Formal Specification of Traffic Scenarios in Scenario-based Testing of Maritime Assistance Systems

Anna Austel, Matthias Steidel, Bernd Westphal

***Abstract*—With the advent of assistance systems that classify into the four degrees of IMO maritime autonomous surface ships comes the challenge how to verify and certify such systems, which, in many cases, rely on AI-components for perception or recommendations. One approach to verification and validation of such systems proposes to define test cases as traffic scenarios, i.e., evolutions over time of traffic situations.**

A necessary prerequisite to apply scenario-based approaches is the correct conduction of specified tests. Especially in physical tests, to ensure correct conduction, the specification has to be communicated clearly to the person conducting the test and the executed test has to be checked against the specification. An appropriate (intuitive and formal) specification of the scenarios to be recreated in testing can provide these clear instructions and an objective criterion to assess correct test conduction.

In this work, we report from a case-study with a graphical scenario specification formalism (as originally defined for the automotive domain and recently extended for maritime scenarios) following the question whether we can formalise relevant situations and scenarios from real test campaigns using this formalism and see how understandable the results are for maritime practitioners not trained in formal methods.

***Index Terms*—Testing, Autonomous Systems, Scenarios, Formal Specifications.**

I. Introduction

In recent years, there is a rising interest in maritime autonomous surface ships (MASS), i.e. ships which, to a varying degree, can operate independent of human interaction. IMO-defined degrees of autonomy range from automated processes and decision support (degree one) to fully autonomous ships (degree four) [1]. Examples for degree one would be ships equipped with situational awareness systems that use a wide range of sensors and usually AI-based techniques for object detection and propose actions to human seafarers.

To ensure safe operation of ships of any degree of autonomy, the employed assistance systems need to be appropriately tested during development and in the context of a certification process. Testing and certification of autonomous vehicles is a well-known, challenging problem in all transportation domains due to the huge input value domain of sensors like video cameras, Lidar, etc. in the open world these systems are supposed to be operated in. The behaviour of AI-based components like perception software frequently employed in autonomous vehicles is not defined by comprehensible rules as in classical programs. Overall, neither traditional test-cases with well-defined input and expected values nor distance-based methods where test conduction consists of the system travelling a predetermined distance, scale to come to trustworthy conclusions [2]–[4].

In the automotive domain, an approach to this problem emerged which is called scenario-based verification and validation [5]. The idea is to structure the test space into typical and corner-case scenarios in which the system under test as well as other traffic participants or environment conditions evolve in a certain way – rather than defining the test goal by a plain number of kilometres to be driven on roads under supervision.

Previous research into scenario-based testing of MASS mostly considers the conduction of a large number of tests in a simulation [6]. In this work we want to address the conduction of physical tests. To get a better understanding of the difficulties particular to this setting, we give a brief description of how physical scenario-based tests are conducted.

As depicted in Figure 1, a test case consists of a specification of how the test is to be conducted and a specification of how the test is to be evaluated (i.e., how to arrive at a verdict).

For test evaluation to provide a meaningful verdict, correct conduction of the test is a necessary prerequisite. In the case of physical tests there are two main difficulties to be addressed: how to clearly communicate the specification to the person conducting the test and how to reliably check whether the test was actually conducted as specified.

In the case of scenario-based testing of autonomous vessels, the test conduction specification is given by a traffic scenario. This specification then serves as instructions for

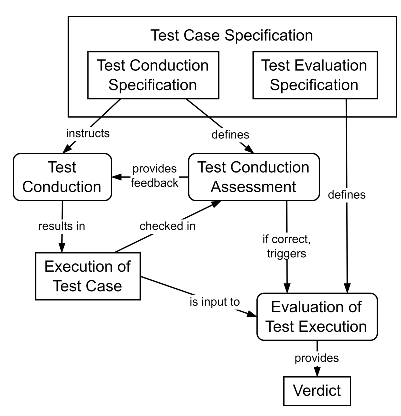

Fig. 1. Steps of testing, from specification over conduction and assessment of conduction to evaluation.

