IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 18, 2020, accepted January 6, 2021, date of publication January 21, 2021, date of current version February 1, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3053159

Criticality Analysis for the Verification and Validation of Automated Vehicles

CHRISTIAN NEUROHR[1], LUKAS WESTHOFEN[1], MARTIN BUTZ[2], MARTIN HERBERT BOLLMANN[3], ULRICH EBERLE[4], AND ROLAND GALBAS[2]

[1]OFFIS, 26121 Oldenburg, Germany
[2]Robert Bosch GmbH, 70465 Renningen, Germany
[3]ZF AG, 88046 Friedrichshafen, Germany
[4]Opel Automobile/Groupe PSA, 65423 Rüsselsheim, Germany

Corresponding author: Christian Neurohr (neurohr@offis.de)

This work was supported by the German Federal Ministry for Economic Affairs and Energy within the project 'VVM - Verification & Validation Methods for Automated Vehicles Level 4 and 5.'

ABSTRACT The process of verification and validation of automated vehicles poses a multi-faceted challenge with far-reaching societal, economical and ethical consequences. In particular, fully automated vehicles at SAE Level 4 and 5 will be expected to operate safely in an arbitrarily complex, infinite-dimensional domain called *open context*. In order to give structure to the open context, we propose a methodical *criticality analysis* that maps an infinite-dimensional domain onto a finite and manageable set of artifacts that capture and explain the emergence of critical situations for automated vehicles. We propose a combined approach of expert-based and data-driven methods to identify relevant phenomena and explain the underlying causalities. Leveraging on abstraction, we define a clearly laid out process that converges towards a manageable set of artifacts based on two assumptions on the nature of traffic. A criticality analysis precedes the design phase of an automated vehicle and is therefore located outside the V-model. As the open context is analyzed independently of a concrete realization, it is relevant for any automated vehicle operating within that domain. Therefore, its results can subsequently be used to derive safety principles and mitigation mechanisms for automated driving and to set up a coherent safety argument for the homologation process.

INDEX TERMS Criticality, verification, validation, homologation, automated vehicles, open context, automotive safety.

I. INTRODUCTION

Automated vehicles (AVs) at SAE Level 4 and 5 [70] are complex systems operating in open context [68]. Therefore, their verification and validation principally necessitates the consideration of all possible traffic situations and influencing factors during the design- and testing phase of such systems. The resulting test space cannot be covered adequately using traditional distance-based statistical approaches to testing [53], [80], usually referred to as *dilemma of completeness*. Recent research projects such as PEGASUS[1] and ENABLE-S3[2] explored a scenario-based approach to verification and validation of AVs at SAE Level 3, where testing is performed by deriving relevant test cases from a manageable set of scenario classes. As a successor of PEGASUS, the project VVM–Verification and Validation Methods for Level 4 and 5 Automated Vehicles[3]–aims at extending this scenario-based approach to AVs at higher levels of automation and more complex environments. In order to define scenario classes that effectively condense verification and validation efforts, the open context needs to be structured systematically.

As one endeavor in this need for structure, the VVM project develops a method, called *criticality analysis*, which analyzes the open context of urban traffic. The method is evaluated for the use case *urban intersection*. The core steps of the criticality analysis, detailed in this publication, include

(i) extracting relevant influencing factors, called *criticality phenomena*,
(ii) improving the understanding of criticality phenomena by identifying underlying *causal relations*,

The associate editor coordinating the review of this manuscript and approving it for publication was Shadi Alawneh.

[1]www.pegasusprojekt.de/en
[2]www.enable-s3.eu
[3]www.vvm-projekt.de/en

(iii) using *abstraction and classification* of causal relations for scenario space condensation.

Figure 1 depicts the criticality analysis as a mapping of the open context onto a finite set of artifacts, as listed above. Long-term goals include the derivation of safety principles and mitigation mechanisms for automated vehicles based on the results of the criticality analysis, in particular on causal relations. However, this will be the focus of future work.

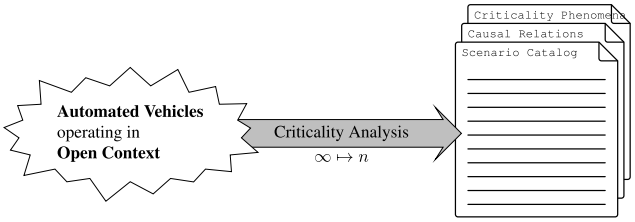


FIGURE 1. The goal of the criticality analysis is to map the infinitely-dimensional open context to a finite set of artifacts by analyzing the underlying structures.

The following section II sketches the automotive verification and validation landscape and relates the presented method to the V-model of ISO 26262. We introduce the high level goals of the criticality analysis, including employed terminology and artifacts, in section III. The basic concept is subsequently presented in section IV. Based on two fundamental but reasonable assumptions on the nature of traffic, we explain our approach to solving the dilemma of completeness. For each step of the criticality analysis we lay out the associated problems and projected solutions in section V. We conclude by proposing future work, in particular, detailing out the steps of the criticality analysis.

II. LANDSCAPE OF AUTOMOTIVE VERIFICATION & VALIDATION

In this section we provide a brief discussion on how the homologation of automated vehicles impacts existing automotive safety processes. Moreover, we introduce the larger context of the criticality analysis, namely the projects VVM, where the criticality analysis is developed, and its sister project SET Level.[4] Both project are successors to the PEGASUS project and, therefore, part of the PEGASUS project family. The section ends with stating our contribution.

A. ADAPTION OF AUTOMOTIVE SAFETY PROCESSES

Although there are a number of advanced driver assistant systems (ADAS), i.e. SAE Level 2, available on the market, the next step towards higher levels of automation, i.e. SAE Level ≥ 3, heavily impacts the processes of verification and validation of such automated vehicles.

For contemporary ADAS, such as advanced cruise control, lane departure warning or emergency brake assistant, intervention is restricted to either lateral or longitudinal control of the vehicle. The human driver is constantly responsible for the driving task [8]. In contrast, automated driving systems (ADS) release the human driver (temporarily) from the driving task and take full control of the vehicle by performing accelerating and braking as well steering actions without any involvement of the human driver. The safety of vehicle operation is therefore the responsibility of the ADS, as the human driver is out of the control loop. Misbehavior of the automation may lead to life-threatening situations [56]. Of course, this responsibility shift presumes valid usage of the ADS. Therefore, the homologation process of such ADS has to assure safe vehicle operation meaning that all relevant hazards are considered and mitigated appropriately.

The functional safety of road vehicles is addressed comprehensively by the well-known automotive safety standard ISO 26262 [47], which focuses on failures and faults of E/E hardware components as well as methods for hazard analysis and risk assessment to identify and quantify the former. The emergence of ADAS made the complementary aspect of the safety of the intended functionality (SOTIF) visible in the automotive industry. The reliance of such systems on sensory input from the open context opens the door to a variety of hazardous situations even in the absense of classical hardware faults. The issue of SOTIF is addressed in the standard ISO/PAS 21448 [48], albeit only up to SAE Level 2.

B. VERIFICATION AND VALIDATION METHODS FOR AUTOMATED VEHICLES

In order to adapt the automotive safety processes for the homologation of higher SAE Levels the VVM project develops methods that facilitate this adaption process. For this, VVM collaborates closely with its sister project SET Level, whose focus is the development of a generic simulation framework to be used for aforementioned V&V activities. Both projects build upon results of the PEGASUS project, which laid the ground work for scenario-based testing of highly automated vehicles (SAE Level ≥ 3). The ultimate goal of VVM is to develop methods and processes that support a safety argument for the homologation of AVs at SAE Level 4 and 5. For exemplary application and proof of concept, the project focuses on the use case 'urban intersection'.

C. CONTRIBUTION OF THE CRITICALITY ANALYSIS

The ISO/PAS 21448 recommends to evaluate the system and its components on unknown hazardous scenarios in the late testing phase (upper-right arm of the V-model), but does not provide a method for identification these unknown hazardous scenarios. For automated vehicles operating in the open context of the traffic world, the set of influencing factors triggering these unknown hazardous scenarios can be arbitrarily large. Finding these influencing factors is a hard task for rule-based systems, but even harder for systems that incorporate machine learning components. The contribution of the criticality analysis is to identify relevant influencing factors for automated vehicles, called criticality phenomena, even before the concept phase (upper-left arm of the V-model) through a systematic analysis of the structure of the open

[4]https://setlevel.de

context. The criticality analysis can be understood as a hazard analysis and risk assessment where the system under consideration is the traffic system in general, and not a concretely defined item. It therefore addresses a whole class of products, that is automated vehicles, instead of a single one. Its results can then be used as a basis for a verification and validation strategy for any product of that class, in particular, to set up a coherent safety argumentation well in advance of testing.

A criticality analysis is expected to be performed by standardization boards and certification bodies in order to gain advanced insights into the emergence of criticality phenomena for automated vehicles. Obtaining such knowledge is a preparatory step for defining a rigorous homologation process. The criticality analysis also addresses corporate and scientific accident researches, as it extends conventional accident research by adding considerable predictive power. This is especially useful when designing an AVs safety concept.

III. HIGH LEVEL GOALS

The overarching analysis goal is the unveiling of how criticality emerges in traffic, generally, and for automated vehicles in particular. For this, we first introduce the term criticality together with several explanatory remarks.

Definition 1 (Criticality): Criticality (of a traffic situation) is the combined risk of the involved actors when the traffic situation is continued.

Remark 1 (Criticality):

(i) In order to determine criticality, probabilities and types of harm, dynamical and behavioral models and actions restrictions of the involved actors are taken into account.
(ii) The time-horizon of criticality of a situation is bound by the fulfillment of the intentions of the involved actors.
(iii) Criticality is inversely correlated with the amount of (sequences of) actions to avoid harm that are available to the involved actors.

While a traffic situation refers to a particular point in time, a scenario describes an evolution over time, cf. subsubsection V-C1. Hence, the definition of criticality can be extended to scenarios by aggregating the criticality of a time sequence of traffic situations. For example, using the maximum or an average over a discrete number of time steps. Note that, for criticality, the system of interest is not a concrete automated vehicle, but the traffic system as a whole. This allows the criticality analysis to derive more universal statements that can be employed in a downstream safety process. We do, however, define a minimal set of high level assumptions on such generic systems for which we analyze criticality. Specifically, we assume that they

1) are regular vehicles found in urban traffic and
2) have a functional sense-plan-act architecture where
   a) the sense-part relies on a set of vehicle sensors (e.g. radar, camera, lidar) and
   b) the plan-part pursues safety and performance goals, adhering to rules and norm behavior.

Based on the above definition of criticality, we propose a break down into six high level goals in order to accomplish the analytical examination of criticality.

(G1) Extract *criticality phenomena*, i.e. observations of traffic that are associated with increased criticality.
(G2) Deliver *explanations* of the criticality phenomena by analyzing the possible underlying causalities.
(G3) Derive a *structuring* of the open context according to these causalities.
(G4) Construct a *catalog of abstract scenarios* based on the classification, including *representative instances*.
(G5) Find an *adequate level of abstraction* for the criticality phenomena, explanations and scenarios.
(G6) Achieve a *convergence* towards a manageable set of criticality phenomena.

Figure 2 gives an overview of the dependencies between the high level goals and their artifacts.

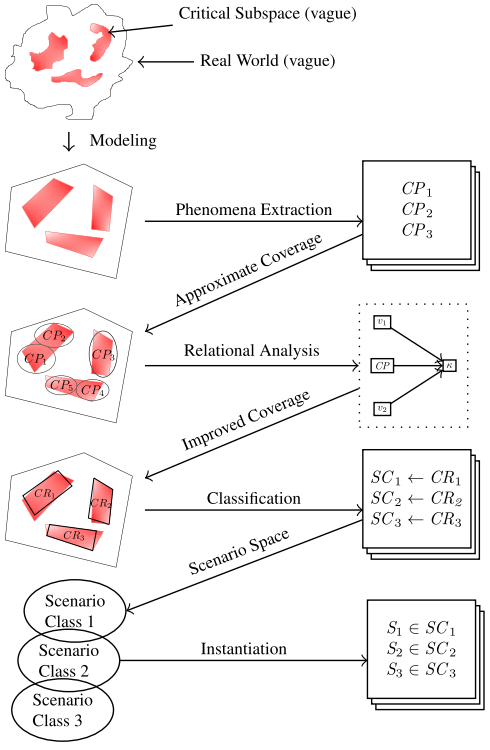


FIGURE 2. Simple graphic representation of the relations between the high level artifacts of the criticality analysis.

Initially, the analyst is confronted with an explicit or implicit model of the highly unstructured traffic world [68]. The state space spanned by this model contains a *critical subspace*. The actual shapes of the critical subspaces can only be made visible by measuring. Such a process is an approximation using explicit or implicit criticality surrogates, e.g. human judgment or computable criticality metrics.

Criticality phenomena (G1, cf. Definition 2) should cover a sufficient part of the critical subspace of the real world, i.e. be indicative of criticality. The quality of this approximation depends on the validity of the employed model and metrics.

Canonically, each phenomenon represents a scenario class, namely those in which the phenomenon is present. As to motivate why the criticality analysis needs to deliver an in-depth explanation of the identified phenomena, consider a scenario catalog on a phenomenological basis. This catalog may not be enough in terms of sufficiency for a safety argumentation: the mere presence of a criticality phenomenon is not necessarily the essence of the criticality in a scenario. For example, the presence of an occluded pedestrian may be uncritical for scenarios where the pedestrian has no physical possibility to enter the driving lane. There may be other – possibly unidentified – criticality phenomena present, e.g. a wet road surface, that lead to an increase in criticality. Albeit this insufficiency, a phenomenological analysis represents a meaningful starting point for subsequent deeper analyses.

We note that each criticality phenomenon has a set of underlying structural relations that lead to an increase in criticality. Those *causal relations* (G2) refer to a set of statements, partially ordered, such that validity of the preceding statements is plausibly causal for a subsequent one. Such relations can similarly construct scenario classes, and capture the core concept more precisely. Continuing the previous example, there exists a synergistic relation between the phenomena 'occlusion' and 'road accessibility'. This leads to a more refined scenario classification, as scenarios that do not include the identified causal relations can be excluded.

Once sufficiently many plausible causal relations are identified, the open context can be structured along the set of all causal relations for all phenomena (G3). For example, one can derive a joint set of all causal relations for all phenomena which then allows to identify even more complex interrelations between phenomena.

Based on this identified structure, the scenario space can be classified into a finite scenario catalog (G4). This involves finding a solution to the problem of representativeness: a sample shall be a sufficient surrogate for the set of all scenarios of that class.

Identifying criticality phenomena and their causal relations requires an adequate level of abstraction (G5). The need for abstraction becomes evident when identifying criticality phenomena in an expert-based approach: one will inherently rely on abstraction to ensure the finiteness of the list. Note that its abstraction level may be too generic to derive useful statements–'using a road is associated with criticality'–or too concrete to be useful in a safety case–'using the roundabout *Place Charles de Gaulle, Paris* at midnight together with two vehicles is associated with criticality'.

Finally, the method shall converge towards a finite and manageable set of phenomena, causal relations and scenarios (G6). The results shall adhere to certain quality requirements, for example, with respect to their explanatory depth.

IV. BASIC CONCEPT

In this paper, we propose a methodical criticality analysis that accomplishes the goals (G1)-(G6) and produces the introduced artifacts. The basic concept of the criticality analysis, as depicted by Figure 3 and presented in the following, will be explained in-depth in section V.

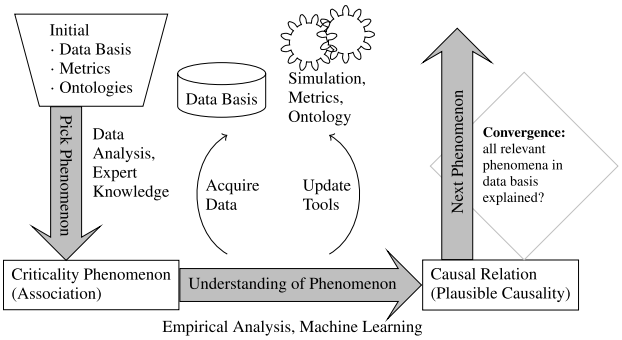


FIGURE 3. Basic concept of the criticality analysis: Pick a criticality phenomenon - Improve understanding of phenomenon - Go to next phenomenon.

The basic concept of the criticality analysis according to Figure 3 can be summarized as: *Pick a criticality phenomenon-Improve understanding of phenomenon-Go to next phenomenon*. Let us elaborate on this three-step-process.

a: PICK A CRITICALITY PHENOMENON

The criticality analysis starts with extracting phenomena that are allegedly associated with increased criticality from the initial information basis, which is the union of all available information about the domain of interest. Within this publication, we are examining the urban traffic domain. Extraction of phenomena is a two-stranded process: a criticality phenomenon can be derived from expert-knowledge such as existing ontologies, approval catalogs or interviews with domain experts, e.g. accident researchers, traffic psychologists and jurists, or through analysis of related data, e.g. accident data bases, under the use of appropriate metrics. Before investigating the phenomenon further, we need to establish its relevance for AVs w.r.t. criticality. Expert-based phenomena need to be supplemented with data and data-based phenomena need to be infused with semantics.

b: IMPROVE UNDERSTANDING OF PHENOMENON

If a criticality phenomenon is deemed relevant, the goal is to improve our understanding of the phenomenon by explaining the underlying causal relation. In particular, we want to build a plausible causal model of how exactly this phenomenon increases criticality. First, an initial hypothesized causal relation explaining the considered phenomenon is specified by an expert. Evidence for the plausibility of the alleged causal relation is gathered using empirical analyses. Iterative learning, based on extending the data basis and continuously updating the ontology, employed metrics and simulation models, leads to an increasingly refined hypothesis supported by accumulating data. If the evidence for the causal relation is statistically sufficient for the ensuing V&V process, it is accepted as a plausible explanation for the phenomenon.

