Generating Structured Set of Encounters for Verifying Automated Collision and Grounding Avoidance Systems

Tom Arne Pedersen[1]*, Chanjei Vasanthan[1]*, Kristian Karolius[1], Øystein Engelhardtsen[1], Koen Pieter Houweling[1], Are Jørgensen[1]

[1]DNV, Veritasveien 1, 1363 Høvik

*E-mail: tom.arne.pedersen@dnv.com, chanjei.vasanthan@dnv.com

Abstract. It is a clear consensus in the maritime industry that ships, both conventional and with automated Collision and Grounding Avoidance (CAGA) systems, shall be navigated following the Convention on the International Regulations for Preventing Collisions at Sea (COLREG), ensuring a level playing field and a common understanding of the responsibilities between conventional and autonomous ships, allowing safe co-existence in the waterways. Simulation-based testing is expected to play an important role to explore and document safety and the ability to adhere to the COLREG, as it allows for rigorous testing of performance and failure handling before the software is installed on the ship. Many terms in COLREG used to specify acceptable behavior are qualitative, such as "ample time" and "apparent maneuver" and need to be viewed in light of a specific situation to be applied in practice. As the COLREG includes a large set of obligations specific to different types of situations, a significant number of ship encounters spread out across the different aspects of the regulations will be necessary to enable any judgments about the overall ability to adhere to these rules. The work presented in this paper provides a method and software toolbox for generating ship encounters to allow a systematic assessment of the vessel's ability to act according to COLREG using simulation-based testing, and it is made freely available to vendors, students, or any other actors involved in the development or assurance of algorithms for automated collision and grounding avoidance systems.

1. Introduction

The road towards the realization of fully autonomous ships has indeed been longer than first anticipated. Several concepts and projects have materialized over the years, although most are still relying on the onboard crew to be operated safely. Smaller passenger ferries such as milliAmpere2 [1], MF Estelle in Stockholm [2], and Greenhopper ferry in Aalborg [3] are all relying on a safety operator, either onboard or onshore, while larger container vessels such as Yara Birkeland [4], and the Asko barges MS Marit and MS Therese [5] will have autonomous functionality while being remotely operated and monitored. In addition, decision support systems on board vessels are expanding, to aid the navigation of the vessel. The support spans from detecting and tracking vessels and other objects in the area to co-piloting the vessels. The Orca AI Co-Captain [6] and Sea Machines SM300 [7] are two examples of such systems.

Neither development nor verification of automated collision and grounding avoidance (CAGA) functions for ships are trivial tasks as these systems need to be able to handle a large number of

different situations while adhering to the COLREG, short for the Convention of International Regulations for Preventing Collision at Sea [8]. One could question why autonomous ships shall adhere to COLREG when many navigators in specific areas are not. According to [9], several navigators are maneuvering their ships using informal rules and by this not following COLREG. Still, it is explicitly stated in the COLREG that all ships shall act in accordance with these regulations, and this will also include autonomous ships.

However, a major challenge in the COLREG is that many terms used to specify acceptable behavior are qualitative, such as "ample time" and "apparent maneuver". Hence, they need to be viewed in light of a specific situation to be applied in practice. In addition, the rules in the COLREG describes only a limited set of basic one-to-one encounters and do not provide explicit requirements for more complex traffic situations involving multiple ship encounters. This makes it challenging to develop an adequate algorithm with the ability to follow the COLREG under arbitrary situations.

Hence, sufficient assurance and testing of the CAGA algorithms will be important to build confidence and trust in the system by ensuring safe, optimal, and collision-free navigation [10]. The testing may be carried out in several ways, where both real-life testing using physical ships and simulation-based testing are possible solutions for the assurance process. Real-life testing will often be time-consuming and expensive. Additionally, it does not easily allow re-running already performed tests to ensure software is still behaving as expected, known as regression testing, essential to handle software updates [11]. On other hand, simulation-based testing facilitates extensive and cost-effective testing, including tests that are too dangerous to be performed in real life, while also enabling regression testing. However, real-life testing is still necessary to verify and ensure that the results gathered using simulation-based testing are valid in the real-world [12].