The authors are with DLR SE, Germany.

test conduction. For manual test conduction, e.g in physical tests, the scenario specification has to be given in a format clearly readable by the test conductor. The complexity of the specification and the influence of factors like wind and current may complicate correct test conduction despite clear instructions. It is important to assess the resulting execution of the test case with respect to the specification, since only correct executions are useful for test evaluation. Here the specification should provide an objective criterion to decide correctness. Given such an objective criterion, test conduction is assessed, repeated if necessary, and correct executions of the test case are finally evaluated with regards to the test evaluation specification to decide pass or fail of the test.

In this work, we report on first steps towards scenario-based testing of MASS. In particular we are concerned with supporting correct test conduction.

To provide an objective criterion for test conduction assessment and human-readable instructions for test conduction we see a need for an appropriate, human-readable and formal specification of test conduction. Note that the same principles can be applied to scenario-based experimentation. Here there is no test evaluation specification, as the goal is data collection instead of reaching a verdict. Still, traffic scenarios have to be conducted as specified.

In Section II we present related work on verification and validation of MASS and on formal specification of maritime traffic scenarios. We then present a more detailed description of scenario-based test conduction in Section III. We show that the visual formalism of maritime Traffic Sequence Charts (cf. Section IV below) is capable of capturing maritime traffic scenarios that are relevant in two particular testing tasks in Section V. We conclude and give an outlook on how such formal specifications of maritime situations could give rise to automatic monitoring components that make testing more objective and efficient in Section VI.

II. Related Work

In this section we want to provide some insight into previous work on scenario-based verification and validation of maritime autonomous surface ships (MASS) and formal specification of abstract maritime traffic scenarios.

Existing research regarding scenario-based testing of MASS is mainly concerned with scenario elicitation, i.e. determining which scenarios the system should be tested in, and the executing large numbers of tests in a simulation. For instance, Porres et al. [7] and Bolbot et al. [8] present methods for the choice of scenarios a collision avoidance system should be tested in for verification. Reiher and Hahn [6] analyse the support of scenario-based verification and validation (V&V) by state-of-the-art marine traffic simulators. In particular, they identify the need of a standardised exchange format for maritime traffic scenarios.

There is some previous work on formal specification of maritime traffic scenarios and the application of such formal specifications in the verification of maritime assistance systems. However this research considers the formalisation of scenario-based traffic rules and uses purely formula-based formalisms. Torben et al. [9] discuss the use of formal methods in the verification of MASS. They consider formal specification of scenario-based traffic regulations and requirements to be a useful tool in this context. Krasowski and Althoff [10] formalise scenario-based maritime traffic regulations using metric temporal logic, a formula-based formalism for the specification of system properties involving time. Torben et al. [11] propose a scenario- and simulation-based test method for MASS based on the use of signal temporal logic (another specification formalism for timed properties) for requirement formalisation and as one example formalise rules from the Convention on the International Regulations for Preventing Collisions at Sea (COLREGs).

III. Scenario-based Testing

Maritime autonomous surface ships (MASS) are complex safety-critical systems relying on AI-based components and operating in the complex open context of maritime traffic [12]. Testing and certification of such systems in a purely distance-based manner, i.e. assuming safety of the system after a set distance travelled without incident, is an inefficient approach as a lot of effort will be spent testing the system in common and uncritical circumstances while rare but critical phenomena can easily be missed entirely. This is an especially valid concern in the maritime domain, where large distances may be travelled without other vessels or infrastructure in sight.

A more efficient and effective approach is scenario-based testing. Here a test engineer carefully designs a catalogue of test specifications, each consisting of a test-scenario specifying test conduction, i.e., the conditions for the system to be tested under and requirements the system under test is expected to adhere to under these circumstances. The test-scenarios are chosen to cover all relevant phenomena including typical and critical scenarios (scenario elicitation). These scenarios are then recreated in a virtual, physical or hybrid test environment (test conduction) and the resulting trajectories are evaluated with respect to the requirements (test evaluation) [13].