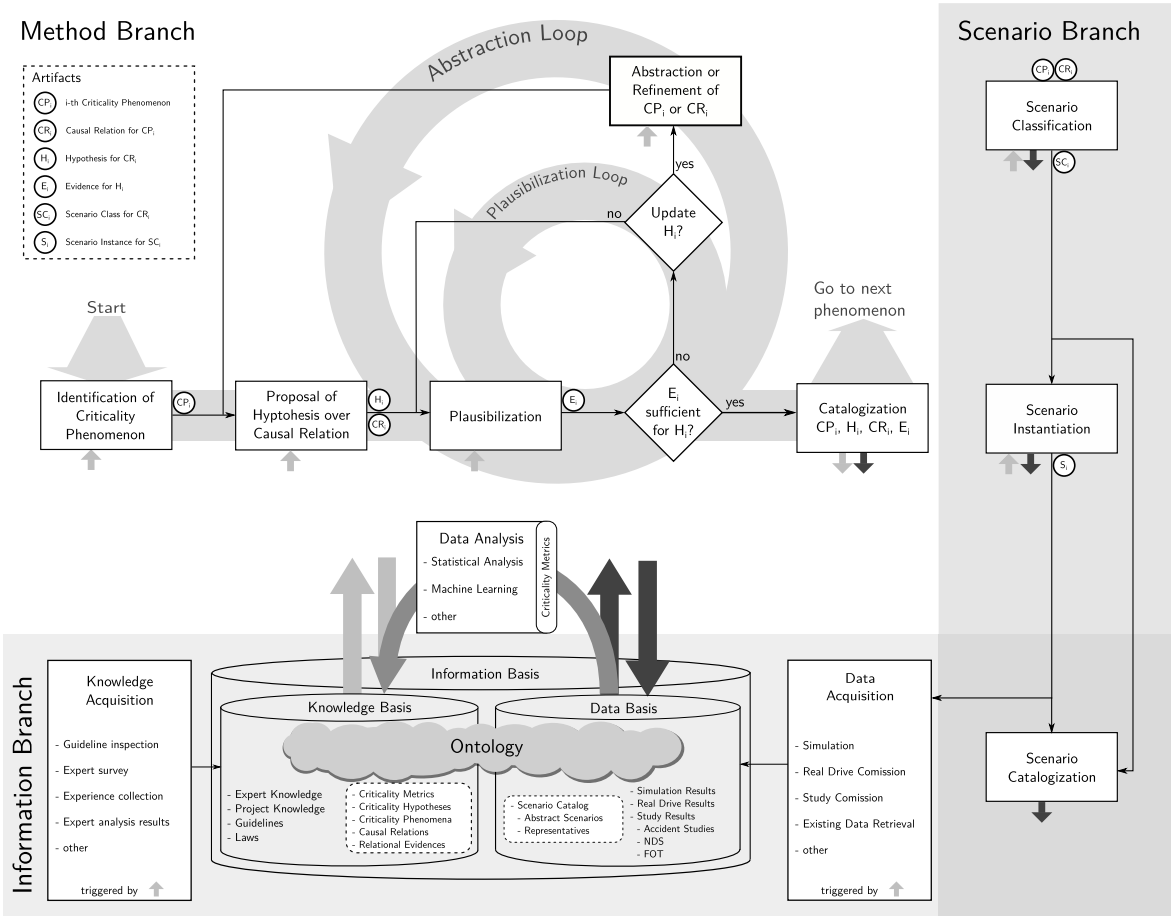


FIGURE 4. **Overview of the procedure of the criticality analysis. It consists of three branches that interact with each other: the method branch (cf. subsection V-A), the information branch (cf. subsection V-B) and the scenario branch (cf. subsection V-C).**

c: GO TO NEXT PHENOMENON

For the criticality analysis to be effective in practice, we need to establish its convergence as a process. Convergence follows from two *fundamental assumptions*:

(A1) The number of relevant criticality phenomena is limited and manageable.
(A2) The relevant criticality phenomena leave traces in a growing information basis.

Assumption (A1) is justified by the observation that, although human drivers have limited driving skills in terms of perception, planning and actuation, they manage to operate reasonably safe in most situations. The advantage of human drivers is their ability to recognize abstract classes of danger, i.e. criticality phenomena, and adapt their driving behavior accordingly. Given the limitations of human consciousness, the number of relevant criticality phenomena cannot be excessively large, if the level of abstraction used to represent the phenomena is chosen adequately. As the relevant criticality phenomena do exist in human consciousness on some level, it is likely that they can be identified using an ever-growing information basis comprised of all different kinds of data recorded with various sensor technologies. Even if relevant phenomena for AVs are different than for humans, they will leave traces once data for automated driving keeps piling up, therefore justifying assumption (A2).

d: CONVERGENCE

Based on the fundamental assumptions (A1) and (A2), the identification of criticality phenomena is expected to converge towards a manageable list of relevant phenomena for automated driving, thus giving structure to the open context. This process is facilitated by the steps of the criticality analysis and their iterative application. In particular, the information basis is expected to increasingly contain data from AVs in field operational tests. Since the impact of AVs on human traffic, i.e. the emergence of mixed traffic, may have profound consequences on the traffic structure, the ever-growing information basis needs to be periodically checked for new criticality phenomena and their relevance to traffic.

V. CRITICALITY ANALYSIS METHOD

This section outlines the process of the criticality analysis, as depicted in the schematic diagram of Figure 4. The basic concept, as introduced in section IV, is expanded into three branches in order to achieve the high-level goals (G1)-(G6)

from section III: method branch, information branch and scenario branch. After a brief introduction to the three branches, the following subsections V-A, V-B and V-C explore the respective branches in detail.

e: METHOD BRANCH

As a more detailed version of the basic concept depicted in Figure 3, this branch is the backbone of the criticality analysis. Essentially, we delineate the process from identification of criticality phenomena over uncovering the underlying causal relations to gathering evidence for their plausibility. Additionally, a subsequent abstraction and refinement loop enables the method to find the fitting level of abstraction for the hypothesized causal relation. All these process steps require input from the information basis, as indicated by the input arrows (light-gray), and may trigger a knowledge or data acquisition. The main tool to generate structured knowledge from the data basis is data analysis, for which we combine criticality metrics with methods from statistical analysis and machine learning. Finally, a catalogization step integrates new artifacts into the information basis.

f: INFORMATION BRANCH

This branch describes the management of information - i.e. knowledge and data - for the criticality analysis and is depicted by the lower part of Figure 4. The center of the information branch is the information basis, which we divide into a knowledge and a data basis. The information branch features two process steps: a knowledge acquisition step that can be triggered by the method branch and feeds the knowledge basis and a data acquisition step that can be triggered by both other strands and feeds the data basis. All process steps of the criticality analysis located outside of the information branch require interaction with the information basis on some level, as indicated by the input and output arrows (light-gray for knowledge and dark-gray for data).

g: SCENARIO BRANCH

The scenario branch describes how scenarios, being the substrate of scenario-based verification and validation, are used within the criticality analysis. For this, we distinguish three process steps: Scenario classification derives abstract classes of scenarios based on the results of the method branch. Scenario instantiation deals with the derivation of more concrete scenarios based on the abstract scenario description, which is required for data acquisition and representative identification. The execution of a scenario – e.g. in simulation – usually requires this step to obtain a logical or even a concrete scenario as an input. Finally, scenario catalogization builds a catalog of critical scenarios that adequately cover all the relevant criticality phenomena and their causal relations.

Subsequently, we detail all branches, using the phenomenon 'occlusion' and its related artifacts as a running example.

A. METHOD BRANCH

The method branch of the criticality analysis, as depicted by Figure 4, is essentially a more detailed version of the basic concept. In the following, we will provide details on the process steps of the method branch. Note that the criticality analysis method operates on the so called *criticality analysis domain*, which defines the search space of the analysis.

1) IDENTIFICATION OF CRITICALITY PHENOMENA

A criticality analysis starts with the identification of artifacts associated with criticality, so-called criticality phenomena.

Definition 2 (Criticality Phenomenon): A criticality phenomenon CP is a concrete influencing factor in a scenario (or a combination thereof) which is associated with increased criticality.

Remark 2 (Criticality Phenomena): Criticality phenomena therefore represent classes of danger.

In our running example, we identify the criticality phenomenon CP_{occ} = 'occlusion'. Per definition, criticality phenomena are related to criticality by association. As to make this association visible, we use criticality properties.

Definition 3 (Criticality Property): Given a temporal logic $\mathcal{L}$, a criticality property $\varphi \in \mathcal{L}$ is a temporal sentence arguing over at least one criticality metric.

For example, $\varphi_{\text{TTC}} = \Diamond\,\text{TTC} < 0.5 \in LTL$ is a linear temporal logic sentence over the Time-To-Collision (TTC) metric [46]. Semantically, it states that at some point in time the TTC falls below 0.5 seconds.

Definition 4 (Scenario Set of a Criticality Phenomenon): $Sc(CP) = \{S \in SC \mid CP \text{ present in } S\}$, where SC is the set of all possible scenarios.

The exemplary phenomenon CP_{occ} thus induces the set of all scenarios in which an occlusion is present: $Sc(CP_{occ})$.

Definition 5 (Satisfaction Relation for Scenarios): For a set of scenarios Sc, a criticality property φ and an extent $\sigma \in [0, 1]$, $Sc \models_\sigma \varphi$ iff. $\forall\, S \in Sc.\ S \models_\sigma \varphi$.

For the definition of $S \models_\sigma \varphi$, we rely on the temporal logic to supply an adequate semantics over the trace of S regarding the extent of satisfaction σ. The simple criticality property φ_{TTC} is satisfied, under LTL semantics, by all scenarios whose traces eventually reach a point in time where the Time-To-Collision falls below 0.5 seconds. Additionally, we introduced the extent of association σ in order to quantify the association of phenomena to criticality. In our example, due to the fact that we have chosen a Boolean logic, the extent of satisfaction σ is 1 for the aforementioned set of scenarios, and 0 otherwise. As to instantiate extents, we can make use of concepts such as *vagueness* – e.g. through fuzzy logics [89] – and *uncertainty* – e.g. through probabilistic logics [44].

Definition 6 (Satisfaction Relation for a Criticality Phenomenon): For a criticality phenomenon CP, a criticality property φ and an extent $\sigma \in [0, 1]$, $CP \models_\sigma \varphi$ iff $Sc(CP) \models_\sigma \varphi$.

Hence, a criticality phenomenon satisfies a criticality property if and only if all its induced scenarios exhibit this

property. In our running example, $CP_{occ} \models_1 \varphi_{\text{TTC}}$ exactly if all scenarios containing an occlusion having a Time-To-Collision of less than 0.5 seconds at some point. Obviously, this relation does not hold – certainly, there exist scenarios with occlusion and a higher Time-To-Collision. At this point, the extent σ can help to quantify the degree of the association to allow a more fine-grained depiction of criticality.

After having semi-formally defined the foundations of criticality phenomena, we turn towards the process of identifying such factors. We propose the following work-flow:

(1) First, acquire and structure knowledge. This includes searching, selecting and rating various available sources. More details on the acquisition of information can be found in subsubsection V-B2.
(2) Search the available knowledge basis for observations associated with criticality. This includes knowledge from domain experts, research and development projects, existing guidelines and laws as well as anecdotes and experiences. If the available knowledge does not adequately cover the entire criticality analysis domain, trigger a knowledge acquisition step.
(3) Initially, describe each criticality phenomenon using the ontological model depicted by Figure 13. This includes potentially interesting abstractions and concretizations, as well as their classification. In order to establish interconnectedness, we recommend denoting the relations to other artifacts and potential synergies, e.g. through tags.
(4) Gather empirical evidence from the data basis in order to estimate the relevance of the phenomena described in (3) for further considerations and to establish confidence in the criticality association. Readjust the level of abstraction, distinguished concretizations and ontological classification according to the available evidence.

After these steps have been carried out by a group of domain experts, the structured results, e.g. in form of a table or database, form the basis for all subsequent considerations.

As an example, we consider CP_{occ} and its concretizations. Initially, we identified occlusion itself at the highest level of abstraction and denoted seven distinguished concretizations, as shown by Table 1. Whether a concretization is listed or not depends on whether its qualitative effect on criticality of traffic is hypothesized to be characteristically different from the others. Note that this initial identification may be updated iteratively during the course of the analysis.

Searching accident databases, such as GIDAS [64], for these different types of occlusion can give empirical evidence that occlusions are indeed relevant phenomena, at least for human drivers. Depending on the sensor setup, occlusions may be relevant for AVs as well. Note that we do not make assumptions on a specific AV realization during the analysis. As we are at the beginning of a criticality analysis, our list of causal relations is currently empty and as such, we determine that we can proceed to the next step of finding an underlying causal relation explaining the phenomenon of occlusion.

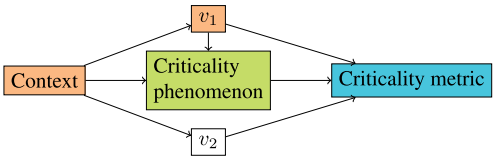


FIGURE 5. An exemplary causal relation.

2) PROPOSAL OF CAUSAL RELATION AND HYPOTHESES

In order to structure the open context of the traffic world according to the emergence of criticality, the first step of the analysis establishes an associative link between the relevant phenomena and criticality. Now, the goal is to infuse those observations with explanations that give plausible reasons for the causal relation between the phenomenon and criticality. To this end, we pick a previously identified phenomenon and improve our understanding of it by exploring the question 'how exactly does this phenomenon influence criticality?'.

From a formal point of view, we can imagine a causal relation as a network of phenomena where each connection between phenomena represents a plausible cause-and-effect relationship. Note that one causal relation might explain several criticality phenomena at the same time, leading to a condensation of artifacts. One method to represent causal relations is the formalism of directed acyclic graphs (DAGs) [65], consisting of nodes that are connected by unidirectional edges without producing circularity. Such causal graphs are employed e.g. in empirical research to explicitly model and communicate assumptions about the world.[5] When designing a study according to a causal graph, those explicit assumptions then allow to control for confounders in a subsequent data analysis. Thus, such causal graphs enable a better approximation when analyzing the causal effect. Additionally, they allow to re-use existing knowledge about other relations that may have already been established.

Due to those advantages, this well-established formalism for specifying and analyzing complex causal relations will be employed for our purposes. Mapping the causal graph terminology to the criticality analysis, nodes represent variables that can be measured in a scenario, including criticality phenomena, and edges correspond to plausible causal implications between these variables. Thus, we define the proposed causal relation as a directed[6] acyclic graph as following:

Definition 7 (Causal Relation): A causal relation $CR = (P, E)$ is a graph where P is the set of nodes – variables described by propositions on traffic scenarios – and $E \subseteq P \times P$ is the set of edges – causal links between the variables.

The referenced variables of the nodes can be defined through the ontological basis as detailed out in subsection V-B. A simple causal relation is shown by Figure 5. The *context* variable is causal for a criticality phenomenon and two

[5]Tools like DAGitty can be used for the analysis of causal graphs [76].

[6]Note that the assumption of acyclicity may not be justified for all variables – think of two traffic participants iteratively influencing each other with their driving behavior. However, our goal here is not to model in detail all circular implications between phenomena, but rather to model the essential paths from a phenomenon to criticality.

TABLE 1. Tabular specification of the phenomenon 'occlusion', including some of its concretizations.

Criticality Phenomenon	Ontological Classification	Estimated Criticality	Tags
Occlusion	Perception	Medium	Limited Perception
Occluded Pedestrian	Perception	High	Limited Perception, Vulnerable Road User
Occluded Bicyclist	Perception	High	Limited Perception, Vulnerable Road User
Occluded Intersecting Vehicle	Perception	Medium	Limited Perception, Trajectory
Occluded Obstacle	Perception	Medium	Limited Perception, Obstacle
Occluded Lane Markings	Perception	High	Limited Perception, Lane Markings
Occluded Traffic Sign	Perception	Depends	Limited Perception, Traffic Sign
Occluded Traffic Light	Perception	High	Limited Perception, Traffic Light

variables v_1, v_2 which in turn influence a criticality metric. If we aim to identify the causal effect of the exposure variable (criticality phenomenon) on the outcome (criticality metric) appropriately, we need to take care of the confounding variables *context* and v_1.

Continuing the running example, we postulate a causal relation $CR_{stat-occ-tp}$ for a stationary occlusion of a traffic participant, $CP_{stat-occ-tp}$, as depicted by Figure 6. Note that $CP_{stat-occ-tp}$ is an concretization of CP_{occ}, but not necessarily a strict abstraction of any other concretization listed in Table 1 due to the occlusion being stationary.