Independent of how the qualification process is performed, it is critical for the confidence in the CAGA system that the ship encounters used to verify the system behavior, are representative and relevant. The ship encounters may be based on operative experience, historical AIS (Automatic Identification System) data, or known incidents in the area where the vessel will be operated [13], [14], [15]. In addition, it should be possible to generate a high number of encounters where the user has full control of which rules are triggered.

Already in 1987, Imazu proposed 22 traffic situations consisting of one, two or three ship encounters, often referred to as the Imazu problem, for testing automated collision avoidance algorithms [16], which has captured increased attention in recent years. The traffic situations cover a range of different requirements in the COLREG with varying complexity. However, a noticeable drawback with the Imazu problem is the fact that the states of the ships are fixed, while in reality even small variations in speed, course, or position may completely change the behavior of the automated navigation system. Another drawback with the 22 situations, which is also the case for other suggested traffic situations used in the literature (see e.g. [17], [18]), is that the ships in multi-ship encounters are set to collide at the same point in temporal- and spatial space, which is unlikely to happen in real life.

In this paper, we propose a new set of standard ship encounters that can serve as a baseline to test and verify basic functionality of a CAGA systems such as adhering to COLREG, and a tool, *Ship Traffic Generator*, to further alternate or generate new set of ship encounters to create multiple variants of a specific traffic situation. Increased complexity may also be added by introducing shallow areas through polygons. The ship encounters aim to cover the navigational aspects of the COLREG for power-driven vessels and the method is made freely available as a toolbox to vendors, students, and other actors involved in the development or assurance of CAGA systems [19]. The tool is intended to be used in combination with smart testing [20], a method for traversing the testing space in an efficient way to reduce the number of necessary test cases.

The remaining part of this paper is organized as follows: Section 2 gives a short overview of the COLREG rules triggered by the scenario set and how it should be used in a test environment. Section 3 introduces the standard ship encounter set, and describes the tool ship traffic generator, followed by a few examples of how the ship encounters are generated using the tool. Finally, section 4 concludes the paper.
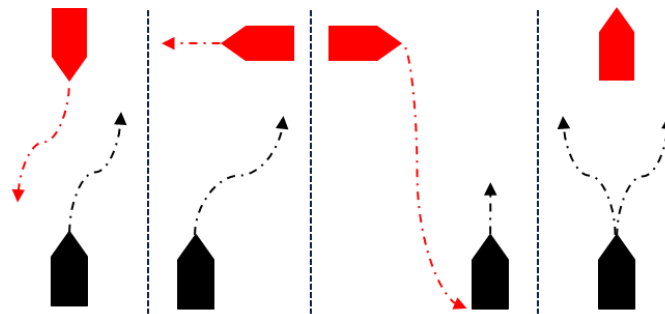
Figure 1: Correct vessel maneuver using COLREG. Own ship is black while target ship is red. From left, head-on (rule 14), crossing from starboard (rule 15/16), crossing from port (rule 15/17) and overtaking (rule 13/16) [21]. Figure based on [21]

2. Assuring COLREG behavior

2.1. COLREG

The Ship Traffic Generator will trigger rules in section II of the COLREG dealing with the *Conduct of vessels in sight of one another*, which are rule 13 – overtaking, rule 14 – head-on, rule 15 – crossing, rule 16 – action by give-way ship and rule 17 – action by stand-on ship, see Figure 1. In addition, rule 2 – responsibility and rule 8 - the maneuvering rule, are important. Rule 2 handles the responsibility between the ships during an encounter, in particular that each ship shall in general follow the rules and act safely, but if needed exempt from the rules in order to avoid immediate danger. Rule 8 describes required actions to avoid collision, meaning whenever a vessel is changing course and/or speed when in a collision situation, this should be done in a certain way. Rules 13, 14 and 15 handle situations where a ship is overtaking, head-on or crossing towards another vessel. Rules 16 and 17 are always combined with rules 13 and 15 to prescribe what a give-way ship and stand-on ship shall do in an overtaking and a crossing situation. In a head-on encounter, on the other hand, both ships shall give-way.

2.2. Chasing the failures

It is a clear consensus in the maritime industry that both conventional ships and ships controlled using automated navigation systems shall be navigated following the COLREG, ensuring a common understanding of the responsibilities between the ships. The CAGA system is part of a larger autonomous navigation system which is a highly complex system. At the same time, it is operating in a complex environment with an infinite number of potential traffic situations, both contributing to the need for massive amounts of testing to achieve a decent level of confidence in the system's behavior. The expectation from the most relevant stakeholders is that autonomous ships shall have a safety level at least equivalent to conventional ships. Translating this into tangible requirements, however, is not straight forward.

