Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving

Simon Ulbrich, Till Menzel, Andreas Reschka, Fabian Schuldt, and Markus Maurer

Abstract—For the design and test of functional modules of an automated vehicle, it is essential to define interfaces. While interfaces on the perception side, like object lists, point clouds or occupancy grids, are to a certain degree settled already, they are quite vague in the consecutive steps of context modeling and in particular on the side of driving execution. The authors consider the scene as the central interface between perception and behavior planning & control. Within the behavior planning & control block, a situation is a central data container. A scenario is a common approach to substantiate test cases for functional modules and can be used to detail the functional description of a system. However, definitions of these terms are often at best vague or even contradictory. This paper will review these definitions and come up with a consistent definition for each term. Moreover, we present an example for the implementation of each of these interfaces.

I. INTRODUCTION

During the last 25 years, the driving abilities of automated vehicles have progressed rapidly. This went along with a huge increase of complexity in terms of the number of involved hardware and software components for automated vehicles. To keep this complexity manageable, many teams use functional system architectures to structure necessary modules into more abstract, functional units. Figure 1 illustrates a functional system architecture for an automated vehicle based on Matthaei & Maurer [1], Reschka et al. [2], and Ulbrich et al. [3]. In this paper, we discuss some key interfaces in such a system architecture.

Many system architectures differentiate between perception modules and modules for navigation, guidance and stabilization tasks (cf. Donges [4]). We summarize them under the term planning & control. A scene is the central link between perception and planning & control modules. A situation is a central interface within planning & control. Thus, what is a scene? What is the difference toward a situation and what is different compared to a scenario?

Defining a consistent terminology is particularly language-specific. This paper defines scene, situation and scenario in English. Moreover, we will define those terms in German in [5].

Sections II, IV and VI review existing definitions and suggest one definition for each term. Sections III, V and VII provide example implementations, while section VIII shows the application of each of the terms to the field of test and simulation. Finally, section IX concludes this contribution.

S. Ulbrich, T. Menzel, A. Reschka, F. Schuldt, and M. Maurer are with the Institute of Control Engineering, Technische Universität Braunschweig, Hans-Sommer-Str. 66, 38106 Braunschweig, Germany {ulbrich, menzel, reschka, schuldt, maurer}@ifr.ing.tu-bs.de
concept does not contradict to include temporal aspects like a time since a previous event (e.g., overtaking a vehicle or being obstructed by a slow front vehicle).

Geyer’s [8] definition suggests to include “optional driving instructions” as a part of the scene. Vice versa, according to Wershofen & Graefe [9] the robot’s goals should be part of the situation. Similarly, Haag [10] and Krüger [11] differentiate between a scene and a situation by the aspect of actions and possible action alternatives. Linked to this, the aspect of a self-representation discussed by Maurer [7], Bergmüller [12] and Reschka et al. [13] is not yet covered. The authors share the opinion, that a scene does not only cover environment aspects, but also the aspect of a self-representation. For automated driving, the authors suggest to make goal-specific driving instructions rather part of the situation, but add the idea of a self-representation to the scene definition. We suggest to understand Geyer’s driving instructions just as information being part of the self-representation and not as goals. Thus, the author will use the term scene in the following way:

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

In this definition, an actor is an element of a scene acting on its own behalf. An observer\(^1\) is a perceiving element within the scene or is observing the scene as a whole. An element might be an actor and observer at the same time. Dynamic elements are elements that are moving, or have the ability to move. The scenery subsumes all geo-spatially stationary elements (cf. section III).

By being based either on observed information or a-priori-information that needs to be associated with observed information, a perceived scene will always be a subjective view of the world. Even if multiple observers share their information, it will not result in an objective representation of the world, but rather the view from multiple subjective observers. Thus, for a scene representation, an actor strives to achieve complete and certain information about the world, but in reality the scene will always be from an/ several observer’s point of view. However, in a simulated world a scene can be complete and uncertainty-free as from an omniscient observer’s point of view.

A scene serves the basic purpose of an interface between environment and self-perception modules on the one side, and application- and mission-specific modules and tasks on the other side. A sequence of scenes is considered here as a scenario and is described in section VI of this article.

---

\(^1\)This is not an observer as in the sense of control engineering.