For easier depiction of the connections between criticality phenomena, causal relations and scenarios, we semi-formally introduce the following relations. Analogously to criticality phenomena, it is reasonable to define the set of scenarios induced by a causal relation.

Definition 8 (Scenario Set of a Causal Relation): $Sc(CR) = \{S \in SC \mid CR \text{ present in } S\}$, where SC is the set of all possible scenarios.

For occlusion, $SC(CR_{stat-occ-tp})$ is, simply put, the set of all scenarios where the static occlusion influenced the participant's perception and subsequently their behavior which lead to an increase in the measured criticality.

As we are interested in whether the proposed causal relations explains the criticality induced by the phenomena, we are defining the explanation relation $\vdash_\rho$ over both scenarios and criticality phenomena.

Definition 9 (Explanation Relation for Scenarios): For a causal relation CR, a set of scenarios Sc, a criticality property φ, an extent ρ and an association $Sc \models_\sigma \varphi$, $CR \vdash_\rho Sc$ iff CR explains the association $Sc \models_\sigma \varphi$ to the extent ρ.

Definition 10 (Explanation Relation for a Criticality Phenomenon): For a causal relation CR, a criticality phenomenon CP, a criticality property φ, an extent $\rho \in [0, 1]$ and an association $CP \models_\sigma \varphi$, $CR \vdash_\rho CP$ iff CR explains the criticality association of $CP \models_\sigma \varphi$ to the extent ρ.

In our running example, we then examine whether the explanation relation $CR_{stat-occ-tp} \vdash_\rho CP_{stat-occ-tp}$ holds to an acceptable extent ρ. The analysis of the extent is subject to the later presented plausibilization methods.

After identifying a possible causal relation CR, a hypothesis H over CR needs to be postulated for analyzing the plausibility of the causal relation.[7]

[7]For the postulation of such a hypothesis, well-understood scientific requirements, such as testability, shall be followed [71]

Definition 11 (Hypothesis over a Causal Relation): A hypothesis H over a causal relation CR, a criticality phenomenon CP and an extent ρ, for which supposedly $CR \vdash_\rho CP$ holds, is defined as $H = (CR, \rho, CP, S)$ where S is a set of additional statements over CR, ρ and CP.

The set of additional statements S can, for example, contain information about the proposed strength and confidence of the causal relation. In our running example, we define $H_{stat-occ-tp} = (CR_{stat-occ-tp}, 1, CP_{stat-occ-tp}$, 'holds for all scenarios with one occluded *tp*').

3) PLAUSIBILIZATION OF CAUSAL RELATIONS

a: ESTABLISHING PLAUSIBILITY

The plausibilization of hypotheses and causal relations carries much of the weight of the method branch. In order to check whether a causal relation adequately explains how a phenomenon influences criticality, we have to gather evidence that supports the corresponding hypothesis H. Plausibilization of CR means that, for each alleged causal link $e = (A, B) \in E$, we provide evidence using either deductive or inductive reasoning.

If for an edge $e = (A, B) \in E$, there exists a possible line of deductive reasoning as to why A is plausibly causal for B, this possibility should be explored foremost. Of course, this highly depends on the nature of the phenomena A and B. Again, as an example, we consider the causal relation $CR_{stat-occ-tp}$ from Figure 6. For many edges in this DAG, convincing lines of deductive reasoning exist such as ('Existence of *ego*','Start position of *ego*') and ('Context','Target speed of *ego*'). Deductive reasoning works particularly well whenever B clearly is dependent on A, or B is a subclass of A. Of course, if more details are required on how the relation actually works, this must be researched additionally.

If for an edge $e = (A, B) \in E$ a deductive line of reasoning cannot be followed directly, we need to gather empirical evidences for e on a representative set of data, which can be generalized using inductive arguments. As an example, consider the edge ('average degree of occlusion of *tp* for *ego*', 'Non-existence of representation of *tp* in world model of *ego*'). In principle, various types of evidence can be used in this step such as observational data from field operational tests, stationary traffic measurements or even driving simulator studies. If insufficient data is available, a data acquisition step may need to be triggered. However, in order to infer causality for an edge $e = (A, B)$, more rigorous evidence

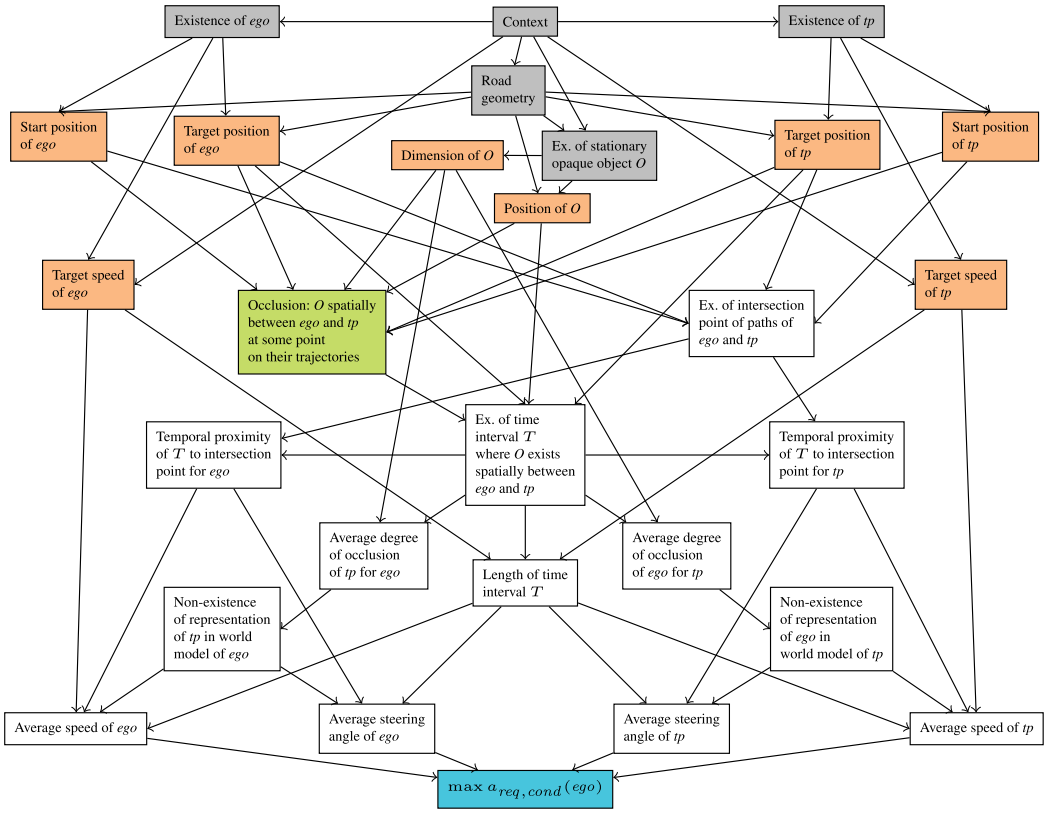


FIGURE 6. Causal relation $CR_{stat-occ-tp}$, represented as a DAG, connecting the criticality phenomenon $CP_{stat-occ-tp}$ to criticality measured via conditional required acceleration ($a_{req,cond}$). Unobserved variables are gray and independent variables are orange. The exposure variable 'occlusion' is marked green. The outcome variable 'max $a_{req,cond}(ego)$' is marked blue.

might be required. This poses a problem since for hypotheses about the development of criticality in traffic, randomized controlled trials are not only unethical (due to conjecturing about the causal relation behind potentially life-threatening events) but also impossible to conduct properly due to the randomness inherent to the open context. However, this might be resolved using a combination of high-fidelity simulation environments and statistical hypothesis testing.

Therefore, plausibilization of causal relations requires the following with regard to criticality metrics:

- Employing metrics that make criticality measurable. The validity of the employed criticality metrics is a crucial factor, c.f. subsubsection V-A6.
- Choosing the right criticality metric for the hypothesis at hand. There exist a myriad of available metrics and depending on the causal relation a certain metric may be required to make that effect visible.
- Criticality metrics are not evaluated on hypotheses or causal relations, but on scenarios. Hence, if we employ inductive reasoning, to test the hypothesis, we need to examine the alleged causal relation for a scenario set.

b: ESTIMATING CAUSAL EFFECTS

One possibility of generating evidence for causal relations and, in particular inductive reasoning, is to employ the framework of statistical hypothesis testing which is a methodical pillar in various empirical sciences such as epidemiology [18]. As already indicated previously, conducting randomized controlled trials in real traffic is infeasible due to ethics and complexity. A potential solution to this dilemma is the use of high-fidelity simulation environments. For example, in a simulation, we can easily control the exposure (e.g. occlusion), randomize the confounders and measure the associated outcome (e.g. mean difference in measured criticality).

In a first step for the estimation of causal effects, we map the terminology used for statistical experimentation to the terminology of scenario-based verification and validation:

- *Population*: The set of all possible scenarios, or a certain subset for which we want to make statements about its criticality. Eventually, we are able to infer criticality statements for the set of all scenarios.
- *Sample*: A set of scenarios drawn from the population.
- *Representative sample*: A set that represents the features of the whole population well.
- *Variable*: Features that can occur in a scenario, e.g. the type of humidity or coefficient of friction.
  - *Independent variable*: A scenario feature of a scenario that the experimenters can control.
  - *Dependent variable*: A scenario feature whose outcome is of interest. For us, this is the criticality as measured by some criticality metric.

- *Confounding variable*: A scenario feature that influences the input parameters and criticality.
- *Unobserved variable*: A scenario feature that is not measured in the experiment. For example, if during a real world drive humidity is not measured, then humidity is an unobserved variable.

We can then use this mapped terminology to outline a general strategy for the estimation of a causal effect. In our running example, we use the following steps in order to generate evidence for the causal link between the criticality phenomenon $CP_{stat-occ-tp}$ and criticality:

1) Model the assumptions of the experiment using a causal graph; this has been done in Figure 6.
2) Identify all confounding variables through graph analysis. For $CR_{stat-occ-tp}$ one minimal sufficient adjustment set consists of the variables from Figure 6 that are marked in orange. These confounders influence both the exposure as well as the outcome variable.
3) Design the experiment to observe the effect from exposure to outcome. For our example, we plan to conduct a simulation-based observation of all previously identified confounding variables, whether an occlusion occurred and the criticality metric $a_{\text{req,cond}}$.
4) Conduct the experiment. For this, we can use a probabilistic simulation to sample representatively from the population of all scenarios. We measure the variables previously mentioned and store those measurements. Note that it is essential for the identified variables to be measurable on some scale. More precisely, we need to be able to formulate properties and a degree of satisfaction similar to Definitions 3 and 5. Otherwise, we cannot derive useful statements with regards to the plausibility of the causal relation and effect size.
5) Finally, we are able to apply methods from the area of data analysis, as described in subsubsection V-A7. These include, for example, methods from descriptive statistics, such as statistical hypothesis testing, and causal estimation methods by controlling for the identified confounders and fitting a regression model to the data in order to estimate the effect size [69].

Note that due to the high possibility of encountering unobserved confounders in the open context, methods need to be in place to tackle unmeasured variables. We will investigate this problem in future research.

4) UPDATE HYPOTHESIS

If, after investigating a causal link either deductively or inductively, the available evidence does not sufficiently support the causal link, the hypothesis for the causal relation may need to be updated or even refused. Possible updates for hypothesis adjustment include the introduction of additional edges and the modification of existing edges by means of abstraction and refinement. For example, $CR_{stat-occ-tp}$ went through various iterations before its depicted form.

5) ABSTRACTION AND REFINEMENT

Besides generating evidence for an established hypothesis, another key concept in handling the open context is the identification of a meaningful level of abstraction for criticality phenomena and causal relations. Specifically, we need a level that is concrete enough to enable an identification of mitigation strategies for critical situations, but is also abstract enough for the artifacts to be manageable in size. Figure 7 shows two causal relations being connected via abstraction.

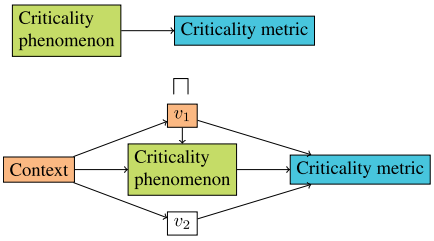


FIGURE 7. A refinement of an abstract causal relation into the more concrete causal relation introduced in Figure 5.

In formal methods, abstraction has long been employed to tackle large and even infinite state spaces [6]. Essentially, instead of analyzing the concrete model, the analysis is run on a smaller and more abstract model. Ideally, such an abstraction is constructed automatically, e.g. through counterexample-guided abstraction refinement [16].

a: ABSTRACTION

Abstraction is the process of bundling a set of criticality phenomena and causal relations into a single criticality phenomenon and causal relation while sufficiently preserving the explanatory strengths. Formally, assume we are given a criticality property φ and a set $\mathcal{C}$ of quadruples of association extents, criticality phenomena, explanatory strengths and their causal relations with $\forall\ (\sigma, CP, \rho, CR) \in \mathcal{C}.\ CP \models_\sigma \varphi \wedge CR \vdash_\rho CP$. We are then interested whether there exist a more abstract phenomenon CP', criticality association σ', causal relation CR' and explanatory extent ρ' for which all four abstraction conditions hold for all $(\sigma, CP, \rho, CR) \in \mathcal{C}$:

(AC1) $Sc(CP') \supseteq Sc(CP)$ (Valid phenomenon abstraction)
(AC2) $CR' \vdash_\rho CP'$ (Valid causal relation abstraction)
(AC3) $CR' \vdash_{\rho'} CP$, $\rho_t < \rho' \leq \rho$ (Sufficiently preserved causal explanation)
(AC4) $CP' \models_{\sigma'} \varphi$, $\sigma_t < \sigma' \leq \sigma$ (Sufficiently preserved criticality association)

where σ_t and ρ_t are acceptable lower thresholds for the extent of the association resp. explanation of the abstract phenomenon and causal relation. Thus, it is expected to lose both associative as well as explanatory power through the process of abstraction, but only up to a certain degree. If abstractions become too generic, phenomena will only be associated weakly with criticality and causal relations will not explain the phenomena sufficiently. We denote any abstraction relation satisfying (AC1) to (AC4) over both the set of criticality phenomena and causal relations by $\sqsubset$.

Let us examine the simple property $\varphi_{TTC} = \Diamond\ TTC < 0.5\ s$. The previously introduced phenomenon $CP_{stat-occ-tp}$ is abstracted to $CP_{occ'}$ = 'there exists an entity that is not perceived by *ego*'. The underlying causal relation explaining occlusion, $CR_{stat-occ-tp}$, is presented in Figure 6. It can be abstracted to the causal relation $E_{occ'} = \{(O$ not represented in *ego*'s world model, Average speed of *ego*$)$, $(O$ not represented in *ego*'s world model, Average steering angle of *ego*$)\}$. Obviously, the phenomenon's abstraction is valid in the sense of (AC1): all scenarios with occlusion are an instance of the perception not being able to perceive something, i.e. $Sc(CP_{stat-occ-tp}) \subseteq Sc(CP_{occ'})$. It is also reasonable to assume that $CR_{occ'} \vdash_\rho CP_{occ'}$ due to the generality of the explanation. Hence, (AC2) holds.

Subsequently, the question of the abstraction's usefulness arises naturally. For this, we examine (AC3) and (AC4): we find that the causal relation $CR_{occ'}$ is not sufficient to explain the criticality of occlusion to a necessary depth. For instance, it takes neither the type (e.g. a bicyclist versus a mosquito) nor the proximity of the occluded object to the ego vehicle (e.g. a few centimeters versus a few hundred meters) into account. Therefore, and based on our threshold ρ_t for the extent of the explanation, we find that the causal explanation is not sufficiently preserved. Additionally, it is imaginable that the criticality association $CP_{stat-occ-tp} \models_\sigma \varphi$ is not preserved, i.e. $CP_{occ'} \not\models_\sigma \varphi$. This is due to the fact that there were occluded objects in every single scenario that has ever occurred, for example people in buildings, animals in trees or simply rocks below the road's surface. As stated earlier, those objects are clearly causing a TTC less than $0.5s$, but the abstract phenomenon does not capture this relevancy adequately. We thus can not expect to find a large statistical association between such a generic unperceivability and φ. Consequently, the abstract phenomenon does not preserve the criticality association in the sense of (AC4). We conclude that we have not yet produced a useful abstraction.

b: REFINEMENT

In case of over-abstraction, statements become too generic. Refinement is a viable option to achieve a better explanatory quality. We define refinement to be the process of creating a set of more concrete criticality phenomena and causal relations given a criticality property φ, an association extent σ, an abstract criticality phenomenon CP, an extent ρ and a causal relation CR with $CP \models_\sigma \varphi \wedge CR \vdash_\rho CP$. Analogously to the abstraction process, we are then interested whether a criticality association σ' as well as a set of refined associations, criticality phenomena, explanatory strengths and causal relations $\mathcal{C}$ exists, for which the four refinement conditions holds for all $(\sigma', CP', \rho', CR') \in \mathcal{C}$:

(RC1) $Sc(CP) \supseteq Sc(CP')$ (Valid phenomenon refinement)
(RC2) $CR' \vdash_{\rho'} CP'$ (Valid causal relation refinement)
(RC3) $(\bigcup_{(CP',\rho',CR')\in\mathcal{C}} CR') \vdash_{\rho'} CP, \rho \leq \rho'$ (Increased causal explanation)
(RC4) $CP' \models_{\sigma'} \varphi, \sigma \leq \sigma'$ (Increased criticality association)

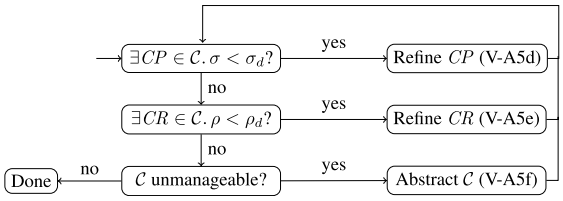


FIGURE 8. The decision flow for determining the direction of refinement or abstraction, given a set of artifacts $\mathcal{C}$.

where $\bigcup$ is the join operator over graphs, i.e. $(V_1, E_1) \cup (V_2, E_2) = (V_1 \cup V_2, E_1 \cup E_2)$. Analogously to abstraction, we again denote any refinement relation by $\sqsupset$.