In the context of testing MASS, test conduction is specified as a maritime traffic scenario, describing an evolution of traffic situations over a period of time. The description may include behaviour and interaction of traffic participants with each other and infrastructure elements as well as environmental conditions. Based on the definitions by Menzel et al. [14], we distinguish between *functional*, *abstract* and *concrete* traffic situations and scenarios. Functional traffic situations and scenarios are natural language descriptions of the general state of and relations between traffic participants and environment at some point of time or the evolution thereof over a period of time respectively. To illustrate this we give examples:

> Situation: *Two vessels meet head-on on the open sea.*
> Scenario: *Two vessels meet head-on on the open sea. Both then alter their course towards starboard and continue on, passing each other on their port sides.*

A concrete traffic situation is a specific (real world) situation at a given point of time and region of space or a complete description of such a situation over a given model of relevant properties. A concrete traffic scenario is an evolution of concrete situations over an interval of time. Abstract traffic situations and scenarios formally describe properties of concrete traffic situations and scenarios respectively, characterizing a (typically infinite) set of concrete traffic situations and scenarios.

In scenario-based testing of autonomous vehicles, test conduction is typically specified in the form of functional scenarios. Correctly conducting a test results in a concrete scenario matching the intentions of the test engineer [13].

As discussed before the goal of this work is to support correct test conduction through an appropriate formal and human-readable specification of test conduction. Formal specification as an abstract traffic scenario provides an objective criterion to assess correctness of test execution. The specification has to be intuitively human-readable to provide clear instructions to the human operator(s) conducting the test.

IV. Scenario Formalisation

We have previously identified the need for an intuitive formal specification of test conduction in the form of abstract maritime traffic scenarios in the context of scenario-based testing and certification of maritime autonomous surface ships. Traffic Sequence Charts (TSCs) [15] are one example of a modelling language tailored towards the intuitive formal specification of traffic scenarios, originally those from the automotive domain. In this section we give a brief introduction to a maritime variant [16] of TSCs (mTSCs), and discuss its application to the formal specification of test conduction.

The TSC formalism is a graphical, spatio-temporal logic for the specification of abstract traffic scenarios.

TSCs specify traffic scenarios in terms of the behaviour and interactions of involved objects such as traffic participants and infrastructure elements in and with their environment. They formalise these properties over an appropriate structural model. The structural model of a TSC is basically a set of classes modelling different types of objects (e.g., ships, quay walls, etc.) with a finite set of typed attributes modelling the relevant properties of the respective objects (e.g., their position, speed over ground, etc.). As their focus is on specifying spatio-temporal properties, TSCs provide dedicated type for positions in 2-dimensional space. Since directions are highly relevant in common maritime traffic scenarios, mTSCs additionally provide a dedicated type for directions. User defined types can be used for further (non-spatial) attributes, e.g., modelling speed, acceleration, or local states of vehicles like enabled light signals. A concrete situation relative to this structural model is then a set of objects with a value for each attribute of the object's class.

TSCs specify abstract scenarios in terms of phases arranged in some temporal and logical structure represented as a Chart. Each phase is characterized by an abstract situation, specified as a so-called *spatial view*. A spatial view then consists of a rectangular canvas on which *object symbols* are placed. Figure 2 for example shows a spatial view with two object symbols. Each object symbol represents an object relevant to the scenario and can be freely chosen to intuitively resemble this object. To support the specification of spatial relations, spatial attributes can be graphically represented as part of an object symbol using dedicated model elements called *position anchors* and *direction anchors*. The red object symbol in the spatial view in Figure 2 represents a ship and has a position anchor representing its position and a direction anchor representing its heading. Spatial views are designed to depict a rough bird's-eye view sketch of the situation they specify. Exact spatial relations are annotated using dedicated model elements called *distance lines*, *direction arrows* and *angle arrows*. Further constraints on attribute values are annotated as predicates. Refer to Figure 2 for examples of all of the aforementioned model elements. The semantics of a spatial view is an object constraint formula which is derived from the diagram through the semantics-defining algorithm (cf. [17]). A concrete situation (a set of objects with values for attributes) then satisfies a spatial view if and only if it satisfies the constraint system under a given binding of objects to symbols.