### III. EXEMPLARY SCENE IMPLEMENTATION

After reviewing definitions of the term scene, this section illustrates the implementation chosen in the Stadtpilot project at TU Braunschweig (cf. Ulbrich et al. [3]). Figure 2 illustrates the components of a scene. A scene consists of the geo-spatially stationary scenery, dynamic elements, and a self-representation of all actors and observers.

![Fig. 2. Example of a (subjective) scene representation of the real world](image)

Deviating from Geyer’s [8] definition of “dynamic elements” being based on the temporal extent of their scene definition, we assume dynamic elements to move (having kinetic energy), or possibly being able to move (having sufficient energy and abilities to move). Past movements (object has stopped at traffic light) are a strong indicator for potential movements in the immediate future. Current perception skills of technical systems are not sufficient to classify stationary elements as dynamic, therefore a statue anchored to the ground may currently not be differentiable from a not moving pedestrian. Hence, a pedestrian may possibly be misclassified as being part of the scenery, or a statue as being part of the dynamic elements.

Similar to Matthaei [14], we consider environment conditions like weather or light to be part of the scenery as they are quasi-stationary for a scene being just a snapshot with an age in terms of milliseconds. Geyer considers the position of traffic lights or variable traffic signs to be part of the scenery, but seems to consider their state being part of the dynamic elements. Once more, based on the snapshot scene definition, we only require the scenery to be geo-spatially stationary, thus a changing speed limit sign or traffic light is still considered being part of the scenery.

The scenery subsumes all geo-spatially stationary aspects of the scene. This entails metric, semantic and topological information about roads and all their components like lanes, lane markings, road surfaces, or the roads’ domain types. Moreover, this subsumes information about conflict areas between lanes as well as information about their interconnections, e.g., at intersections. Apart from the before mentioned environment conditions, the scenery also includes stationary elements like houses, fences, curbs, trees, traffic lights, or traffic signs.

The scene representation is completed by a self-representation containing the current skill levels and general
system skills, as well as the states and attributes of all actors and observers. The skills may be represented in a very basic form like a timeout signal from a sensor system or in a sophisticated form of a skill graph as proposed by Reschka et al. [13]. For observers, the field of view and occlusions are an essential part of its skills. The actors’/observers’ states and attributes entail information about the position relative to the road network, dynamic motion information, and even information from the (vehicle’s) data busses like whether an indicator is currently activated or not.

The scene is completed by information being model-incompliant for dynamic elements or the scenery. This may be unclassifiable, untrackable or unsegmentable measurements or information about object types not being considered at the design time. So far, many implementations simply ignore these information. However, their existence and possibly even partial, imperfect information may be function-relevant from a functional safety point of view.

Figure 3 illustrates an exemplary screenshot of a subjective scene representation for an automated vehicle with elements and their relationships (e.g. between a dynamic element and a lane). Similar context models or world models with semantic relations have been presented by Homeier et al. [15], Ulbrich et al. [3], and Schmidt et al. [16].

IV. DEFINING THE TERM ”SITUATION”

While the usage of the term scene is inconsistent already, the usage of the term situation is often even more undetermined. According to Wershofen & Graefe [9, p. 3] cited by Maurer [7], a situation is the entirety of circumstances, which are to be considered by a robot for its selection of an appropriate behavior pattern in a particular moment.

In psychology, Wirtz [17, p. 1430] defines a situation as the entirety of circumstances which result in a certain behavior of a human. Wirtz uses the term situation for a person plus its psychological setting.

Reichardt [18, p. 35] defines a situation as the union of subsets of the internal and external situation. The internal situation consists of a subset describing the (automated) vehicle’s state and its user input. The external situation consists of the environment information describing the street, obstacles and traffic signs. He limits his situation definition to the “world of discourse” where the automated vehicle is used. He clarifies that this is just a subset of the real world.

According to Haag [10] cited by Pellkofer [19], the difference between a scene and a situation is the aspect of (possible) actions². Krüger [11] also cited by Pellkofer [19] defines a situation as an extended (system) state, in which an element is not only seen as a physical object, but also its actions and action alternatives are considered to estimate the temporal development of a situation.