As an example, let us consider the abstract phenomenon CP_{drive} = 'driving on a road'. It is associated with criticality to some extent $\sigma > 0$ – but it does not provide a meaningful insight into criticality. In the end, we cannot derive useful safety principles or mitigation mechanisms from this, besides the trivial principle of avoiding to drive at all. Thus, a refinement to multiple more concrete phenomena, e.g. including the phenomenon $CP_{stat-occ-tp}$, will provide value in finding more adequate strategies for criticality mitigation.

c: CHOOSING THE DIRECTION OF ABSTRACTION AND REFINEMENT

At a given point in the analysis, it needs to be checked whether an abstraction or refinement is advisable. Once the decision is made, new artifacts need to be produced that satisfy the abstraction and refinement conditions.

Starting with the question of the direction of abstraction and refinement, it is important to note that the experts establishing the initial list of criticality phenomena have already set an original level of abstraction. Hence, for a given phenomenon CP that has a criticality association to some extent σ, i.e. $CP \models_\sigma \varphi$, we need to decide whether σ is sufficient. This can be done by means of a predetermined, desired association extent σ_d, which is then checked against σ, i.e. $\sigma_d \leq \sigma$. The threshold for such an extent is to be determined a-priori, answering the question to which level of detail criticality shall be identified. Obviously, setting such a threshold heavily relies on the downstream usage of the artifacts and desired level of safety, involving complex societal and ethical issues. For the criticality analysis, we assume that appropriate discussions have been conducted, leading to σ_d. Similarly, the extent of the desired explanatory power, ρ_d, needs to be identified a-priori and checked against the current level of explanation, ρ, i.e. checking $\rho_d \leq \rho$.

As indicated by Figure 8, if one of the inequations is violated, a refinement is advisable either for the phenomenon or causal relation. This is due to the fact that criticality association or explanatory strength has to be increased. Otherwise, we conclude the artifacts to be sufficiently concrete. We finally check if the phenomena list has grown too large, implying that an abstraction may reduce complexity.

Once the direction of abstraction or refinement is determined, new artifacts for each of the three steps have to be created. This process is presented in the following paragraphs.

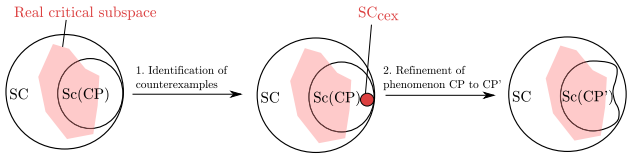


FIGURE 9. Schematic depiction of the counterexample-guided abstraction refinement process for criticality phenomena. *SC* is the set of all scenarios, *Sc(CP)* are scenarios with phenomenon *CP* present and SC_{cex} are counterexample scenarios not or only weakly associated with criticality.

d: REFINE CP

Based on existing approaches, we propose a counterexample guided abstraction refinement method [19], [22].

As shown in Figure 9, we initially note that $CP \not\models_\sigma \varphi$ for all $\sigma \geq \sigma_d$, i.e. the current extent of association with criticality is not sufficient. Hence, we can, e.g. through simulation, identify a set of counterexample scenarios $SC_{cex} \subseteq \{Sc \in Sc(CP) \mid \nexists\ \sigma \geq \sigma_d.\ Sc \models_\sigma \varphi\}$. Those scenarios hence partially contain the reasons for the abstract scenario's weak association with criticality. We are then left with two options:

1) Conclude that even a concretization of CP can not be associated with well enough criticality and remove it from the list, e.g. if $Sc(CP) = SC_{cex}$.
2) Refine the criticality phenomenon to exclude those counterexample scenarios, i.e. $Sc(CP) \setminus SC_{cex}$.

We then use the new set of scenarios $Sc(CP) \setminus SC_{cex} = Sc(CP')$ to identify more concrete associations with criticality. They induce a new phenomenon CP' which is a refined phenomenon with a greater criticality association, as a whole class of scenarios with a low associative extent has been excluded. In order to define CP', we can, for example, learn, e.g. through a classifier, new features for this set of induced scenarios and possibly even refine it into multiple subclasses.

Continuing our running example, we previously argued that CP_{occ} may not satisfy a desired associative extent. We can identify a class of counterexample scenarios where the *ego* is driving straight and an occluded pedestrian walks behind parked cars, but has no possibility of reaching the lane. Here, criticality is not meaningfully increased. We can then define CP'_{occ} to be the conjunction of the seven concretizations of occlusion in table Table 1. The associative extent of these phenomena can be experimentally or analytically shown to be increased when compared to $Sc(CP_{occ})$.

e: REFINE CR

Assume we are confronted with a phenomenon which is explained plausibly albeit not concretely enough, the underlying causal relation needs to be refined, cf. Figure 7.

Similar to refining criticality phenomena, we can also use counterexample guided abstraction refinement for causal relations. Note that at this point in the process, it has been established that $CP \models_\sigma \varphi$ with some $\sigma \geq \sigma_d$, but $CR \vdash_\rho CP$ holds only for $\rho < \rho_d$. Here, we are interested in refining the explanations, i.e. finding counterexample scenarios where the causal relation is not sufficient to explain the increased criticality association of CP. We thus identify such a scenario set $SC_{cex} \subseteq \{Sc \in Sc(CP) \mid \nexists\ \rho \geq \rho_d.\ Sc \vdash_\rho CR\}$ through e.g. simulation or a database search. Once counterexample scenarios have been identified, we can refine by

1) adding nodes to CR or
2) adding edges to CR or
3) refining the variables of the nodes of CR.

One possibility is to use the ontology's classifications and relations for the refinement. Another option, as sketched in [20], is to use machine learning techniques to derive new explanatory statements over the counterexample scenarios.

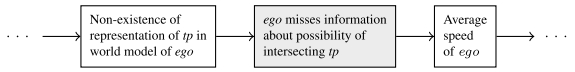


FIGURE 10. Refinement of the causal relation of Figure 6. A node (gray) is added stating that the absence of information is causal for occlusion to be a criticality phenomenon.

Consider the exemplary causal relation $CR_{stat-occ-tp}$ of Figure 6. Here, we may find that (Non-existence of representation of *tp* in world model of *ego*, Average speed of *ego*) may be a bottleneck in the explanation process. This means that there exists a set of scenarios where the direct influence of the non-existence in the world model cannot explain the resulting average speed. As an illustration, we consider a counterexample scenario Sc_{cex} with an occluded traffic participant for which no representation is present in the world model of *ego*. But in those scenarios, the *ego* has acquired the knowledge that, due to the opaque object O, there is an occluded area. It thus believes that there is a possibility of *tp* existing behind O. The *ego* adapts its speed accordingly and successfully mitigates a critical situation.

For the refinement of this counterexample, we use an ontological approach. The ontology provides us with knowledge about the perception chain, e.g. that faults in the feature extraction may result in their negligence in the scene modeling and hence in an erroneous world model. In an expert-based approach, we use this information to identify that the criticality of the phenomenon $CP_{stat-occ-tp}$ is additionally dependent on whether the *ego* has the information about the possibility of an occlusion. This is depicted by the refined causal relation in Figure 10. Here, a node was added stating that the non-existence of a representation of *tp* leads to the *ego* not having information on the possibility that some intersecting *tp* may exist. As the counterexample scenario Sc_{cex} explicitly states that the *ego* has such information, we have effectively excluded it from $Sc(CR_{stat-occ-tp})$.

f: ABSTRACT $\mathcal{C}$

Increasing the level of abstraction may be necessary in case the number of concretizations grows too large. In this case, we can merge phenomena that can be explained by the same causal relations. For example, let us suppose that 'occluded traffic sign' and 'occluded traffic signal' are both present in $\mathcal{C}$ and their criticality explanation is 'presence of an important traffic rule can not be perceived by the *ego*'. We are then able

to abstract 'occluded traffic sign' and 'occluded traffic signal' into 'occluded traffic rule signalization', as the extent of this explanation was previously deemed as accepted.

An additional tool in the abstraction of traffic phenomena is the usage of domain knowledge formalized by means of an ontology, as presented in subsection V-B. Here, the hierarchy provides an essential tool to enable abstraction also over causal relations. In our running example, our ontology states that pedestrians and bicyclists are subsets of vulnerable road users. One option is then to merge those phenomena into a new, abstract phenomenon 'occluded vulnerable road user'. Obviously, (AC1) to (AC4) need to be checked, e.g. through a formal argument, simulation or data analysis.

6) CRITICALITY METRICS

In section IV we introduced a qualitative definition of *criticality* in order to explain the basic concept of the criticality analysis and its high level goals. Going into more detail, in particular for the plausibilization of causal relations, we argued that criticality needs to be measurable in order to collect empirical evidences. The main tool for this are *criticality metrics*. We already encountered some criticality metrics in the beginning of this section, namely TTC and $a_{\text{req,cond}}$.

Definition 12 (Criticality Metric): A criticality metric is a function, evaluated on a (discrete) set of measurements (in the case of real-world data) or values generated by a simulation engine (in the virtual case) with the goal to evaluate the criticality of a traffic situation or scenario.

Thus, criticality metrics are excellent tools to assess criticality, condensed into a single number. Depending on the employed metrics and associated criticality thresholds, the union of all measurements delivers a coverage of the actual, unknown critical subspace, as depicted by Figure 1. The choice of such metrics and thresholds, however, introduces a significant bias into the criticality analysis. Subsequently, any measured criticality and classification based on it is always to be seen relative to the employed metrics and thresholds.

a: APPLICATIONS WITHIN THE CRITICALITY ANALYSIS

As depicted by Figure 4, criticality metrics are inherently attached to the process step 'Data Analysis', but are employed at various other occasions. Therefore, within the criticality analysis, we identify the following applications for metrics:

(AP1) For *plausibilization* of causal relations (cf. subsubsection V-A3), we need to make criticality measurable in order to gather empirical evidence. Plausibilization can therefore trigger a data analysis step. Note that each causal relation may require different criticality metrics in order to measure its effects appropriately.

(AP2) In a data acquisition step (cf. paragraph V-B2.d), to enlarge the available data basis by commissioning real world drives, huge amounts of data are produced. For the identification of critical scenarios, criticality metrics can be applied either live for *selective data recording* during recording or for *data filtering* afterwards.

(AP3) More generally, criticality metrics can be used for any kind of *data analysis* (cf. subsubsection V-A7) to identify critical situations or scenarios in already recorded data from stationary measurements, naturalistic driving studies or accident databases. The results of such an analyses provide evidence for the relevance of phenomena, if they can be identified within the data set.

(AP4) For *scenario classification* (cf. subsubsection V-C2) it is of interest to label scenarios as critical or uncritical. This requires evaluating criticality metrics on scenarios together with thresholds that enable such a labeling.

(AP5) When optimization algorithms are used for *scenario instantiation* (cf. subsubsection V-C3), criticality metrics can be used as objective function [63]. This approach is particularly useful to identify critical parameter combinations in a parameterized scenario class.

(AP6) During the *refinement process* (cf. paragraph V-A5.d, paragraph V-A5.e), criticality metrics enable us to exclude scenarios with a low criticality and thus refine criticality phenomena and causal relations.

(AP7) During the *abstraction process* (cf. paragraph V-A5.f), criticality metrics allow to check whether the extent of the criticality association and its explanation has been decreased only by a tolerable level.

b: REQUIREMENTS ON CRITICALITY METRICS

When analyzing traffic in the real world or in a simulation it is obvious that scenarios vary in several ways w.r.t. the emergence of criticality. Besides the ranges and definitions of the parameters describing the logical level of a scenario, these parameters have a large impact on the associated hazards and risks, e.g. harm done to passengers. Generally, criticality metrics shall reflect the real criticality accurately as to get an assessment of how safe or unsafe a traffic situation or scenario is. While requirements on criticality metrics depend on the desired application, we refer to the three requirements defined by Junietz *et al.*, which roughly cover the requirements imposed by the criticality analysis [51]. Depending on the scope, other sources elicit slightly different requirements [43]. Additional requirements on criticality metrics are likely to be exhibited when executing a criticality analysis.

c: ANALYSIS, CALIBRATION AND ENGINEERING OF CRITICALITY METRICS

As mentioned previously, there exist a myriad of criticality metrics in the literature. One goal of the criticality analysis is to analyze these metrics with respect to the previously referenced requirements and their applications (AP1)-(AP7) within the criticality analysis. For an non-exhaustive

overview on the literature, we consider the following classification of existing metrics with respect to their outputs:

- *Time-Scale Metrics:* Time-To-Collision (TTC) [46], Time-To-Maneuver (TTM) [75], Time-To-React (TTR) [75], [82], Time Headway (THW) [36], [49], Post-Encroachment-Time (PET) [2], [66], Time-To-Closest-Encounter (TTCE) [29]
- *Distance-Scale Metrics:* Headway (HW) [49], Distance-To-Closest-Encounter (DTCE) [29]
- *Motion-Scale Metrics:* Required (Longitudinal / Lateral) Acceleration (a_{req}) [49], [54]
- *Ratio-Scale Metrics:* Brake Threat Number (BTN), Steer Threat Number (STN) [30]
- *Probability-Scale Metrics:* Crash Probability via Monte Carlo methods [12], [31] [3], Crash Probability using Stochastic Reachable Sets [4], Crash Probability by Scoring Multiple Hypothesis [62]
- *No-Scale Metrics:* Minimum Driving Task Difficulty Score [51], Potential Functions as Superposition of Scoring Functions [50], [87]

While *plausibilization* (AP1) requires establishing what is considered a meaningful effect size in our context, *data acquisition* (AP2), *data analysis* (AP3), *scenario classification* (AP4) and *instantiation* (AP5) as well as *refinement* (AP6) and *abstraction* (AP7) rely on thresholds to determine if a situation or scenario is considered critical. For example, Junietz *et al.* [52] consider manually labeled critical scenarios in order to learn binary classifiers. Our goal is to pick up and extend existing approaches for the applications within the criticality analysis. Based on these results, new metrics that are well-suited for the criticality analysis can be developed manually or learned artificially. This issue of calibration of criticality metrics, however, remains as a challenge for future research. In future work, we plan on investigating those issue using nanoscopic traffic simulation software [43].

d: CRITICALITY METRICS FOR THE RUNNING EXAMPLE

In order to assess the plausibility of the exemplary phenomenon CP_{occ}, we simulated a logical scenario on an urban intersection, as described in subsubsection V-B2. As to identify whether criticality does increase when an occlusion is present in the simulation, we employed a combined a predictive, scaled variant of the Post-Encroachment-Time (PET), based on [2], which we label *Scaled Predictive-Encroachment-Time* (SPrET). Predictive approaches of the PET have been proposed in earlier works [57], [61], but have to our knowledge not been presented in a formalized way. The SPrET is additionally combined with a variant of the required acceleration a_{req} [49] adapted for intersection scenarios rather than car following scenarios.

First, we use a constant velocity model to assess whether the projected paths $\text{tr}_i(s,t) = v_i(t)s + p_i(t)$ of two traffic participants with positions $p_i(t)$ and velocities $v_i(t)$, $i = 1, 2$, do intersect at time t by solving $\text{tr}_1(s_1, t) = \text{tr}_2(s_2, t)$ for s_1 and s_2. The SPrET at time t is then defined as

$$\text{SPrET}(t) = \begin{cases} (s_1 + s_2) \cdot |s_1 - s_2|, \\ \quad \text{if } \exists\, s_1, s_2 > 0 : \text{tr}_1(s_1, t) = \text{tr}_2(s_2, t), \\ \infty, \ \text{otherwise.} \end{cases} \quad (1)$$

Extending the definition to scenarios, the SPrET of a time series $T = (t_0, \ldots, t_n)$ is defined as the minimum, i.e. $\text{SPrET}(T) = \min_{t \in T} \text{SPrET}(t)$.

By definition, the SPrET predicts the Post-Encroachment-Time (i.e. $|s_1 - s_2|$) under the assumption that all traffic participants will follow a straight path defined by their current velocity. Intuitively, this answers the question: how close will the encounter of both traffic participants be if no one reacts to the other? When the SPrET approaches zero, a change of behavior should be considered by at least by of the actors. Moreover, we scale the predicted value by the factor $(s_1 + s_2)$ to avoid over-estimating the criticality when the vehicles are far away from the predicted intersection point. This allows us to effectively model prediction uncertainty, which we assume increases with the distance to the predicted intersection.