Given a functional description of a test-scenario, formalisation using mTSCs consists of choosing an appropriate structural model and object symbols, identifying the temporal and logical structure of the scenario, formalising the abstract situation characterizing each phase and designing an appropriate visual representation as a Spatial View.

The primary concern in this formalisation is to make sure to match the test engineers intention for the test-scenario. The method employed here is basically the classical requirements engineering approach to the analysis and formalisation of properties for clients without a background in formal methods [18]: (a) As requirements analyst, conduct interviews with the client to identify relevant structural aspects of the problem domain and collect known constraints and examples (what in software engineering would be use cases or user stories). (b) Develop a formal model that represents the analyst's understanding of the relevant properties. (c) Validate the proposed model with the client by eliciting further examples that are supposed or not supposed to satisfy the formal description of the properties and check the model for whether the computed satisfaction matches the expectation. In case of mismatches, iterate Steps (b) and (c) (or (a), if necessary).

To enable automatically checking correctness of test conduction with respect to a test-scenario formalised as a TSC, data regarding attribute values needs to be (made) available. Where possible, the structural model can be chosen to work with available data. Finally, the visual syntax (including object symbols and spatial views) should be designed to be intuitively readable for the test conductor and provide useful instructions for test conduction.

V. Case Studies

To evaluate whether maritime Traffic Sequence Charts (mTSCs) are an appropriate formalism to support conduction of scenario-based tests and experiments, we conduct two case studies.

In the first case study we formalise the specification of an experimental setup for evaluating performance of an assistance system under development in a specific traffic situation. The relatively simple nature of the specification to be formalised allows us a more detailed discussion of our formalisation procedure. For the second case study we chose a test for a commercially available perception system, where test conduction is specified as a more complex traffic scenario.

A. Case Study 1: Experimental evaluation of object detection system

Consider a ship-mounted object detection system in an early stage of development. The developer wants to experimentally collect data regarding the system's performance in locating a single stationary object. In particular they wish to explore the influence of the positioning of the stationary object relative to the ship on detection accuracy. For this purpose they plan on conducting a set of experiments in which they fix the object to be detected and the approximate velocity of the ship while varying the relative positioning of the ship to the reference object, in particular their distance and relative bearing. One arbitrarily chosen example of the experimental setups is given by the following functional description:

A stationary reference object is located at a distance of a few meters straight ahead or slightly towards the port side of the ship, which is moving at a slow speed over ground.

Note that in this example the goal is experimentation instead of acceptance testing. While compared to scenario-based (acceptance) testing the scenarios to be executed are chosen with a different objective and the evaluation is not with regards to a pass-fail criterion, the actual execution of scenarios still fundamentally works the same way. Given a functional scenario description we have to make sure to produce a concrete scenario in accordance with the intention behind the functional scenario to then be able to use it in an evaluation step. Scenario-based experimentation is thus another application of our method for ensuring correct test conduction.

Furthermore the given functional scenario describes only a single traffic situation instead of a more complex scenario with evolving conditions over time. We chose this as our first example for its simplicity as this allows us a greater focus on the details of the formalisation process.

The above description of a traffic situation is imprecise and prone to misinterpretation. Terms like *a few*, *slightly* and *slow* are imprecise and it is up to subjective interpretation whether the mentioned distance between objects refers to their minimum physical separation, the distance between some unspecified reference positions or some other measure of distance. This illustrates the need for formal specification.

To begin with, we identified two relevant objects, the ship and the reference object. In an interview with the test engineer we identified the reference object's and ship's central positions and the ship's heading as further relevant properties, as their intended meaning for the mentioned distance was the distance between the centres and the term *straight ahead or slightly towards the port side* refers to the angle between the ship's heading and the direction from the ship's position to the reference object's position. We modelled this as elements of a structural model for mTSCs as two object types *ship* and *reference* where *ship* has attributes *sog* (representing a ship's speed over ground in knots), *pos* (representing the central position of a ship as a *position*) and *hdg* (representing a ship's heading as a *direction*) and *reference* has a single attribute *pos* (representing its central position as a *position*).