Pellkofer [19, p. 4] defines a situation to be the sum of all behavior decision relevant aspects. Relevant for the behavior decision making are the current scene, the intentions and actions of all subjects in a scene (including the ego-vehicle), and the abilities of the ego-vehicle, which represent the decision alternatives. In contrast to him, we do not consider abilities as decision alternatives, but as an input to the decision process to derive decision alternatives.

Mock-Hecker [20, p. 4] considers the traffic situation to be an extract of the traffic (world) at a certain point of time. It entails the actions and plans of traffic participants. To him, a situation does not only represent the current state but even its probable future development.

Once more, an important aspect for the technical usage of a situation is its applicability. As it was done for the scene, the authors suggest to consider a situation, similar to Mock-Hecker [20, p. 4], as a snapshot of the entirety of circumstances, which are to be considered by a robot (actor) for its selection of an appropriate behavior pattern in a particular moment. Revisiting the overtaking example from section II, a situation would not last for several seconds an overtaking scenario might take, but rather be once more a snapshot. Again, such a definition avoids the technical challenge of determining what kind of situation it currently is and how long it lasts before the world changes into another situation.

Another challenge for the definition of the term situation arises from the system architecture elements called “situation assessment” or “situation analysis” as a submodule of the guidance block (cf. figure 1). A situation assessment uses a situation as an input and interprets the situation or particular aspects of it. Thus, its results may be considered as an augmentation of the prior situation, which provided further details regarding certain aspects. Reichel [21], [22] and Siedersberger [23] coin the term of “situation aspects” for these.

Eco [24, p. 65] cited by Maurer [7, p. 95] considered the transition from a signal to a meaning³ as the central signification process conducted by humans. According to Maurer [7, p. 95], a situation assessment could be considered as such a signification process in a technical system.

Geyer et al. [8] clarifies that a “situation is defined by the set of criteria, that need to be true to conduct an associated action”. As for their scene definition, the end of a situation is defined by a change of one criterion, that describes the situation. The authors agree with Geyer et al. [8] that, “depending on the action, the same scene can evolve into different situations.” In the illustration of the proposed ontology in Geyer et al. [8, p. 185], the situation seems to entail the scene fully. To the authors, the instructions in Geyer’s scene should be part of the situation and not the scene. Moreover,

²German: “Handlungsaspekt”.
³German: “Übergang vom Signal zum Sinn”.
the situation should result from an information selection and augmentation of the scene information based on the mission-specific or permanent (c.p. [9]) goals and values of the automated vehicle.

Angenendt [25, p. VIII] assumes a situation is more than just a snapshot of the traffic scene with infrastructure and environment representing measures. Above this, a situation contains information regarding the behavior of traffic participants and resulting informal rules of conduct. He uses the concept of a behavior setting, to entail informal rules followed by traffic participants [25, p. 23]. The authors agree that the “behavior setting” is an integral part of the situation. We suggest to subsume it under goal- and value-related information.

Von Benda [26, p. 1] defines a traffic situation as a limited extract of the entire traffic scene. The driver with its point of view experiences such an extract. A traffic situation from a driver’s point of view is to her the environment of the human machine system. She assumes a situation pertains for a certain amount of time until a new situation starts with the interaction with a new event.

Dickmanns [27, p. 448] defines a situation as “the collection of environmental and all other facts that have influence on making proper (if possible 'optimal') behavior decisions in the mission context. This also includes the state within a maneuver being performed [...] and all safety aspects.” For the authors situation definition, we agree with Dickmanns driving function relevance criteria for situation information.

Schmidt [16] differentiates between a true world model, a true situation for an individual observer and a subjective situation from an individual observer’s point of view. While agreeing that a true situation may exist in a perfect simulated world, a real situation representation in a technical system will always be not all-encompassing, uncertain and from a subjective point of view (cf. section II).

While acknowledging that it is hard to find a general definition of the term situation, the authors suggest the following definition:

A situation is the entirety of circumstances, which are to be considered for the selection of an appropriate behavior pattern at a particular point of time. It entails all relevant conditions, options and determinants for behavior. A situation is derived from the scene by an information selection and augmentation process based on transient (e.g. mission-specific) as well as permanent goals and values. Hence, a situation is always subjective by representing an element’s point of view.