The SPrET addresses the temporal dimension of criticality, i.e. the time gap between actors passing an intersection point, but not the dynamical dimension. For this, we evaluate the required deceleration for each actor once the SPrET falls below a certain threshold. As to assess this, let $p_{\text{int}} = \text{tr}_1(s_1, t) = \text{tr}_2(s_2, t)$ denote the predicted intersection point. For each time t, we define the *Conditional Required Acceleration* $a_{\text{req,cond}}$ in m/s^2 as the deceleration that is required for actor $i = 1, 2$ to stop in front of p_{int}:

$$a_{\text{req,cond}}(t, i) = \begin{cases} || \dfrac{v_i(t)^2}{2(p_i(t) - p_{\text{int}})} ||, \ \text{if } \text{SPrET}(t) < 3s^2, \\ 0, \ \text{otherwise.} \end{cases} \quad (2)$$

The threshold value $3s^2$ has been chosen by the author's human judgment of criticality. Analogously to SPrET, $a_{\text{req,cond}}$ of a time series T is defined as $a_{\text{req,cond}}(T, i) = \max_{t \in T} a_{\text{req,cond}}(t, i)$.

The combination of both metrics is a useful criticality surrogate in urban intersection scenarios, as it assesses two important dimensions of criticality. Additionally, the prediction does not rely on intersecting trajectories, i.e. predicted collisions, over time in order deliver a meaningful value. This avoids problems that arise for intersection scenarios as is the case with a simple TTC metric: when assuming straight paths under constant velocity, if no collision will happen, the TTC will be infinite even if the projected paths do intersect.

Figure 11 gives two example graphs for SPrET and $a_{\text{req,cond}}$ measurements. In the first of these concrete instances of the running example of Figure 16, the ego vehicle approaches the intersection while the oncoming bicyclist is critically occluded by parking vehicles. The bicyclist becomes visible to the *ego* only shortly before entering the intersection area. The ego vehicle cannot react sufficiently, as the $a_{\text{req,cond}}$ spikes upwards at the end of the scenario. In the second instance, the bicyclist is not occluded and the ego vehicle is

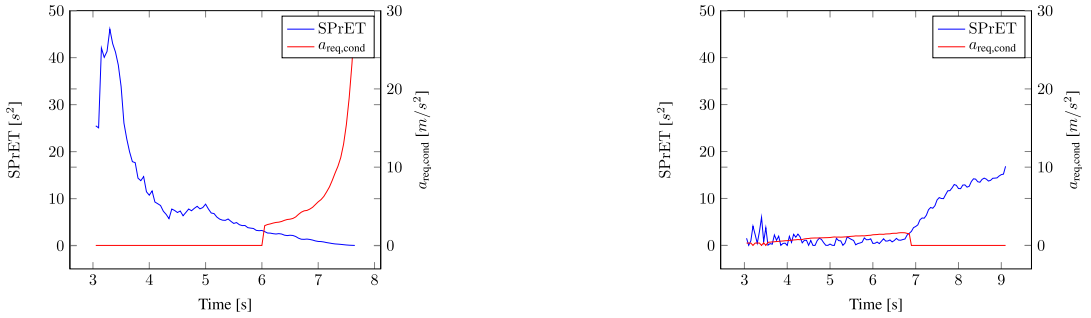


FIGURE 11. SPrET and $a_{\text{req,cond}}$ over time for a critical occlusion (left) and an uncritical non-occlusion (right) scenario.

able to initiate an early reaction at around $t = 7s$. This is indicated by a subsequent increase in SPrET.

7) DATA ANALYSIS

Data analysis is the main connection between the method and information branch and creates new knowledge by means of analyzing parts of the available data basis. Here, criticality metrics present an important tool for the analysis of scenario data. Details on the data basis can be found in subsection V-B.

As indicated in Figure 3, we use techniques from the fields of *statistical analysis* and *machine learning*.

In order to derive valid statements from the results of real world drives or simulation runs, methods from *statistical analysis* can be applied. Descriptive statistics is helpful for expressing or summarizing certain features of a data set. In particular, accident databases can be analyzed using database functionality such as filtering for certain criticality phenomena. However, to infer properties of a data set, e.g. for the plausibilization of causal relations, we have to employ methods from inferential statistics such as the Kolmogorov-Smirnov test, Student's t-test, analysis of variance and χ^2*-test*. Moreover, we can conduct correlation analyses between phenomena and criticality metrics, e.g. using Spearman's ρ. In this way, we can measure the sensitivity of a criticality metric to a parameter and use the result for feature selection when training regression models or classifiers.

On the side of *supervised machine learning*, we can use data sets of criticality metrics evaluated on the parameter space, e.g. of a logical scenario, to fit a regression model. This model can then be used to predict the value of the criticality metric for unseen parameter combinations and lends itself as objective function for optimization techniques. If thresholds for criticality metrics are introduced, we can use such data to fit classifiers for criticality. In case the experiment has been designed according to a plausible causal graph, one can use the regression model to estimate causal effects [38] – under the assumption that confounding variables have been identified and controlled for. Finally, we can apply *unsupervised machine learning* to different data sets, e.g. for clustering of scenarios or anomaly detection in real drive data.

Optimally, the results from machine learning are interpreted and infused with expert knowledge and combined with the results from other statistical analyses. The strength of machine learning techniques, namely finding correlations and fitting models using vast amounts of data, is enabled by a growing data basis. However, in order to infer causality from correlations it is mandatory to apply statistics diligently.

For our running example, we analyze the data gathered by simulation, as described in subsubsection V-B2. We evaluated the $n = 1000$ instantiations for SPrET and $a_{\text{req,cond}}(ego)$. For SPrET, separating the data set into 'no occlusion' (530 scenarios) and 'occlusion' (470 scenarios), we obtained means and standard deviations of $3.25\ (\pm 9.46)s^2$ and $2.75\ (\pm 8.74)s^2$, indicating a 15% decrease in SPrET and therefore, a small increase in criticality. In order to verify the effect of 'occlusion' we checked for differences between the distributions using a two-sample Kolmogorov-Smirnov test ($D = 0.12, p < 0.05$).

For the analysis of $a_{\text{req,cond}}(ego)$, in order to limit the impact of highly-dynamical accident scenarios, we introduced $9.81m/s^2$ as an upper bound for $a_{\text{req,cond}}(ego)$, a value over-approximating the point of no return. Note that the 'no occlusion' scenarios were not affected by this modification. Based on this, we obtained means and standard deviations of $1.10\ (\pm 0.75)m/s^2$ and $3.15\ (\pm 3.10)m/s^2$, i.e., on average in the modified sample, $a_{\text{req,cond}}(ego)$ was around 2.9 times higher for occlusion scenarios, indicating a very strong increase in criticality. Again, we verified the effect using a two-sample Kolmogorov-Smirnov test ($D = 0.40, p < 10^{-35}$) and calculated Cohen's $d = 0.93$. Conducting a correlation analysis between variables and $a_{\text{req,cond}}(ego)$ using Spearman's ρ revealed several interesting significant ($\alpha = 10^{-9}$) correlations, as can be seen in Table 2. In particular, the exposure variable 'occlusion' was significantly correlated with the outcome variable $a_{\text{req,cond}}(ego)$ ($\rho = -0.35, p < 10^{-29}$), which further substantiates the alleged causal relation.

This initial data analysis point towards a statistically significant effect of the criticality phenomenon 'occlusion' on criticality, at least for the running example. In particular, the effect of 'occlusion' on the temporal dimension, measured with SPrET, is rather small. However, its effect on the dynamical dimension, measured with $a_{\text{req,cond}}(ego)$, vastly increases criticality and leads to accident scenarios. More experiments

TABLE 2. Significant ($\alpha = 10^{-9}$) results of correlation analysis between variables and $a_{\text{req,cond}}(ego)$ using Spearman's ρ.

Variable	Correlation (ρ)	p-value
Occlusion	0.29	$p < 10^{-20}$
Duration of occlusion	0.26	$p < 10^{-15}$
ego starting position (x)	-0.24	$p < 10^{-14}$
bicyclist starting position (y)	-0.35	$p < 10^{-29}$
bicyclist target speed	0.42	$p < 10^{-44}$
Position of O (y)	0.20	$p < 10^{-9}$

are required to confirm this effect and, eventually, to infer a causal link between occlusion and criticality.

8) CATALOGIZATION

After examining the relation of a phenomenon to a criticality property φ, the following artifacts have been produced:

- CP, σ s.t. $CP \models_\sigma \varphi$, the phenomenon associated to φ
- CR, ρ s.t. $CR \vdash_\rho CP$, the relation explaining CP
- H, the hypothesis for CR
- E, the evidences supporting H
- a description of the abstract scenario class $Sc(CP)$
- a description of the abstract scenario class $Sc(CR)$
- a set of scenario instantiations for these classes
- concrete simulation and measurement data

These artifacts will be cataloged to enable re-use and traceability during and after the analysis. The following sections consider, among other aspects, the storage and access of such a catalog of method artifacts and scenarios.

B. INFORMATION BRANCH

The information branches supplies the method branch with relevant information. For the analysis, we differentiate between two categories of information: data and knowledge. Knowledge, in this context, is understood to be general facts about the world, whereas data represents concrete instances of those facts. We can thus extract knowledge from data by means of analyzing the relations of instances in the data – an inherently inductive process. Additionally, knowledge is generated from existing knowledge itself through deduction. The access to such information provides the foundation to identifying and substantiating relevant causal relations.

The handling of information is depicted in three steps: Firstly, we lay the foundations of our information branch approach by introducing *ontologies*, a key enabler in formalizing and reasoning over the available information.

Subsequently, we demonstrate the connection of the method to the *real world*: it is addressed how relevant knowledge about the traffic domain can be formalized, stored and utilized as well as how data and knowledge about the real world are incorporated into the method branch.

Lastly, we depict the generation, storage and usage of the *method-related information*, i.e. the artifacts produced during the criticality analysis. For this, meta-models for the most common entities that occur during the analysis, e.g. criticality phenomena, are introduced. Such models need to be stored and conveniently accessed – both for machines and humans – thus enabling consistent knowledge re-use and even automated deduction throughout the process, especially considering the iterative nature of the criticality analysis.

1) REPRESENTING INFORMATION THROUGH ONTOLOGIES

a: INTRODUCING ONTOLOGIES

Explicating information for storage, access, and reasoning purposes is a central concept of the field of information science. In this context, an ontology is a digital and formal tool for such an explication process [42]. In a nutshell, for a given domain, an ontology stores a representation of classified concepts and entities as well as their properties and relations to each other. Thus, it is a formal model of a given conceptualization. Specifically, knowledge can be represented in terms of universal axioms over the structured entities, which in turn can be instantiated by measured or synthetic data. Ontologies are capable of handling both knowledge (i.e. general facts) and data (i.e. instances of those facts).

b: THEORETICAL AND TECHNICAL LIMITATIONS OF ONTOLOGIES

Although ontologies can express a vast amount of knowledge, there exist both theoretical and technical limitations to their expressiveness. Firstly, a common realization of ontologies is the Web Ontology Language (OWL) 2, which is semantically based on the description logic (DL) $\mathcal{SROIQ}$. DLs are a decidable fragment of first-order logic. Hence, in exchange for computability, it is not possible to express all possibly relevant facts. Besides the theoretical limitations, there exist practical considerations: often, describing temporal or spatial properties on a concrete level can become a tedious task when using plain description logics. In practice, one can fall back to abstract predicates, e.g. *to the side of*.

c: ONTOLOGIES IN THE CRITICALITY ANALYSIS

We identify two relevant ontological domains:

1) *The automotive urban traffic domain*, including concepts such as vehicles, infrastructure, and weather.
2) *The criticality analysis domain*, including artifacts from subsection V-A, e.g. criticality phenomena.

We will further substantiate ontologies for both domains in their respective upcoming paragraphs.

2) INFORMATION ABOUT THE REAL WORLD

In section IV, we bootstrapped the criticality analysis by means of a model of the real world. For reasons of traceability, reliability and comparability, we strongly suggest making this model explicit. We propose an ontological basis.

a: ONTOLOGICAL INFORMATION REPRESENTATION

We make use of a traffic ontology that allows the analysts to argue over a standardized technical language for the urban traffic domain, which we denote by the *Automotive Urban Traffic Ontology* (A.U.T.O.). The implementation of this

ontology is based on the 6 layer model [72] and subject of future work. Though, at this point, we motivate the need for having such a domain ontology: it enables expert knowledge to be stored digitally such that a sound and sufficiently complete model of the traffic world can be constructed. This model represents the entities relevant for reasoning about criticality. Note that at the beginning, such a model will be filled with entities deemed to be relevant. If new entities and relations are discovered to be in scope for the safety of automated driving, the ontology can be extended iteratively.

Another aspect considers abstractions: entities – e.g. bicyclist and pedestrian – can be combined to abstracted classes – e.g. vulnerable road user. In other words, the abstraction and refinement relation can be ontologically explicated.

b: INFORMATION ACQUISITION

Although ontologies are an essential tool in such an information representation process, the method additionally needs to consider how to obtain information. This concerns the knowledge – e.g. the taxonomy of relevant entities, their relations and abstractions – but also the data, meaning concrete facts about those entities, typically observed in the real world but possibly also deducted from existing knowledge.

c: KNOWLEDGE ACQUISITION

Existing knowledge can be harvested from various sources. For the workings of road traffic, those include:

- Guidelines and laws, e.g. driving school catalogs
- Traffic research analyses conducted by domain experts [7], [67], [77], [81], [73], [78]
- Surveys and opinions of domain experts
- Subjective experience reports

One can also employ knowledge on autonomous vehicles:

- AV disengagement analyses [11], [25], [33], [83]
- Autonomous vehicle accident analyses [34], [86], [88]

d: DATA ACQUISITION

Besides knowledge, data needs to be gathered as well to support the hypotheses proposed in the method branch. Existing data sources include:

- Data from automated or manual drives equipped with automation-grade sensors [1], [15] [37], [39] [40], [74]
- Naturalistic driving studies, e.g. [9], [28] [55]
- Accident data bases [13], [64]

For the exemplary criticality phenomenon, occlusion, we gathered evidential support by examining an accident data base. For this, we coordinated an analysis of the representative GIDAS data base through the traffic accident research at the TU Dresden.[8] We restricted the population to inner-city accidents involving at least one passenger car from 2007 to 2019, leading to a total number of 12997 cases. Occlusion was present in 2978 accidents, as depicted in Table 3. This implies that occlusion occurred – but may not be causal–in over a fifth of all accidents in Germany in the given time frame. An analysis of the identified concretizations shows that the occlusion of traffic participants constitutes the largest share. Some concretizations were not identifiable. In the case of occluded obstacle, there exist no relevant accidents in the data base. Note that due to the overlap – some accidents contain multiple concretizations – the sum of the cases for the concretizations exceeds the total occlusion number.

[8] www.vufo.de

TABLE 3. Absolute and relative cases for the representative GIDAS data set of inner-city accidents with passenger car involvement. Includes the projected total number of inner-city accidents with passenger car involvement in Germany. Occluded lane markings and traffic lights were not identified.

Criticality Phenomenon	Absolute Cases	Relative Cases	Projection
Occlusion	2978	22.9%	36746
Occluded Pedestrian	600	4.6%	7401
Occluded Bicyclist	1076	8.3%	13280
Occluded Intersecting Vehicle	844	6.5%	10413
Occluded Obstacle	0	0%	0
Occluded Lane Markings	-	-	-
Occluded Traffic Sign	313	2.4%	3865
Occluded Traffic Light	-	-	-

This data can subsequently be used for the assessment of relevancy of criticality phenomena in the method branch. We conclude that occlusion is relevant, as it is associated with a large part of all accidents. On the other hand, occluded obstacles were not relevant for *human* inner-city accidents.

Due to the fact that existing information on autonomous vehicles acting in open traffic environments is still sparse, additional data needs to be generated when conducting a criticality analysis. For example, we identified occlusion to be both relevant for machines and humans, although there exists subtle differences. Humans are often affected by visual obstructions of their line of sight, a phenomenon whose effects can be seen in data bases such as GIDAS. For AVs, objects can be occluded by other phenomena as well, e.g. a fine spray mist that is almost invisible to the human eye.

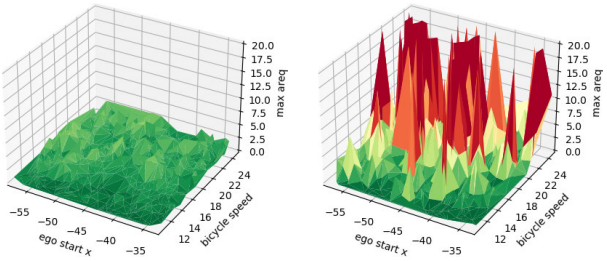
Depending on the required quality, validity, level of detail, sample size and cost factors, one can address this issue using:

- Synthetic data from nano-, micro- and macroscopic traffic simulations, where simulated scenarios can be derived from real world drives or be synthetically defined.
- Measurement data from vehicle drives on proving grounds and in open-context environments.
- Partially synthetic data of X-in-the-loop methods, filling the continuum of fully synthetic and real drive data.