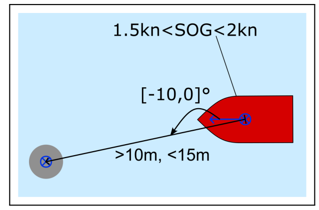

Fig. 2. Spatial view formalising the test-situation from case study 1.

With these attributes available, the intended characteristics of the traffic situation could be refined to:

- The distance between object and ship has to be between 10 and 15 meters, measured from the centre.
- The obstacle has to lie within a relative bearing (angle from ship's heading to direction from ship to object) between -10 and 0 degrees.
- Speed over ground has to be between 1.5 and 2 knots.

While this refined description might provide clear instructions (having defined imprecise terms like *a few* and how to measure distances), they are still neither formal nor intuitive. We can however now express these constraints intuitively and formally using a spatial view as shown in Figure 2.

To construct this spatial view, we first designed appropriate object symbols to represent the ship and the reference object graphically. For the ship we chose a symbol resembling the silhouette of a ship viewed from above as spatial views are meant to resemble a concrete example of the specified traffic situation from a bird's-eye perspective. As the pos attribute is of *position* type we visually represent it using a position anchor. Since the attribute refers to the central position of the ship we placed the anchor in the middle of the symbol. Similarly the hdg attribute is represented by a direction anchor which we placed from the centre of the symbol towards the part resembling the ship's bow. As for the reference object the exact shape or orientation is irrelevant for the situation, we chose a circular symbol as a simple symmetrical shape and place a position anchor for the pos attribute in the centre of the symbol. As the ship is of central relevance to the test conduction we coloured it red to draw attention.

Next we describe the elements of the spatial view with semantical significance and how they relate to the informal description of the situation. The presence of ship and reference symbol in the spatial view, specifies the presence of ship and reference object in the situation. The distance arrow annotated with a distance constraint between the position anchors of the object symbols formalises the constraint on the distance between the objects. Specifically, it contributes the sub-formula $dist(\mathit{ship.pos}, \mathit{obj.pos}) < 15\,m$ to the semantics of the spatial view. For the second constraint we added the direction arrow from the ship symbol's position anchor to the reference symbol's position anchor and the angle arrow with the appropriate constraint from the ship symbol's direction anchor to the direction arrow, resulting in the sub-formula

$angle(ship.hdg, direction(ship.pos, obj.pos)) \in [-10, 0]^\circ$. The final constraint is expressed through the predicate attached to the ship symbol, meaning $1.5\,kn < ship.sog < 2\,kn$.

The semantics of the spatial view only depend on the presence of the elements described above, leaving the design placement and scale of object symbols themselves and within the spatial view up to the person formalising the situation. We discussed intuitive design of the object symbols above. To ensure intuitive readability we further took care in the placement and scale of the object symbols within the spatial view. We place, rotate and scale the object symbols in accordance with the geometry of the specified situation. To do so we collected additional information about the dimensions of the ship and reference object to be used in the experiment from the test engineer and scaled the symbols and the distance between the symbols to approximate the actual proportions. We chose to place the reference symbol at a sufficiently large angle from the ship symbols direction angle to keep the visual elements separate and easy to distinguish while keeping the angle in the graphical representation close to the specified angle.

The resulting spatial view delivers an intuitive visualization of the specified situation in addition to correctly formalising the intended situation. It provides an objective criterion for checking whether the intended situation was recreated correctly. To evaluate this formula for a given real situation (a given ship in a given environment), we only need a component that provides the real values for the attributes in the model (here: positions, heading, speed over ground). If the test ship carrying the object detection system under test has corresponding sensors (built-in or as additional equipment), the formula obtained from the spatial view can be continuously evaluated during the test drive, and objectively assess whether the situations necessary for correct test conduction have been reached or not, and as an added bonus provide continuous live feedback to the test conductor. The spatial view provides clear and more intuitive instructions to a test conductor compared to, e.g., a list of formulas as an alternative representation of the formalised situation.