A situation consists of several situation aspects to be interpreted or comprehended by situation assessment modules. A situation is input and output of such modules at once.

According to the authors definition of a situation, it can be fully derived from a scene and the system’s goals and values, as illustrated by the Venn diagram in figure 4. There is a wide overlap between a scene and a situation to include, e.g., all relevant parts of the scenery, all relevant dynamic elements, and all relevant aspects of the self-representation. This information selection helps to simplify the situation representation and by this the driving function development and computational complexity. Moreover, the situation is implicitly or explicitly augmented, e.g., by goals and values. For instance, by explicitly labeling the usefulness of roads or lanes to reach the mission goal or implicitly by characterizing a playing child on the side to be more relevant than a flying around plastic bag. The remaining part of the situation, not overlapping with the scene or the goals and values, represents situation aspects evaluated and populated with information by situation assessment modules.

V. EXEMPLARY SITUATION IMPLEMENTATION

The implementation of a situation deviates from a scene by the before mentioned goal- and value-specific information selection and augmentation. According to figure 4, there is a significant overlap between a scene and a situation regarding the types of information. The major difference is that only driving function relevant information are part of the situation according to the system’s goals and values. Figure 5 provides an example for an implementation.

---

4German: "Verkehrsgeschehen"
5Cf. Wershofen & Graefe [9].
A simple example in figure 6 clarifies the difference: An automated vehicle approaches an intersection with a bike riding on an edifically separated bike lane heading in the same direction. If the mission requires the automated vehicle (blue) to pass the intersection straight and the bike has physically no chance to leave its bike lane, it might be irrelevant for the driving function. Thus, the bike would not be part of the situation representation. If the mission requires a right turn and thus a crossing of the bike lane, the same bike is very relevant for the driving function and needs to be part of the situation representation. The scene representation needs to contain the bike at all times as it is independent of the automated vehicle’s goals and values.

Apart from the aspects discussed in the scene representation already, a situation needs to contain function relevant goals and values. These may be transient like the current mission, or driving commands and preferences given by an operator to the automated vehicle. In a partial automation, such driving commands may be commanded maneuvers like lane changes or a changed time gap for longitudinal distance keeping. These goals and values may also be permanent like regulatory or societal constraints. For computational efficiency, it may be worthwhile not to list every paragraph of the road traffic regulations every few milliseconds as a part of the situation representation, but rather aggregated information, like which country’s or which traffic domain’s road traffic regulations shall be applied.

VI. DEFINING THE TERM “SCENARIO”

The term “scenario” is often found in the context of simulation and testing, and in the functional description of driver assistance systems.

According to Jarke et al. [29] there are “three major disciplines that use scenarios - strategic management, human-computer interaction, and software and systems engineering - to deal with description of current and future realities.”

Go & Carroll [30] remark that the usage of scenarios in any field is quite different, but the elements of a scenario are similar. According to [30], “a scenario is a description that contains (1) actors, (2) background information on the actors and assumptions about their environment, (3) goals or objectives, and (4) sequences of actions and events.”

The Oxford dictionary [31] defines a scenario as a “postulated sequence or development of events” or the “written outline of a film, novel, or stage work giving details of the plot and individual scenes”.

Geyer et al. [8] define that “a scenario includes at least one situation within a scene including the scenery and dynamic elements. However, a scenario further includes the ongoing activity of one or both actors. According to the movie and theater metaphor previously introduced, the term scenario can be understood as some kind of storyline - including the expected action of the driver - but does not specify every action in detail.” To the authors, Geyer et al. should consider multiple actors, instead of both, for a general definition.

According to figure 8, a scenario contains scenes, actions & events and goals & values. The authors suggest the following definition:

