *Formal Modeling and Analysis of Timed Systems: 13th International Conference, FORMATS 2015, Madrid, Spain, September 2-4, 2015, Proceedings*. Springer, 93–107.

[14] M. Fränzle, T. Gezgin, H. Hungar, S. Puch, and G. Sauter. 2011. Using Guided Simulation to Assess Driver Assistance Systems. In *FORMS/FORMAT 2010*. Springer, 195–206.

[15] C. Frese and J. Beyerer. 2010. *Planning Cooperative Motions of Cognitive Automobiles Using Tree Search Algorithms*. Springer, 91–98.

[16] S. Gerwinn, E. Möhlmann, and A. Sieper. 2018. Statistical Model Checking for Scenario-based verification of ADAS. In *Control Strategies for Advanced Driver Assistance Systems and Autonomous Driving Functions*. to appear.

[17] D. Harel and R. Marelly. 2003. *Come, Let's Play: Scenario-Based Programming Using LSC's and the Play-Engine*. Springer.

[18] T. A. Henzinger. 1996. The theory of hybrid automata. In *Proceedings 11th Annual IEEE Symposium on Logic in Computer Science*. IEEE Computer Society, 278–292.

[19] H. Hungar, F. Köster, and J. Mazzega. 2017. Test Specifications for Highly Automated Driving Functions: Highway Pilot. Presentation at Vehicle Test & Development Symposium 2017 (20.-22.06.2017), Stuttgart, elib.dlr.de/117384/. (2017).

[20] N. Kalra and S. M. Paddock. 2016. Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability? www.rand.org/pubs/research_reports/RR1478.html. (2016).

[21] C. G. Keller, C. Hermes, and D. M. Gavrila. 2011. Will the Pedestrian Cross? Probabilistic Path Prediction Based on Learned Motion Features. In *Pattern Recognition - 33rd DAGM Symposium, Frankfurt/Main, Germany, August 31 - September 2, 2011. Proceedings (LNCS)*, Vol. 6835. Springer, 386–395.

[22] J. F. P. Kooij, N. Schneider, and D. M. Gavrila. 2014. Analysis of pedestrian dynamics from a vehicle perspective. In *2014 IEEE Intelligent Vehicles Symposium Proceedings, Dearborn, MI, USA, June 8-11, 2014*. IEEE, 1445–1450.

[23] L. Laurenti, A. Abate, L. Bortolussi, L. Cardelli, M. Ceska, and M. Kwiatkowska. 2017. Reachability Computation for Switching Diffusions: Finite Abstractions with Certifiable and Tuneable Precision. In *Proceedings of the 20th International Conference on Hybrid Systems: Computation and Control (HSCC '17)*. ACM, 55–64.

[24] K. Lemmer and J. Mazzega. 2017. PEGASUS: Automatisiertes Fahren effektiv absichern. Presentation at 19. VDA Technischer Kongress (05.-06.04.2017), Berlin. (2017).

[25] Federal Ministry of Transport and Germany Digital Infrastructure. 2017. BMVI2017. www.bmvi.de/SharedDocs/EN/PressRelease/2017/084-ethic-commission-report-automated-driving.html. (2017).

[26] S. Puch, B. Wortelen, M. Fränzle, and T. Peikenkamp. 2013. Evaluation of Drivers Interaction with Assistant Systems Using Criticality Driven Guided Simulation. In *Digital Human Modeling and Applications in Health, Safety, Ergonomics, and Risk Management. Healthcare and Safety of the Environment and Transport: 4th International Conference, DHM 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part I*. Springer, 108–117.

[27] A. Pütz, A. Zlocki, J. Bock, and L. Eckstein. 2017. Die Datenbank als Freigabewerkzeug automatisierter Fahrfunktionen im Kreislauf relevanter Szenarien. Presentation at 8. Tagung Fahrerassistenz (22. - 23.11.2017), München. (2017).

[28] S. Ratschan. 2014. Safety verification of non-linear hybrid systems is quasi-decidable. *Formal Methods in System Design* 44, 1 (01 Feb 2014), 71–90.

[29] P. Scovanner and M. F. Tappen. 2009. Learning pedestrian dynamics from the real world. In *IEEE 12th International Conference on Computer Vision, ICCV 2009, Kyoto, Japan, September 27 - October 4, 2009*. IEEE Computer Society, 381–388.

[30] M. Segelken. 2007. Abstraction and Counterexample-Guided Construction of ω-Automata for Model Checking of Step-Discrete Linear Hybrid Models. In *Computer Aided Verification: 19th International Conference, CAV 2007, Berlin, Germany, July 3-7, 2007. Proceedings*. Springer, 433–448.

[31] S. Silvetti, A. Policriti, and L. Bortolussi. 2017. An Active Learning Approach to the Falsification of Black Box Cyber-Physical Systems. In *Integrated Formal Methods: 13th International Conference, IFM 2017, Turin, Italy, September 20-22, 2017, Proceedings*. Springer, 3–17.

[32] J. E. Stellet, M. R. Zofka, J. Schumacher, T. Schamm, F. Niewels, and J. M. Zöllner. 2015. Testing of Advanced Driver Assistance Towards Automated Driving: A Survey and Taxonomy on Existing Approaches and Open Questions. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. 1455–1462.

[33] VIRES Simulationstechnologie GmbH. 2015. OpenDRIVE. (2015). www.opendrive.org www.opendrive.org, accessed: 2017-11-10.

[34] VIRES Simulationstechnologie GmbH. 2016. OpenCRG. (2016). www.opencrg.org www.opencrg.org, accessed: 2017-11-10.

[35] VIRES Simulationstechnologie GmbH. June 29th, 2016. OpenSCENARIO - Bringing content to the road. 2nd OpenSCENARIO Meeting. (June 29th, 2016). www.openscenario.org/docs/OSCUserMeeting20160629pub.pdf www.openscenario.org, accessed: 2017-11-10.

[36] C. F. Wakim, S. Capperon, and J. Oksman. 2004. A Markovian model of pedestrian behavior. In *Proceedings of the IEEE International Conference on Systems, Man & Cybernetics: The Hague, Netherlands, 10-13 October 2004*. 4028–4033.

[37] H. Winner. 2017. Presentation at Automated Vehicles Symposium 2017 (10.-14.07.2017), San Francisco. (2017).

[38] T. Wongpiromsarn, U. Topcu, and R. M. Murray. 2013. Synthesis of Control Protocols for Autonomous Systems. *Unmanned Systems* 1 (2013), 21–39. www.seas.upenn.edu/~utopcu/pubs/WTM-us13.pdf

[39] F. Yan, M. Eilers, M. Baumann, and A. Lüdtke. 2016. Development of a Lane Change Assistance System Adapting to Driver's Uncertainty During Decision-Making. In *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '16 Adjunct)*. ACM, 93–98.

[40] F. Yan, M. Eilers, A. Lüdtke, and M. Baumann. 2016. Developing a model of driver's uncertainty in lane change situations for trustworthy lane change decision aid systems. In *Intelligent Vehicles Symposium (IV), 2016 IEEE*. IEEE, 406–411.

[41] F. Yan, M. Eilers, A. Lüdtke, and M. Baumann. 2017. Building driver's trust in lane change assistance systems by adapting to driver's uncertainty states. In *IEEE Intelligent Vehicles Symposium, IV 2017, Los Angeles, CA, USA, June 11-14, 2017*. IEEE, 529–534. http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=7987634