B. Case Study 2: Testing a vessel recognition system

For the second example, we consider testing of an AI system which is capable of identifying and classifying other vessels in its vicinity based on camera images. The system-under-test is a commercial product that is currently used on ships. Consequently, this is a black box test with the aim of checking the functionality of the system. The system is to be tested in a port environment so that complex encounter situations with other ships can be expected. Initially we were provided with the following description of how the test should be conducted: "The own ship is to pass two target ships on its starboard side. Both target ships are anchored on a line roughly perpendicular to the own ship's course."

The intention here is to start with both vessels in plain view of the system with some free space between the two and having this distance grow smaller over time until one vessel is first partially and finally completely occluded from view of the own ship by the other.

After additional interviews we arrived at a more detailed description of the test conduction.

Two reference vessels are positioned about 25 meters apart and kept close to stationary throughout the test. The own ship approaches the reference vessels on its starboard side at a course approximately perpendicular to the line through the reference vessels, starting at a distance of about 100 meters to the closer of the two vessels and ending circa in line with the reference vessels at a distance of about 25 meters to the closer of the two.

As in the first case study, this functional description is still hard to understand and imprecise.

For formalisation we use an object type *ship* similar to the one from the previous example. Instead of an attribute *hdg* modelling a ship's heading we include an attribute *cog* modelling a ship's course over ground, also as a *direction*. We model the own ship and the reference vessels as instances of this object type and represent them using the object symbol from the first example, with the only difference that the direction anchor now represents the *cog* attribute. We colour the symbol red for the own ship and grey for the reference vessels.

The temporal structure of the scenario is a sequence of three subsequent phases, beginning with the starting conditions immediately followed by a phase of approach and then a phase with the final conditions.

Using the same procedure of refining the specification, constructing the appropriate semantical elements and choosing intuitive arrangements of symbols in the spatial views as before we arrive at the mTSC shown in Figure 3.

Semantically this is satisfied by a concrete scenario on a time interval $[b, e]$ if and only if there are times $b < t_1 < t_2 < e$ such that for all $b \leq t < t_1$ at time t in the concrete scenario the first spatial view is satisfied, for all $t_1 \leq t < t_2$ the second spatial view is satisfied at time t and correspondingly the third spatial view is satisfied at all times t with $t_2 \leq t < e$.

The semantics of the first spatial view as derived using the semantics-defining algorithm for instance is equivalent to the following formula:

$$
\begin{aligned}
& dist(own.pos, r_1.pos) \in [100, 110]m && (1)\\
\wedge\ & dist(r_1.pos, r_2.pos) \in [20, 30]m && (2)\\
\wedge\ & angle(own.cog, dir(r_1.pos, r_2.pos)) \in [80, 100]^\circ && (3)\\
\wedge\ & angle(own.cog, dir(own.pos, r_1.pos)) \in [10, 20]^\circ && (4)\\
\wedge\ & r_1.sog < 0.2kn && (5)\\
\wedge\ & r_2.sog < 0.2kn && (6)
\end{aligned}
$$

Provided data regarding the positions of all three involved ships and the course of the own ship, satisfaction of this formula can be objectively assessed. Similarly satisfaction of the other two spatial views and consequently the whole TSC can be assessed using the same data.

Overall the mTSC correctly specifies the scenario as intended by the test engineer and in addition to an objective criterion to decide correctness of the test execution, provides clearer and more readable instructions compared to the functional description as well as the equivalent formula.