For this, nanoscopic traffic simulation software such as openPASS [26] or CARLA [27], possibly in combination with microscopic traffic simulations like Eclipse SUMO [58], may be used. However, in order to generate meaningful evidences, this approach relies heavily on the validity of the simulation environment, the utilized models and their interactions [10], [63]. Therefore, if the plausibilization of a causal relation is based solely on simulative data, this part of the

TABLE 4. Parameters corresponding to confounding variables used in the logical scenario for the exemplary simulation. All parameters were drawn from a uniform distribution.

Parameter	Range
ego start position (x, y)	$[-58, -33] \times [-29, -28]$
ego target position (x, y)	$[50, 55] \times [-29, -28]$
ego target speed (km/h)	$[25, 60]$
bicyclist start position (x, y)	$[31, 32] \times [3, 15]$
bicyclist target position (x, y)	$[-50, -45] \times [-34, -33]$
bicyclist target speed (km/h)	$[10, 25]$
Dimension of O (discretized as number of parking cars)	$\{0, 1, 2, 3, 4, 5, 6, 7\}$
Position of O (x, y)	$[2, 20] \times ([-35, -34] \cup [-26, -25])$



Group A: no occlusion present Group B: occlusion present

FIGURE 12. Visualization of the simulation results for the running example. Group A shows scenarios without occlusion and Group B shows samples with an occlusion present.

hypothesis needs to be continuously validated with real world data. The continuous validation of such hypotheses can be considered a part of an AVs life cycle. For example, suppose that repeated simulative experiments delivered satisfactory evidence supporting $CR_{stat-occ-tp}$. For the validation of this hypothesis with real world data, an AV needs either be able to recognize $CR_{stat-occ-tp}$ online or record the scenario for subsequent analysis.

Finally, during the collection process, the question whether enough data with enough diversity was collected in order to deliver confidence in the hypotheses will arise. Similar questions are currently being examined [23], [45] [84].

e: INFORMATION ACQUISITION FOR THE RUNNING EXAMPLE

In order to further substantiate the running example with evidences, we obtained data[9] for the plausibilization of occlusion using the CARLA Simulator [27]. We defined a logical scenario for the running example depicted in Figure 16. The ranges of the parameters are stated in Table 4.

Note that the variables were chosen exemplarily. When conducting such an experiment on a large scale, *all* confounding variables have to be controlled for in order to generate statistically relevant evidence. For example, we did not consider weather as a confounding variable, although it most likely affects both occlusion as well as criticality, e.g. in scenarios where rain reduces the visibility as well as the friction between road and tires. Additionally, valid probability distributions and density functions shall be chosen.

For the simulation, we employed simplistic models. The bicyclist was set to a non-reactive model with constant speed, whereas the *ego* vehicle implements a basic driving function. Its perception relies on an array of several obstacle sensors pointing to the vehicle's front with a combined field of view of 90° and a viewing distance of 35 meters. The sensors' view can be blocked by visual obstructions such as parking cars. The *ego*'s planner conducts a light braking maneuver or an emergency brake, depending on positions and velocities, if a moving object, as perceived by one of the obstacle sensors, is predicted to cross the intersection within a certain time span.

[9]The data set is published at https://github.com/lu-w/criticality-analysis.

For the experiment, we sampled $n = 1000$ instantiations from the logical scenario (cf. Table 4) and executed them using CARLA. For each simulation run we measured whether an occlusion between *ego* and bicyclist happened or not, as well as the maximum $a_{\text{req,cond}}$.

Figure 12 shows a three-dimensional visualization of the criticality, as measured with $a_{\text{req,cond}}(ego)$, plotted against the two high-impact confounders '*ego* start position' (x) and '*bicycle* speed', which were determined by correlation analysis of the sample data. Here, less critical scenarios correspond to low $a_{\text{req,cond}}(ego)$-values (green) while more critical scenarios correspond to high $a_{\text{req,cond}}(ego)$-values (red). A detailed analysis of the data is presented in subsubsection V-A7.

3) INFORMATION EMERGING DURING THE CRITICALITY ANALYSIS

Besides the traffic domain, it is necessary to store knowledge about the current state of the analysis, e.g. the relation of criticality to the modeled traffic entities. For example, phenomena can be tagged, and identified causal relations between entities can be related to those phenomena.

For *criticality phenomena*, we propose the model schematically depicted in Figure 13 using the VOWL specification.[10] Each phenomenon *CP* can be described through an abstract scenario, i.e. $Sc(CP)$, e.g. the set of all scenarios where there exists an occluding object between two traffic participants. It is extended by its association with criticality and the underlying causal relation *CR* leading to the phenomenon *CP*. Additionally, tags can relate a phenomenon to other phenomena and traffic entities to support the analyst.

Moreover, *evidences* play a central role in the analysis. Firstly, there needs to be evidence for the relevance of the phenomena, answering: Is the given phenomenon even associated with criticality? In case of a positive association, we are then interested in finding evidences for the hypothesized causal relations. Finally, after proposing the underlying causal relation, evidences supporting their plausibility need to be gathered. These evidences can be from different sources. For instance, the influence of the friction coefficient on the Brake-Threat-Number can easily be supported by an

[10]http://purl.org/vowl/spec/v2

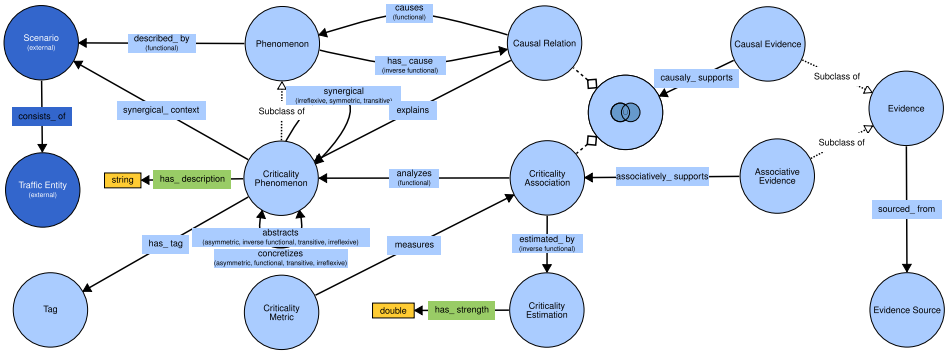


FIGURE 13. Ontological model of the entity 'criticality phenomenon' and its relations, visualized in VOWL.

analytical argument. More complex relations, e.g. that the number of crossing arms is causal for an increase in criticality, will be accompanied by empirical evidence, e.g. through simulation. Moreover, by using an ontological basis, we can use inference to deductively generate logical evidences.

Lastly, *causal relations* can also be represented ontologically.[11] The traffic ontology specifies the referenced classes, e.g. *physical objects*, and their relations, e.g. *before*. In the ontology, the identified causal relations are additionally annotated by their evidences to increase traceability.

C. SCENARIO BRANCH

This section covers the scenario branch of the criticality analysis, as shown in Figure 4. For various process steps of the method we require an adequate description of the interesting happenings within traffic. The point of view of a scenario is different from a criticality phenomenon or a causal relation. The scenario is focused on describing the happenings, or even just the scenery and the actions of the involved actors. Clearly, scenarios are a key element to scenario-based verification and validation (cf. [63], [32]), where a vehicle under test is confronted with interesting or challenging situations in order to assess its performance. The criticality analysis uses scenario classification to complement the abstraction and refinement process described in subsection V-A. There are many options how classification of scenarios can be realized, e.g. according to criticality phenomena, causal relations or clustering based on expected behavior. Naturally, scenarios are also employed for the specification of real-world drives or simulations used for data acquisition.

Before we elaborate on the requirements on the scenario description resulting from the proposed workflow, we briefly introduce already existing terminology from previous works such as the PEGASUS project. Based on this, we outline the methodical aspects and requirements on the scenario description and introduce the concept of *abstract scenarios*. In particular, we emphasize the tight interconnection between the scenario description and the developed traffic ontology described in subsubsection V-B2. The ontology enables a consistent use of names for entities as well as the efficient clustering of entities into classes such as 'vulnerable road user'. These ontological classes play an important role for the specification of the scenario classes as well.

1) SCENARIO DESCRIPTION

Scenarios are used to describe the evolution of traffic scenes over time. Therefore, they are employed to talk about interesting things we experience driving or participating in traffic, to communicate a challenging setup, to talk about what contributed to a challenging situation or as an instruction for recreation in a simulation, in proving grounds or for observation in live traffic flow. There exists a variety of terminologies on scenarios in the literature. We base our discussion on the terminology used within the PEGASUS project, i.e. the terms *scene, situation* and *scenario*, as introduced by Ulbrich *et al.*, cf. [79], as well as the three different qualifications of the term scenario – functional, logical and concrete – established by Menzel *et al.* [60]. There are already numerous standardization activities in this area, e.g. ISO/PAS 21448 [48], which covers the safety of the intended functionality (SOTIF) and therefore extends the well-known functional safety norm ISO 26262 [47], or the development of ASAM OpenSCENARIO [35]. Our terminology is intended to support a development process to fulfill the requirements imposed by the standards especially for vehicles in urban traffic.

In order to describe the evolution of traffic happenings, we consider snapshots of the environment, called scenes. A *scene* contains the scenery, the static and dynamic traffic elements, the self-representation of all actors and observers, as well as relations between those entities. In contrast to a scene, a scenario describes a time span. A *scenario* is a description of the evolution of a sequence of scenes over time, starting with an initial scene. In order to characterize the temporal development of the scenario, we use actions and events as well as goals and values of the involved actors.

Within PEGASUS, three different types of scenario descriptions were established by Menzel *et al.* [60]: the

[11] At the end of a criticality analysis, there will be a large number of propositions on phenomena connected by a web of causal relations. One can use graph databases to store this network of causal relations.

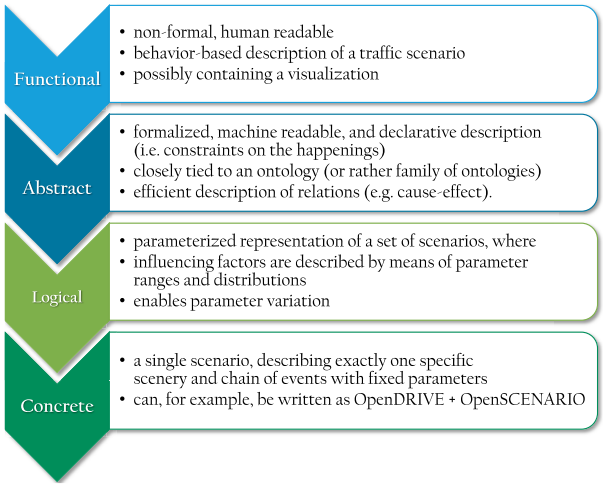


FIGURE 14. An overview of the different qualifications of scenario description.

functional scenario, the *logical scenario* and the *concrete scenario*. We use the following notion of these different qualifications of scenarios, compare Figure 14 for an overview.

A *functional scenario* is a behavior-based description of a traffic scenario. It uses a controlled natural language and may employ the terminology of an ontology as a basis. Optionally, it may include a visualization. We use this form of description as a first step to sketch out a scenario, which is later on used as an easily readable representation of what is supposed to happen in an informal, abstract way. Although the original concept had already foreseen a close connection with an ontology and the formal depth that goes with it, see e.g. [60], today's understanding is generally reduced to the natural linguistic aspect. In order to avoid misunderstandings, we restrain our definition of functional scenario to this notion.

The next more concrete qualification is the logical scenario. While initially defined in [60], we use the definition of the DIN SAE SPEC 91381 [17]. We point out a slight difference in the use within the criticality analysis, caused by the specifications focus on testing, whereas we consider the analysis of phenomena associated with critical situations.

A *logical scenario* is a model of the time sequence of scenes whose parameters are defined as ranges, starting at a specific point in time. However, we omit the constraint that the behavior of the main actor is not further specified.

Within the criticality analysis, on top of the actions of the surrounding traffic we usually have the behavior of the main actor also specified. Hence, we deviate from the standard in that regard. The criticality analysis looks at the traffic system itself as the system of interest, when identifying causal relations. Thus, we do not necessarily have a single vehicle of interest (or even system under test) within a scenario. Even in the cases where the focus is put on a single vehicle, we still have to specify its actions. This enables us to express the actions of *ego* that contributed to the temporal evolution and the resulting change in criticality.

Finally, a *concrete scenario* is an instance of a logical scenario where all parameters are evaluated to a single value.

a: REQUIREMENTS FROM POINT OF VIEW OF THE CRITICALITY ANALYSIS

The method branch of the criticality analysis, cf. V-A, is centered around identifying criticality phenomena and analyzing causal relations related to criticality in the context in which the *ego* will be operating in. Within this workflow, scenarios play various distinguished roles. In particular, they

- describe the happenings in traffic in a way that is comprehensible and can be interpreted for humans,
- enable the creation of a knowledge base, structured as a scenario database, and
- are also the currency that connects the workflow to simulation tools and real-world drives.

The method branch (cf. subsection V-A) of the criticality analysis is based on finding hypotheses for causal relations which contribute to the criticality of a traffic situation. Simulation and real-world drives are used to generate evidences for the hypotheses under consideration. For the plausibilization of such hypotheses, the need for specific additional evidences may arise. Thus, we derive scenarios related to the proposed causal relation in order to analyze the correctness or the validity boundaries of the relation.

The knowledge-driven part of the data acquisition step, described in the information branch (cf. subsection V-B), uses scenarios as a means for the experts to talk about sequences of scenes and situations in traffic or sequences of phenomena leading to a critical situation.

Causal relations cannot be easily represented in a logical scenario, since they are generally a more abstract concept than a parameterized model. Functional scenarios, on the other hand, having the complete expressiveness of natural language, can describe cause-and-effect relations. However, to reach the necessary degree of automation and thus scaleability in the workflow, it is necessary for the description to both human- and machine-readable, resulting in the need for a formal and declarative scenario specification level.

b: ABSTRACT SCENARIOS

For the presented reasons, we introduce the *abstract scenario*, situated in between the functional and logical scenario with respect to the abstraction level, as depicted in Figure 14.

Definition 13 (Abstract Scenario): An abstract scenario is a formalized, declarative description of a traffic scenario focusing on complex relations, particularly on causal relations.

The semantic of the description will be closely tied to an ontology, drastically increasing the precision of the employed terminology. The declarative specification on the abstract level enables the focusing of the relevant aspects of the scenario while sparing the details which are unimportant for the causal relation for the later refinement steps.

A number of of specification schemes are currently analyzed and extended regarding their use within the criticality

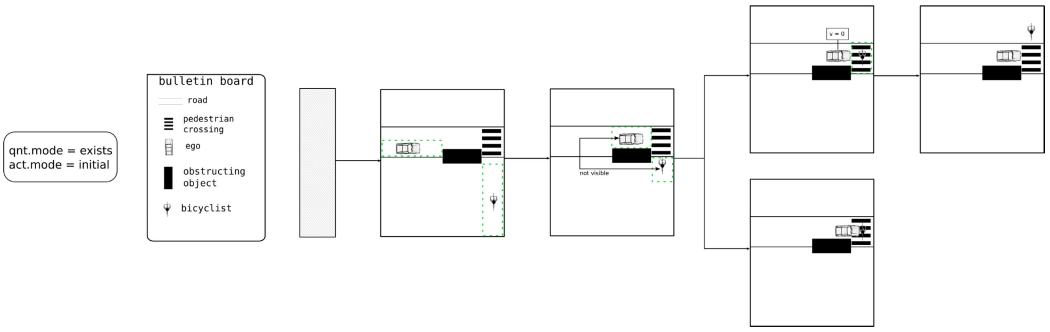


FIGURE 15. A Traffic Sequence Chart for an abstract scenario in which an object occludes a bicyclist at a pedestrian crossing.

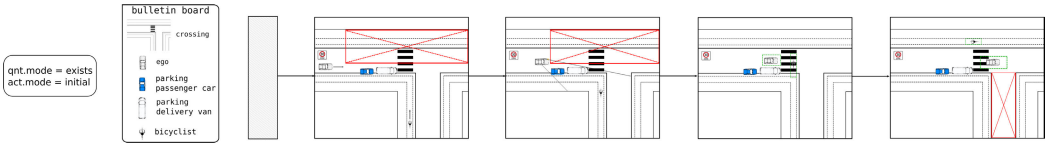


FIGURE 16. A TSC specifying a subset of the abstract scenario of Figure 15.

analysis. Promising candidates as building blocks for this task are Traffic Sequence Charts (TSCs) [21] as well as modeling with zone graphs in the SCODE for Open Context Analysis (SOCA) approach [14]. In the following, we briefly introduce these specification schemes under consideration.

Figure 15 and Figure 16 illustrate how an abstract scenario involving the running example of occlusion can be specified as a Traffic Sequence Chart. The abstract scenario of Figure 15, 'obstructing object occludes a bicyclist at a pedestrian crossing', depicts a sequence of traffic constraints from left to right, including a branching of the scenario into two possible evolutions. In the upper evolution, the *ego* brakes in front of the pedestrian crossing and the bicyclist crosses the street safely, while in the lower evolution, the *ego* collides with the bicyclist. From this abstract scenario space, we can easily specify a subset of scenarios, e.g. by choosing a single path through the TSC and concretizing the 'obstructing object' to 'parking deliveriy van', as can be seen in Figure 16.