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

Scenes in a scenario are linked by actions and events. According to figure 7, a scenario is a single path of a temporal sequence of actions & events (edges) and scenes (nodes) out of the tree representing the entirety of all possible future scenarios for a given initial scene. Other than scenes, a scenario spans a certain amount of time. A scenario needs to include at least one (initial) scene and actions & events to fully specify a path in figure 7. However, a scenario may also be specified by a complete set of scenes, while the actions and events just covering the elapse of a specified time.
Using the theater metaphor, a scenario is typically described by several scenes with prescribed actions & events in between. In the real world, any actions & events are to some degree uncertain. Hence, human actors may even slightly adopt their behavior to achieve a prescribed key or closing scene in a theater play. Likewise, a driving scenario may contain certain key scenes, e.g., a narrowly defined crossing sequence at an intersection. At the extreme, a scenario is described by a storyboard describing every minimal detail like in a cartoon movie. Vice versa, a scenario may also only specify that all actors shall start heading for their goals while following a certain set of prescribed actions & events without specifying any future scenes. However, possibly allowed uncertainty in the behaviors may result in an entirely open outcome of such a scenario after a certain amount of time. If no uncertainty is allowed, both forms of description will be a dual way to specify exactly the same scenario.

Depending on what a scenario is used for, it may also be sufficient to specify only situations instead of entire scenes plus goals & values. This may be true for a test setup solemnly designed to test, e.g., a situation assessment as in the situation-based open-loop test described in section VIII.

For simulation and test of an automated vehicle or its modules, test-cases may be specified. Each of them entails a scenario and pass-fail criteria to evaluate it. Furthermore, the functional description of the system (use-case) needs to be defined in the early phases of the system design according to the V-model, e.g., in the ISO 26262 standard development process [32, Part 3]. A use-case entails a description of the functional range and the desired behavior, the specification of system boundaries, and the definition of one or several usage scenarios. While these scenario descriptions might be rough and incomplete in the first phase, they may be detailed to achieve fully testable test-cases in the development process.

VII. EXEMPLARY SCENARIO IMPLEMENTATION

The components of a scenario implementation are illustrated in figure 8. A scenario consists of at least one scene, actions & events and goals & values. An example for a scenario is illustrated by figure 9. It depicts an initial (simplified) scene for a lane change on a two lane highway. The goals and values are illustrated by the checkered flag. As a goal of the automated vehicle (blue), it is assumed to reach the end of the left lane as soon as possible without any collisions. The initial scene is illustrated by the leftmost birds-eye view. As actions and events the activation of the indicators, the changing of a lane and a lane following are visualized by the clapper board between the scenes. In simulation environments like Virtual Test Drive (VTD) tools already exist to specify scenarios.

VIII. APPLICATION IN TEST AND SIMULATION

The definitions can be applied for the purpose of testing, e.g., the behavior planning for automated vehicles. According to the V-model, e.g., in the ISO 26262 standard development process [32, Part 3], it is necessary to test on different levels of abstraction. After use-cases have been specified in the design phase of system development, test-cases need to be specified. Both require the specification of scenarios. Figure 10 illustrates different levels of testing for a driving function.

On a very basic level, unit tests are executed to test the correct functionality of software components, like particular software functions and parts of code. This is depicted by single parts of a jigsaw puzzle of the driving function in figure 10. As a next step, tactical behavior planning modules, e.g., for situation assessment may be tested with all its parameters and settings in an open-loop test by a situation-based testing. In fact, while it is possible to specify entire scenarios with goals & values, actions & events, and entire scenes, it may be sufficient and easier to specify situations only as a middle ground. Situation-based open-loop testing generates driving situations from the test-case description and evaluates the behavior response without feeding this behavior response back into future situations.

If all test cases are passed successfully, scenario-based testing is used to test the behavior planning in its interaction with strategic level modules and stabilization level modules as a whole and in a closed-loop setup. Scenario-based closed-loop testing specifies an entire scenario in a test case. This
includes scenes, actions & events to alter these scenes and goals & values for situation extraction and as an input for the driving function. The control and behavior response from the driving function is used to influence future scenes and by this implicitly future situations, as well. As a last step, testing is completed by real driving tests.

IX. CONCLUSIONS

In this paper, we reviewed definitions for the terms scene, situation and scenario. Based on the requirements for an automated vehicle, the authors came up with a definition for each term. Moreover, we provided hands-on implementations and demonstrated the usage of a scene, situation and scenario for test and simulation.

It remains to challenge those definitions in multi-agent environments and in the context of vehicle-to-vehicle communication and vehicle-to-infrastructure communication. Moreover, the authors would like to stimulate a discussion among several stakeholders to come up with an industry-wide, consistent nomenclature in our discipline.

REFERENCES