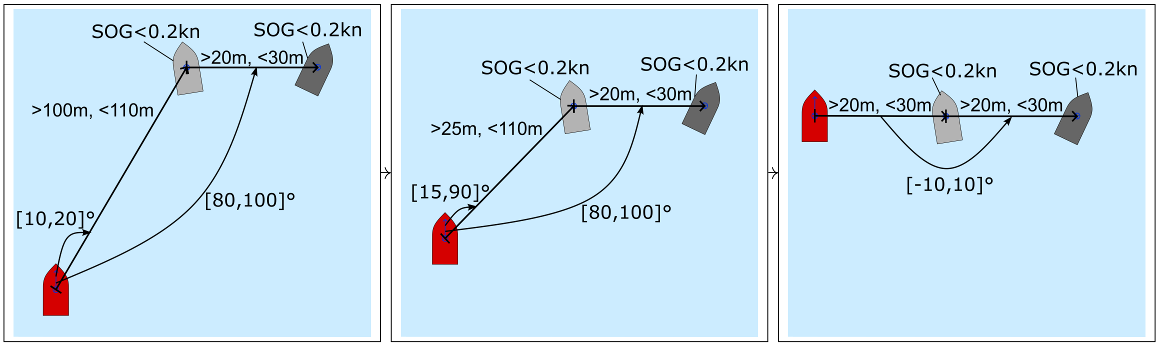

Fig. 3. Maritime Traffic Sequence Chart formalising the test scenario from Case Study 2. Description of how the test should be conducted in three consecutive phases with invariant conditions.

This shows that mTSCs can be used to intuitively specify a range of different test scenarios in a way that can be used as instructions for a test conductor.

VI. Conclusion and Future Work

Scenario-based verification and validation as well as certification of maritime autonomous surface ships (MASS) is a promising alternative to distance-based methods. Here, a test specification consists of a test-scenario describing the conditions the system should be tested in and requirements the system must adhere to under these circumstances. For a test to provide meaningful results the test-scenario first needs to be conducted correctly. Correct test conduction is supported by precise specifications of the relevant traffic scenarios, first of all as clear instructions for the test conductor. A formal specification of test-scenarios additionally provides an objective criterion for checking correct conduction of the tests.

While formalisation of traffic scenario properties has been attempted before, these attempts use notation as mathematical formulas (note that the formula denoted by the spatial view shown in Figure 2 could also have been constructed manually). We propose the use of a visual formalism like Traffic Sequence Charts (TSCs) as an alternative. The main advantage is better readability for domain experts who may not necessarily be experts in formal methods (like developers of such systems, users, certification authorities, etc.). Additionally TSCs are less error-prone in writing since a number of checks can already be applied on the graphics (like correct endpoints of distance arrows).

In an initial Case Study we have shown in two examples, that TSCs can be used to intuitively and formally specify test-scenarios to support correct test conduction. We have seen that specifying test conduction using mTSCs provides useful instructions to people conducting physical tests. As data regarding all attributes used in our formalisations can easily be made available, they also provide an objective criterion for deciding correctness of a test execution, which can be easily determined.

As a next step we plan on integrating online monitors built based on the mTSC specification [19] and installed on the ship conducting a test to provide live feedback during test conduction. This can reduce the effort of physical testing as incorrect test executions can be terminated early.

Furthermore, scenario-based requirements (as for example collision avoidance rules) on the behaviour of a MASS could in future also be specified using TSCs and automatically be evaluated using TSC monitors.

Acknowledgments

The work was conducted within the ‘AMISIA – Advanced Port Maintenance: Intelligent, Sustainable and Automated Dredging’ project, which is funded by the German Federal Ministry for Digital and Transport (BMDV), in the IHATEC programme under grant ID 19H21003D.

References

[1] Maritime Safety Committee, “Outcome of the regulatory scoping exercise for the use of maritime autonomous surface ships (MASS),” International Maritime Organization, Tech. Rep. MSC.1/Circ.1638, 2021, https://www.imo.org.

[2] P. Koopman and M. Wagner, “Challenges in autonomous vehicle testing and validation,” *SAE International Journal of Transportation Safety*, vol. 4, pp. 15–24, 04 2016. [Online]. Available: https://doi.org/10.4271/2016-01-0128

[3] M. Brinkmann, A. Hahn, A. Lamm, and E. Böde, “Learning from automotive: Testing maritime assistance systems up to autonomous vessels,” 06 2017. [Online]. Available: https://doi.org/10.1109/OCEANSE.2017.8084951

[4] N. Kalra and S. M. Paddock, “Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?” *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 182–193, 2016. [Online]. Available: https://doi.org/10.1016/j.tra.2016.09.010

[5] B. Neurohr and E. Möhlmann, “Scenario-based verification and validation of automated transportation systems,” *INSIGHT*, vol. 25, no. 4, pp. 47–50, 2022. [Online]. Available: https://doi.org/10.1002/inst.12411