The SOCA method introduced by Rittel *et al.* [14] is based on an abstraction from concrete road geometries and concrete population with objects as well as other traffic participants called *zone graphs*. This concept is used together with the SCODE method for the description and analysis of traffic situations. SCODE is a morphological analysis used at BOSCH, which is based on the essential analysis described in [59] as well as the morphological analysis of F. Zwicky [90].

Figure 17 shows the zone graph for the running example on a sketch of a T-intersection, displaying the different types of zones that have to be considered in that situation. The zones the *ego* will be passing along its intended path, marked by the red arrows, are called *driving zones* (Y, F, G and H in this example). The *position zones* K and L denote the areas where other vehicles can get in conflict with the driving path. The zones M and N together with their sub-zones Q and P are the

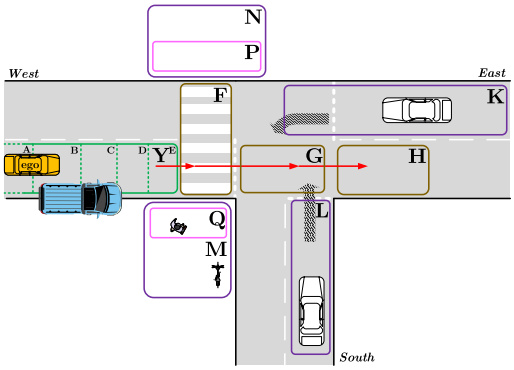


FIGURE 17. A zone graph for the abstract scenario 'obstructing object occludes a bicyclist at a pedestrian crossing'.

areas of interest with respect to pedestrians using the crosswalk F. The sub-zones Y.A to Y.E are *dynamic driving subzones* that capture the different braking distances for yielding at the yield line. Then, E can be considered as maximum braking not sufficient, D as emergency braking possible, C as uncomfortable braking possible, B as comfortable braking possible, and A as no braking necessary.

Based on this graph, a behavior specification can be derived using a Zwicky-Box (morphological box). Each row of the box represents one of the decision dimensions used for the specification of the behavior of the *ego*. Each dimension consists of a set of (exclusive) valid alternatives. shows a The Zwicky-Box for the running example, as shown in Table 5, uses the position of the *ego* together with the allowed behavior for each of the driving zones for the specification. Note that driving zone Y is not present, because it is the initial zone of the situation and thus a stop in front of the zone would be meaningless. The definition of the behavior is given by a consistent partitioning of the space spanned by

TABLE 5. The Zwicky-Box used for the specification of the behavior modes on the zone graph of Figure 17. Alternatives marked in yellow define one rule of the mode 'comfortable stop'. Dimensions without a selection are interpreted as 'don't care' and either alternative is contained in the rule.

AllowedBehavior-4-F	pass				stop in front of			
AllowedBehavior-4-G	pass				stop in front of			
AllowedBehavior-4-H	pass				stop in front of			
Position	H	G	F	Y.E	Y.D	Y.C	Y.B	Y.A

these dimensions and can be specified by using rules (logical expressions) on the dimensions. These rules are clustered into modes depending on their allowed or expected behavior.

Thus, the resulting models contain the information under which circumstances the vehicle will have to perform certain maneuvers, e.g. a comfortable stop or an emergency brake. Within the criticality analysis, this information can be to exploited to link the scenarios space that is spanned by such a specification to criticality phenomena or even to causal relations. However, exploring this link and its explanatory strength in more detail are subject to ongoing research.

2) SCENARIO CLASSIFICATION

In order to structure the infinite space of possible scenario space, we aim at deriving scenario classes. Of course, there exist myriad possibilities for scenario classification. However, within the criticality analysis, scenario classification must be closely tied to the abstraction and refinement step (cf. subsubsection V-A5), as described in the method branch.

The core idea is to establish a relation between criticality phenomena, their causal relations and scenario classes to obtain (i) a definition of done and (ii) an input on where to look for new phenomena if we are not yet done. With regard to the high level goals of section III, establishing this relation contributes directly to structuring the open context (G3) with an adequate level of abstraction (G5) while aiming for convergence towards a manageable set of classes (G6).

As to achieve that target, we require an argument for completeness of the covered scenario space, i.e. whether the union of all scenario classes covers the complete scenario space. Furthermore, we need an idea about the completeness of the complete scenario space itself, i.e. about the coverage of (the relevant) aspects of the world by the scenario space. Note that an absolute completeness with respect to the real world is likely not achievable, due to the hard problem of formalizing reality. However, the coverage of a closed formalized scenario space spanning the operational domain of a vehicle seems comparatively achievable.

As already indicated by Figure 2, criticality phenomena can be used as a classifier, i.e. they separate the scenario space into scenarios where the phenomenon plays a role and those where it does not. The same holds true for causal relations. Since a causal relation often involves multiple phenomena, studying combinations of phenomena with respect to the scenarios in which they are present can also be understood as an attribute or tag to a scenario class. Combinatorial testing methods use an heuristic approach to compute a set of scenarios, as to cover pairwise or, more generally, n-wise combinations of the basic phenomena [5]. This concept has already been applied successfully to analyze the domain model for deep neural networks in the context of perception, identifying relevant influencing factors in [41]. Combinatorial testing can be used in the abstraction and refinement step subsubsection V-A5 of the method branch in order to identify causal relations based on combinations of criticality phenomena. Another classification criteria is the understanding of criticality as a combination of complexity and dynamics, as proposed by Damm and Galbas [20].

In the area of virtual testing, using scenario classes to discern situations where small deviations of the environment lead to significant changes in the behavior is well established. This results in scenario classes containing the same expected behavior for a vehicle of interest. This concept can also be employed in the criticality analysis as the capabilities of a vehicle regarding its abstract actions, such as maneuvers, are rather limited. Basically, the atomic actions used to compose maneuvers are steer left or right, accelerate or decelerate as well as give signals, e.g. indicator lights or sounds. A maneuver then can be understood as a short sequence of activities of a vehicle [24]. A vehicle may execute maneuvers such as 'change lane', 'turn right on intersection' and 'keep distance to lead vehicle'. However, the absolute number of maneuvers can be assumed to be reasonably limited. The problematic part with respect to the development of automated driving functions is the vast space of possibilities that could cause an AV to perform one of these maneuvers. For example, a 'lane change' could either be caused by the desire to turn at the next intersection, as part of overtaking another traffic participant, in order to enter or leave a highway, or a variety of other reasons. Thus, the expected behavior of a vehicle of interest is a key aspect of scenario classification.

For the running example of 'occlusion', the associated scenario class, as defined in 4, consists of the set of scenarios where an occlusion is present, written as $Sc(CP_{occ}) = \{S \in SC \mid CP_{occ} \text{ present in } S\}$. A representative of that class is depicted by Figure 18.

3) SCENARIO INSTANTIATION

The derivation of single concrete instances from a scenario class can be understood as a two step process; the methodically more recent step being the sampling of logical scenarios from an abstract scenario. Abstract scenarios are used to describe scenario classes. We want to point out that an abstract scenario is only required to be a declarative description, leading to a non-trivial problem during concretization. The selection of representative logical scenarios as well as the estimation of the achieved coverage is similar to the instantiation problem within a logical scenario. However, the issue here is even harder, due to the fact that the abstract scenario is just spanned by constraints specified in the declarative description versus the parameter space of a logical scenario.

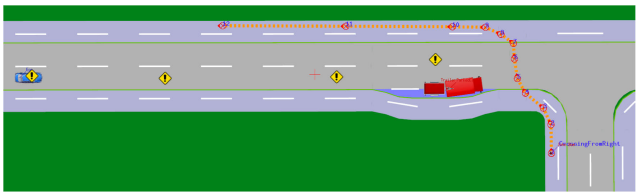


FIGURE 18. Graphical representation of a concrete instance of the scenario 'static occlusion of a bicyclist at a pedestrian crossing'. The underlying OpenSCENARIO file specifies the path of both participants as well as their initial conditions.

The second instantiation step, deriving concrete scenarios from a logical scenario, has been researched e.g. in the PEGASUS project. However, sampling the infinite continuous space of a logical scenario remains subject to current research [63]. The optimal sampling strategy for this space depends on the question that shall be answered. For the purpose of checking a causal relation, one can setup a cost function for an optimizer in a suitable way to aim for a fast falsification of the hypothesis. This will be an issue for future research.

Nevertheless, the classical approaches for instantiating a logical scenario, like random sampling, grid-based sampling, distribution-based evaluation [85], optimization or search-based approaches, can of course be applied. In the end, the goal is the identification of sufficient evidence to support or discard a hypothesis for a causal relation. The question which of these instantiation approaches render the best results to achieve the efficient creation of supporting evidence, while paying attention to the fact that we have to avoid introducing a bias into the criticality analysis here, is to be answered in further development. To be able to study this issue some kind of coverage argument will be necessary, especially for the consideration of sufficiency as well as avoiding the bias.

In section III, we introduced six high level goals, (G1) to (G6). The instantiation described here has to contribute to the goal (G4), i.e. finding sufficiently *representative instances*. In both steps, from abstract to logical as well as from logical to concrete, the selection of the suitable representatives is a key issue. At this point, it is important that the classes and spaces spanned by the logical scenarios are partitioned suitably such that a representative instance can even be defined. This depends also on the question that is posed to the partition. There might be different classifications necessary for different use cases of the scenario catalog, requiring different instantiations. For example if the focus of an analysis lies on perception, different parameters of the logical scenario will be of interest as opposed to the analysis of a planner decision.

Picking up our running example, let us consider the previously discussed T-intersection with a crosswalk. Figure 18 shows a concrete scenario instance at a crossing with a parked truck causing the occlusion in front of the crosswalk.

A logical scenario can be spanned by various parameters for the variation of the concrete scenario. These can include the starting positions of the ego vehicle, the position and velocity of the bicyclist, and the position of the parked truck.

4) SCENARIO CATALOGIZATION

The third step in the scenario branch consists of the catalogization of the emerging scenarios, as it is important to save and document scenarios that appeared in the criticality analysis. We propose to create this catalog of the created scenarios to document the relations between the different levels of abstraction – abstract, logical and concrete scenarios. It is necessary to track the criticality phenomena or causal relations within the scenarios to allow for short iteration cycles when changes arise. A key issue is tracking the association of the criticality phenomena and causal relations to scenario classes, see Definition 4 in subsection V-A. This can be either in the traffic system as a whole, e.g. a new traffic participant emerges (e.g. an e-scooter), or it can simply be new knowledge to be brought into the analysis, e.g. as result of an accident. The scenario catalog helps in finding other interesting scenarios when looking within the data accumulated in the information branch. A further application of the catalog is simply the documentation of the connection between the data basis used within the information branch and the scenario representation. This is of use when checking if a scenario has already been included in the analysis as well as rerunning the data acquisition in a test drive or simulation.

Such a scenario catalog is important for a long-term argumentation about coverage. To enable its identification, we link to the data analysis in the method branch. A scenario catalog is also useful during the design and testing phase of an autonomous system as well as for its release (homologation).

VI. CONCLUSION

The method proposed in this work, called *criticality analysis*, provides a blueprint for structuring the open context with the goal of safe operation of automated vehicles at SAE Level 4 and 5 under manageable effort. The criticality analysis has been divided up into three branches – method-, information and scenario branch – each consisting of several process steps that have been sketched in this work. The method enables analysts to identify in which cases criticality arises in the traffic system and then provide the underlying explanations in order to establish safety principles and mitigation mechanisms for those cases. Future research within the VVM project will address the process steps of the proposed method in-depth and provide a proof-of-concept by exemplarily conducting a criticality analysis for the use case 'urban intersection'.

This entails various activities in the three branches. In the method branch, we focus on providing a comprehensive set of relevant criticality phenomena and associated causal relations. Based on these causal relations, safety principles and mitigation mechanisms for the homologation of automated vehicles can be derived. On the side of data analysis we will employ criticality metrics for the evaluation of scenarios from real-world data as well as synthetic data with respect to the requirements and applications within the criticality analysis. The information branch has to convert the results of aforementioned data analyses into available knowledge. In particular, providing details on the architecture, implementation

and applications of the automotive urban traffic ontology. Further work on the scenario branch includes developing a concept for scenario classification, analysis of different scenario instantiation processes and tackling the problem of representativeness, eventually leading to a comprehensive but manageable abstract scenario catalog.

REFERENCES

[1] S. Agarwal, A. Vora, G. Pandey, W. Williams, H. Kourous, and J. McBride, "Ford multi-AV seasonal dataset," *Int. J. Robot. Res.*, vol. 39, no. 12, pp. 1367–1376, Oct. 2020.

[2] B. L. Allen, B. T. Shin, and P. J. Cooper, "Analysis of traffic conflicts and collisions," *Transp. Res. Rec.*, vol. 667, pp. 67–74, 1978.

[3] D. Althoff, J. J. Kuffner, D. Wollherr, and M. Buss, "Safety assessment of robot trajectories for navigation in uncertain and dynamic environments," *Auton. Robots*, vol. 32, no. 3, pp. 285–302, 2012.

[4] M. Althoff, O. Stursberg, and M. Buss, "Model-based probabilistic collision detection in autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 2, pp. 299–310, Jun. 2009.

[5] 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, no. 1, pp. S52–S57, 2019.

[6] C. Baier and J.-P. Katoen, *Principles of Model Checking*. Cambridge, MA, USA: MIT Press, 2008.

[7] J. S. Baker and H. L. Ross, "Concepts and classification of traffic accident causes," *Int. Road Saf. Traffic Rev.*, vol. 9, no. 31, pp. 11–18, 1961.

[8] J. Becker, S. Kammel, O. Pink, and M. Fausten, "Bosch's approach toward automated driving," *At-Automatisierungstechnik*, vol. 63, no. 3, pp. 180–190, Jan. 2015.

[9] J. Bock, R. Krajewski, T. Moers, S. Runde, L. Vater, and L. Eckstein, "The inD dataset: A drone dataset of naturalistic road user trajectories at German intersections," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Oct. 2020, pp. 1929–1934.

[10] E. Böde, M. Büker, U. Eberle, M. Fränzle, S. Gerwinn, and B. Kramer, "Efficient splitting of test and simulation cases for the verification of highly automated driving functions," in *Proc. Int. Conf. Comput. Saf., Rel., Secur.* Cham, Switzerland: Springer, 2018, pp. 139–153.

[11] A. M. Boggs, R. Arvin, and A. J. Khattak, "Exploring the who, what, when, where, and why of automated vehicle disengagements," *Accident Anal. Prevention*, vol. 136, Mar. Art. no. 2020, Art. no. 105406.

[12] A. Broadhurst, S. Baker, and T. Kanade, "Monte Carlo road safety reasoning," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2005, pp. 319–324.

[13] E. Brühning and S. Berns, "IRTAD-international road traffic and accident database (stand: 1993)," *Zeitschrift für Verkehrssicherheit*, vol. 39, no. 3, pp. 121–124, 1993.

[14] M. Butz, C. Heinzemann, M. Herrmann, J. Oehlerking, M. Rittel, N. Schalm, and D. Ziegenbein, "SOCA: Domain analysis for highly automated driving systems," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, 2020, pp. 1–6, doi: 10.1109/ITSC45102.2020.9294438.

[15] H. Caesar, V. Bankiti, H. A. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "Nuscenes: A multimodal dataset for autonomous driving," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Aug. 2020, pp. 11618–11628.

[16] E. Clarke, O. Grumberg, S. Jha, Y. Lu, and H. Veith, "Counterexample-guided abstraction refinement," in *Proc. Int. Conf. Comput. Aided Verification*. Berlin, Germany: Springer, 2000, pp. 154–169.

[17] *Terms and Definitions Related to Testing of Automated Vehicle Technologies*, Standard DIN SAE Consortium. Spec 91381, 2019, p. 20.

[18] G. Cumming, *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. Evanston, IL, USA: Routledge, 2013.

[19] W. Damm, M. Fränzle, S. Gerwinn, and P. Kröger, "Perspectives on the validation and verification of machine learning systems in the context of highly automated vehicles," in *Proc. AAAI Spring Symposia*, 2018, pp. 512–515.

[20] W. Damm and R. Galbas, "Exploiting learning and scenario-based specification languages for the verification and validation of highly automated driving," in *Proc. 1st Int. Workshop Softw. Eng. AI Auton. Syst.* New York, NY, USA: ACM, 2018, pp. 39–46.

[21] W. Damm, E. Möhlmann, and A. Rakow, "A scenario discovery process based on traffic sequence charts," in *Validation and Verification of Automated Systems*. Cham, Switzerland: Springer, 2020, pp. 61–73.

[22] W. Damm, E. Möhlmann, T. Peikenkamp, and A. Rakow, "A formal semantics for traffic sequence charts," in *Principles of Modeling* (Lecture Notes in Computer Science: Lecture Notes in Artificial Intelligence Lecture Notes Bioinformatics), vol. 10760. Cham, Switzerland: Springer, 2018, pp. 182–205.

[23] E. D. Gelder, J.-P. Paardekooper, O. O. D. Camp, and B. D. Schutter, "Safety assessment of automated vehicles: how to determine whether we have collected enough field data?" *Traffic Injury Prevention*, vol. 20, no. 1, pp. S162–S170, 2019.

[24] E. D. Dickmanns, *Dynamic Vision for Perception and Control of Motion*. New York, NY, USA: Springer-Verlag, 2007.