[6] D. Reiher and A. Hahn, “Review on the current state of scenario- and simulation-based V&V in application for maritime traffic systems,” in *OCEANS 2021: San Diego – Porto*. New York: IEEE, 2021, pp. 1–9. [Online]. Available: https://doi.org/10.23919/OCEANS44145.2021.9705781

[7] I. Porres, S. Azimi, and J. Lilius, “Scenario-based testing of a ship collision avoidance system,” in *2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 2020, pp. 545–552. [Online]. Available: https://doi.org/10.1109/SEAA51224.2020.00090

[8] V. Bolbot, C. Gkerekos, G. Theotokatos, and E. Boulougouris, “Automatic traffic scenarios generation for autonomous ships collision avoidance system testing,” *Ocean Engineering*, vol. 254, p. 111309, 2022. [Online]. Available: https://doi.org/10.1016/j.oceaneng.2022.111309

[9] T. Torben, Ø. Smogeli, I. B. Utne, and A. J. Sørensen, "On formal methods for design and verification of maritime autonomous surface ships," *Proceedings of the World Maritime Technology Conference*, vol. 1, no. 7, pp. 251–261, 2022, accepted: 2023-03-14T13:54:07Z. [Online]. Available: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3058210

[10] H. Krasowski and M. Althoff, "Temporal logic formalization of marine traffic rules," in *2021 IEEE Intelligent Vehicles Symposium (IV)*. Nagoya, Japan: IEEE, 2021, pp. 186–192. [Online]. Available: https://doi.org/10.1109/IV48863.2021.9575685

[11] T. R. Torben, J. A. Glomsrud, T. A. Pedersen, I. B. Utne, and A. J. Sørensen, "Automatic simulation-based testing of autonomous ships using gaussian processes and temporal logic," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 237, no. 2, pp. 293–313, 2023.

[12] R. Abudu and R. Bridgelall, "Autonomous ships: A thematic review," *World*, vol. 5, no. 2, pp. 276–292, 2024. [Online]. Available: https://doi.org/10.3390/world5020015

[13] C. Neurohr, L. Westhofen, T. Henning, T. de Graaff, E. Möhlmann, and E. Böde, "Fundamental considerations around scenario-based testing for automated driving," in *2020 IEEE Intelligent Vehicles Symposium (IV)*. New York: IEEE, Oct. 2020, pp. 121–127. [Online]. Available: https://doi.org/10.1109/IV47402.2020.9304823

[14] T. Menzel, G. Bagschik, and M. Maurer, "Scenarios for development, test and validation of automated vehicles," in *2018 IEEE Intelligent Vehicles Symposium (IV)*. Changshu, China: IEEE, 2018, pp. 1821–1827. [Online]. Available: https://doi.org/10.1109/IVS.2018.8500406

[15] W. Damm, E. Möhlmann, T. Peikenkamp, and A. Rakow, "A formal semantics for Traffic Sequence Charts," in *Principles of Modeling - Essays Dedicated to Edward A. Lee on the Occasion of His 60th Birthday*, ser. Lecture Notes in Computer Science, M. Lohstroh, P. Derler, and M. Sirjani, Eds., vol. 10760. Springer, 2018, pp. 182–205. [Online]. Available: https://doi.org/10.1007/978-3-319-95246-8_11

[16] A. Austel, J. S. Becker, and G. Hake, "A visual formalism for the specification of maritime traffic scenarios," 2024.

[17] W. Damm, S. Kemper, E. Möhlmann, T. Peikenkamp, and A. Rakow, "Traffic Sequence Charts - from visualization to semantics," Reports of SFB/TR 14 AVACS, Number 117, 2017.

[18] C. Rupp, *Requirements-Engineering and Management*, 7th ed. Hanser, München, 2021.

[19] D. Grundt, A. Köhne, I. Saxena, R. Stemmer, B. Westphal, and E. Möhlmann, "Towards runtime monitoring of complex system requirements for autonomous driving functions," in *FMAS*. Ithaca, NY: arXiv.org, 2022, pp. 53–61. [Online]. Available: https://doi.org/10.48550/arXiv.2209.14032