[25] V. V. Dixit, S. Chand, and D. J. Nair, "Autonomous vehicles: Disengagements, accidents and reaction times," *PLoS ONE*, vol. 11, no. 12, Dec. 2016, Art. no. e0168054.

[26] J. Dobberstein, J. Bakker, L. Wang, T. Vogt, M. Düring, L. Stark, J. Gainey, A. Prahl, R. Mueller, and G. Blondelle, "The eclipse working group openpass-an open source approach to safety impact assessment via simulation," in *Proc. 25th ESV Conf.*, 2017, pp. 1–8.

[27] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," 2017, *arXiv:1711.03938*. [Online]. Available: http://arxiv.org/abs/1711.03938

[28] R. Eenink, Y. Barnard, M. Baumann, X. Augros, and F. Utesch, "UDRIVE: The European naturalistic driving study," in *Proc. Transport Res. Arena*, 2014, pp. 1–10.

[29] J. Eggert, "Predictive risk estimation for intelligent ADAS functions," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 711–718.

[30] A. Eidehall, "Multi-target threat assessment for automotive applications," in *Proc. 14th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2011, pp. 433–438.

[31] A. Eidehall and L. Petersson, "Statistical threat assessment for general road scenes using Monte Carlo sampling," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 1, pp. 137–147, Mar. 2008.

[32] ENABLE-S3 Consortium, *Testing and Validation of Highly Automated Systems*, A. Leitner, D. Watzenig, and J. Ibanez-Guzman, Eds. Cham, Switzerland: Springer, May 2019. [Online]. Available: https://link.springer.com/book/10.1007%2F978-3-030-14628-3

[33] F. Favarò, S. Eurich, and N. Nader, "Autonomous vehicles' disengagements: Trends, triggers, and regulatory limitations," *Accident Anal. Prevention*, vol. 110, pp. 136–148, Jan. 2018.

[34] M. F. Favarò, N. Nader, O. Sky Eurich, M. Tripp, and N. Varadaraju, "Examining accident reports involving autonomous vehicles in California," *PLoS ONE*, vol. 12, no. 9, 2017, Art. no. e0184952.

[35] *ASAM OpenSCENARIO*, Assoc. Standardization Automat. Measuring Syst. (ASAM), Munich, Germany, 2020.

[36] T. W. Forbes, "Human factors in highway design, operation and safety problems," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 2, no. 1, pp. 1–8, Feb. 1960.

[37] L. Fridman, E. Daniel Brown, M. Glazer, W. Angell, S. Dodd, B. Jenik, J. Terwilliger, A. Patsekin, J. Kindelsberger, L. Ding, S. Seaman, A. Mehler, A. Sipperley, A. Pettinato, B. Seppelt, L. Angell, B. Mehler, and B. Reimer, "MIT advanced vehicle technology study: Large-scale naturalistic driving study of driver behavior and interaction with automation," *IEEE Access*, vol. 7, pp. 102021–102038, 2019.

[38] M. J. Funk, D. Westreich, C. Wiesen, T. Stürmer, M. A. Brookhart, and M. Davidian, "Doubly robust estimation of causal effects," *Amer. J. Epidemiol.*, vol. 173, no. 7, pp. 761–767, Apr. 2011.

[39] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231–1237, Sep. 2013.

[40] J. Geyer, Y. Kassahun, M. Mahmudi, X. Ricou, R. Durgesh, S. Andrew Chung, L. Hauswald, V. H. Pham, M. Mühlegg, and S. Dorn, "A2D2: Audi autonomous driving dataset," 2020, *arXiv:2004.06320v1*. [Online]. Available: https://arxiv.org/abs/2004.06320

[41] C. Gladisch, C. Heinzemann, M. Herrmann, and M. Woehrle, "Leveraging combinatorial testing for safety-critical computer vision datasets," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 324–325.

[42] N. Guarino and P. Giaretta, "Ontologies and knowledge bases," in *Towards Very Large Knowledge Bases*. Amsterdam, The Netherlands: IOS Press, 1995, pp. 1–2.

[43] S. Hallerbach, Y. Xia, U. Eberle, and F. Koester, "Simulation-based identification of critical scenarios for cooperative and automated vehicles," *SAE Int. J. Connected Automated Vehicles*, vol. 1, no. 2, pp. 93–106, Apr. 2018.

[44] H. Hansson and B. Jonsson, "A logic for reasoning about time and reliability," *Formal Aspects Comput.*, vol. 6, no. 5, pp. 512–535, Sep. 1994.
[45] L. Hartjen, R. Philipp, F. Schuldt, and B. Friedrich, "Saturation effects in recorded maneuver data for the test of automated driving," *Highway Res. Rec.*, 1972.
[46] J. C. Hayward, "Near miss determination through use of a scale of danger," Dept. Pennsylvania Transp. Traffic Saf. Center, Pennsylvania State Univ. Park, University Park, PA, USA, Tech. Rep., 1972.
[47] *Road vehicles—Functional Safety*, Standard ISO 26262, 2011.
[48] *Road Vehicles—Safety of the Intended Functionality*, Standard ISO/PAS 21448, 2019.
[49] J. Jansson, "Collision avoidance theory: With application to automotive collision mitigation," Ph.D. dissertation, Linköping Univ. Electron. Press, Linköping, Sweden, 2005.
[50] J. Ji, A. Khajepour, W. W. Melek, and Y. Huang, "Path planning and tracking for vehicle collision avoidance based on model predictive control with multiconstraints," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 952–964, Feb. 2017.
[51] P. Junietz, F. Bonakdar, B. Klamann, and H. Winner, "Criticality metric for the safety validation of automated driving using model predictive trajectory optimization," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 60–65.
[52] P. Junietz, J. Schneider, and H. Winner, "Metrik zur Bewertung der Kritikalität von Verkehrssituationen und-szenarien," in *Proc. 11th Workshop Fahrerassistenzsysteme*, 2017, pp. 1–12.
[53] N. Kalra and S. M. Paddock, "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?" *Transp. Res. Part A, Policy Pract.*, vol. 94, pp. 182–193, Dec. 2016.
[54] R. Karlsson, J. Jansson, and F. Gustafsson, "Model-based statistical tracking and decision making for collision avoidance application," in *Proc. Amer. Control Conf.*, vol. 4, 2004, pp. 3435–3440.
[55] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein, "The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2118–2125.
[56] B. Kramer, C. Neurohr, M. Büker, E. Böde, M. Fränzle, and W. Damm, "Identification and quantification of hazardous scenarios for automated driving," in *Proc. Int. Symp. Model-Based Saf. Assessment*. Cham, Switzerland: Springer, 2020, pp. 163–178.
[57] A. Laureshyn, C. Johnsson, T. D. Ceunynck, A. Svensson, M. D. Goede, N. Saunier, P. L. W. lodarek, R. V. D. Horst, and S. Daniels, "Review of current study methods for VRU safety. Appendix 6—Scoping review: Surrogate measures of safety in site-based road traffic observations: Deliverable 2.1–part 4," Deliverable 2.1 InDev Consortium, Tech. Rep., Oct. 2016.
[58] P. A. Lopez, E. Wiessner, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flotterod, R. Hilbrich, L. Lucken, J. Rummel, and P. Wagner, "Microscopic traffic simulation using SUMO," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2575–2582.
[59] S. M. McMenamin and J. F. Palmer, *Essential System Analysis*. New York, NY, USA: Yourdon Press, 1984.
[60] T. Menzel, G. Bagschik, and M. Maurer, "Scenarios for development, test and validation of automated vehicles," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1821–1827.
[61] M. G. Mohamed and N. Saunier, "Motion prediction methods for surrogate safety analysis," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2386, no. 1, pp. 168–178, Jan. 2013.
[62] E. S. Morales, R. Membarth, A. Gaull, P. Slusallek, T. Dirndorfer, A. Kammenhuber, C. Lauer, and M. Botsch, "Parallel multi-hypothesis algorithm for criticality estimation in traffic and collision avoidance," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 2164–2171.
[63] C. Neurohr, L. Westhofen, T. Henning, T. D. Graaff, E. Möhlmann, and E. Böde, "Fundamental considerations around scenario-based testing for automated driving," 2020, *arXiv:2005.04045*. [Online]. Available: http://arxiv.org/abs/2005.04045
[64] D. Otte, C. Krettek, H. Brunner, and H. Zwipp, "Scientific approach and methodology of a new in-depth investigation study in Germany called Gidas," in *Proc. Int. Tech. Conf. Enhanced Saf. Vehicles*. Washington, DC, USA: National Highway Traffic Safety Administration, 2003, p. 10.
[65] J. Pearl, *Causality*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[66] L. N. Peesapati, M. P. Hunter, and M. O. Rodgers, "Can post encroachment time substitute intersection characteristics in crash prediction models?" *J. Saf. Res.*, vol. 66, pp. 205–211, Sep. 2018.
[67] E. Petridou and M. Moustaki, "Human factors in the causation of road traffic crashes," *Eur. J. Epidemiol.*, vol. 16, no. 9, pp. 819–826, Sep. 2000.
[68] A. Poddey, T. Brade, J. E. Stellet, and W. Branz, "On the validation of complex systems operating in open contexts," 2019, *arXiv:1902.10517v1*. [Online]. Available: https://arxiv.org/abs/1902.10517
[69] M. A. Pourhoseingholi, A. R. Baghestani, and M. Vahedi, "How to control confounding effects by statistical analysis," *Gastroenterol. Hepatol. Bed Bench*, vol. 5, no. 2, p. 79, 2012.
[70] *Definitions for Terms Related to on-Road Motor Vehicle Automated Driving Systems*, Standard J3016, SAE International Standard, Taxonomy SAE, 2014.
[71] T. Schick, *How to Think about Weird Things*. New York, NY, USA: McGraw-Hill, 1995.
[72] M. Scholtes, L. Westhofen, L. R. Turner, K. Lotto, M. Schuldes, H. Weber, N. Wagener, C. Neurohr, M. Bollmann, F. Körtke, J. Hiller, M. Hoss, J. Bock, and L. Eckstein, "6-layer model for a structured description and categorization of urban traffic and environment," 2020, *arXiv:2012.06319*. [Online]. Available: http://arxiv.org/abs/2012.06319
[73] M. C. Simon, T. Hermitte, and Y. Page, "Intersection road accident causation: A European view," in *Proc. 21st Int. Tech. Conf. Enhanced Saf. Vehicles*, 2009, pp. 1–10.
[74] P. Sun *et al.*, "Scalability in perception for autonomous driving: Waymo open dataset," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2446–2454.
[75] A. Tamke, T. Dang, and G. Breuel, "A flexible method for criticality assessment in driver assistance systems," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2011, pp. 697–702.
[76] J. Textor, B. V. D. Zander, M. S. Gilthorpe, M. Liśkiewicz, and G. T. Ellison, "Robust causal inference using directed acyclic graphs: The R package 'dagitty,'" *Int. J. Epidemiol.*, vol. 45, no. 6, pp. 1887–1894, 2016.
[77] J. R. Treat, N. S. Tumbas, S. T. McDonald, D. Shinar, and R. D. Hume, "Tri-level study of the causes of traffic accidents. Executive summary," Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep., May 1979.
[78] N. Uchida, M. Kawakoshi, T. Tagawa, and T. Mochida, "An investigation of factors contributing to major crash types in japan based on naturalistic driving data," *IATSS Res.*, vol. 34, no. 1, pp. 22–30, Jul. 2010.
[79] S. Ulbrich, T. Menzel, A. Reschka, F. Schuldt, and M. Maurer, "Defining and substantiating the terms scene, situation, and scenario for automated driving," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 982–988.
[80] W. Wachenfeld and H. Winner, "The release of autonomous vehicles," in *Autonomous Driving*. Berlin, Germany: Springer, 2016, pp. 425–449.
[81] W. A. Wagenaar and J. T. Reason, "Types and tokens in road accident causation," *Ergonomics*, vol. 33, nos. 10–11, pp. 1365–1375, Oct. 1990.
[82] S. Wagner, K. Groh, T. Kuhbeck, M. Dorfel, and A. Knoll, "Using Time-to-React based on naturalistic traffic object behavior for scenario-based risk assessment of automated driving," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1521–1528.
[83] S. Wang and Z. Li, "Exploring causes and effects of automated vehicle disengagement using statistical modeling and classification tree based on field test data," *Accident Anal. Prevention*, vol. 129, pp. 44–54, Aug. 2019.
[84] W. Wang, C. Liu, and D. Zhao, "How much data is enough? A statistical approach with case study on longitudinal driving behavior," *IEEE Trans. Intell. Veh.*, vol. 2, no. 2, pp. 85–98, Jun. 2017.
[85] N. Weber, D. Frerichs, and U. Eberle, "A simulation-based, statistical approach for the derivation of concrete scenarios for the release of highly automated driving functions," in *Proc. Automot. Electron. 11th GMM-Symp.*, 2020, pp. 1–6.
[86] T. Winkle, "Safety benefits of automated vehicles: Extended findings from accident research for development, validation and testing," in *Autonomous Driving: Technical, Legal and Social Aspects*, M. Maurer, J. C. Gerdes, B. Lenz, and H. Winner, Eds. Berlin, Germany: Springer, 2016, pp. 335–364.
[87] M. T. Wolf and J. W. Burdick, "Artificial potential functions for highway driving with collision avoidance," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2008, pp. 3731–3736.
[88] W. Ye, C. Wang, F. Chen, S. Yan, and L. Li, "Approaching autonomous driving with cautious optimism: Analysis of road traffic injuries involving autonomous vehicles based on field test data," *Injury Prevention*, vol. 27, no. 1, pp. 42–47, Jan. 2021.

[89] L. A. Zadeh, "Fuzzy sets," in *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers*, L. A. Zadeh, Ed. Singapore: World Scientific, 1996, pp. 394–432.
[90] F. Zwicky, "The morphological approach to discovery, invention, research and construction," in *New Methods of Thought and Procedure: Contributions to the Symposium on Methodologies*. Berlin, Germany: Springer, 1967, pp. 273–297.

CHRISTIAN NEUROHR received the B.Sc. and M.Sc. degrees in mathematics from Technische Universität Kaiserslautern, Germany, in 2011 and 2013, respectively, and the Ph.D. degree (Dr. rer. nat.) from Carl von Ossietzky Universität Oldenburg, Germany, in 2018. After a short period as a Visiting Researcher with the MAGMA Group, The University of Sydney, he started his current occupation as a Postdoctoral Researcher at OFFIS e.V., Oldenburg, Germany, where he is working in the area of scenario-based verification and validation of automated vehicles. Specifically, he is the Scientific Coordinator of the sub-project 'Criticality Analysis' within the project VVMethods.

LUKAS WESTHOFEN received the B.Sc. and M.Sc. (Hons.) degrees in computer science from RWTH Aachen University, in 2015 and 2019, respectively, specializing in the topics of probabilistic programs and software verification. He is currently pursuing the Ph.D. degree with OFFIS e.V., Oldenburg, Germany. His general work focuses on developing methods to establish confidence in the safety of automated vehicles. More specifically, his research interests include the formalization of knowledge as well as its exploitation for safeguarding automated driving functions.

MARTIN BUTZ received the Ph.D. degree (Dr. rer. nat) in mathematics from the University of Regensburg, in 2012. Since then, he has been working as a Research Engineer in the field of formal methods and software architecture at corporate research of Robert Bosch GmbH in Renningen, Germany. The main focus of his current research is verification and validation of autonomous driving functions, in particular, the application of formal modeling techniques for open context problems.

MARTIN HERBERT BOLLMANN received the Dipl.-Ing. degree in mechanical engineering from Technical University Dresden, in 2012. Until 2019, he worked as a Calculation and Reliability Engineer at ZF Group in Friedrichshafen, Germany. He currently works as a Test Architect, developing test strategies and searching for test methods that are aligned with SOTIF and able to fulfill requirements coming from safety-related standards. Moreover, he is the industrial coordinator of the sub-project 'Criticality Analysis' within the project VVMethods.

ULRICH EBERLE studied physics at Stuttgart University and received the Ph.D. degree for a thesis conducted at the Max-Planck-Institut for Intelligent Systems. In 2003, he joined GM and Opel. His work at the time focused on technology strategy development on alternative powertrains, energy storage systems, and infrastructure. He was involved in strategic activities of the German National Innovation Program on hydrogen storage, and was a member of the founding team of 'New Energy World,' an EU Joint Technology Initiative. Within this context, he also served on the 'Task Force Elektromobilität' of the German automotive industry association VDA. His current assignment at Opel and Groupe PSA centers around novel advanced simulation-based and hybrid-reality methodologies to develop, test, and validate conventional and AI-based Automated Driving functions. Within VVMethods, he is the industrial coordinator of the sub-project 'VVMethods Framework.'

ROLAND GALBAS received the Dipl.-Ing. degree in electrical engineering, with a focus on control techniques for renewable energies from the University of Kassel, in 1988. Since then, he has been working at Robert Bosch GmbH Schwieberdingen/Abstatt, Germany, in various areas, including active safety system development, vehicle dynamics management and requirement development, and positions such as Developer, Team Lead, Senior Manager, and Project Lead. He was involved in the creation of the German initiative for automated and connected vehicles 'VDA-Leitinitiative' and the Project Lead for Robert Bosch GmbH in the EU projects SCOUT and CARTRE. Since 2019, he has been a Project Consortium Lead for the VVMethods project.