It is not possible to anticipate how a CAGA system could fail. It could be missing requirements, error in the software, lacking or faulty training data to name a few reasons. Situations which are difficult for human navigators may be easily handled by the CAGA system, but the opposite may also be the case; situations which human navigators handle without difficulty may be impossible for the CAGA system to solve. In addition, rare and unpredictable events could also arise as a consequence of deploying CAGA system. One may assume that traffic situations follow a long-tailed distribution [22]. Traffic situations with a high or medium probability to happen, will most probably happen during the first months or years of operation. Situations that have lower probability of occurrence could potentially happen after 10, 100 and maybe 1000 years of operation, see Figure 2.

In general, before a ship is delivered from the yard, the ship undergoes sea trials. However, for an autonomous ship, pilot operation after the sea trial with increased manning is necessary to evaluate and validate the behavior of the ship. Although sea trials and pilot operations uncover failures that are
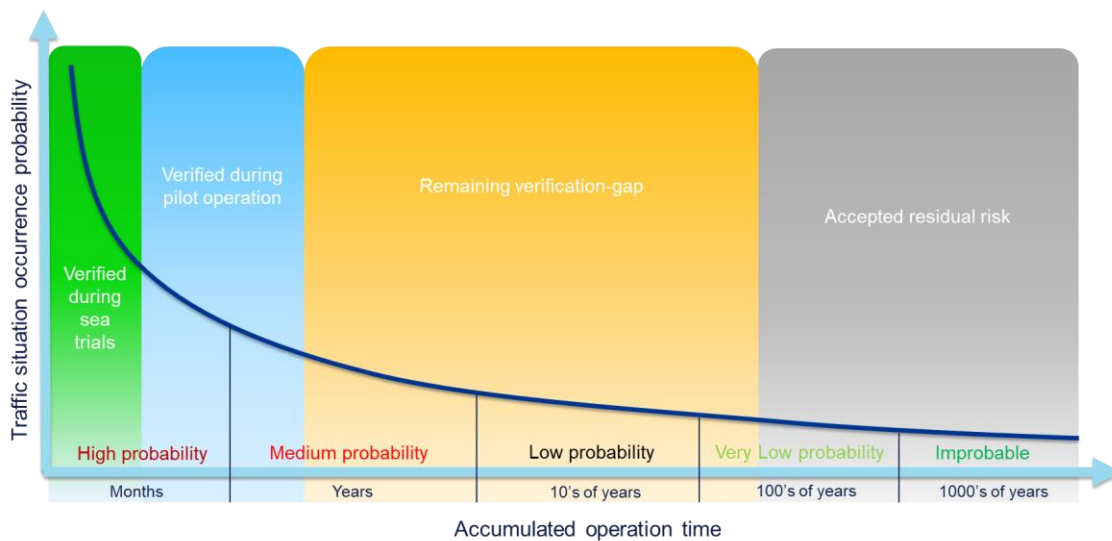
Figure 2: Chasing the failures. Verification during sea trials and pilot operation are not sufficient and numerous latent and unknown failures may still be in the system.

triggered with high and medium probability, there may still be many latent and unknown failures residing in the system as merely a fraction of the operation space has been tested. Simulation-based testing can be used to enable a much broader search in the operation space and can uncover failures with larger orders of magnitude in a more cost-effective manner. The verification process of the CAGA system aims to reduce the probability of safety-critical failures in the system down to an acceptable level through massive systematic and targeted testing, combined with the strengths of other verification tools. The acceptable level will vary dependent on the consequence of such failures.

2.3. Simulation-based testing

The aim of simulation-based testing is mainly to allow cost-effective, large-scale testing to identify and eliminate major faults in the CAGA system. This will systematically reduce the probability of failures and increase confidence in the system behavior. Figure 3 illustrates a possible test setup for verifying the CAGA system. The right side details the simulator setup which should include simulation models of the own ship and target ship(s). Own ship is navigated by the CAGA system under test, while the different target ships can be controlled by an equivalent or different CAGA system adhering, partly adhering or simply not adhering to the COLREG. The test manager and the
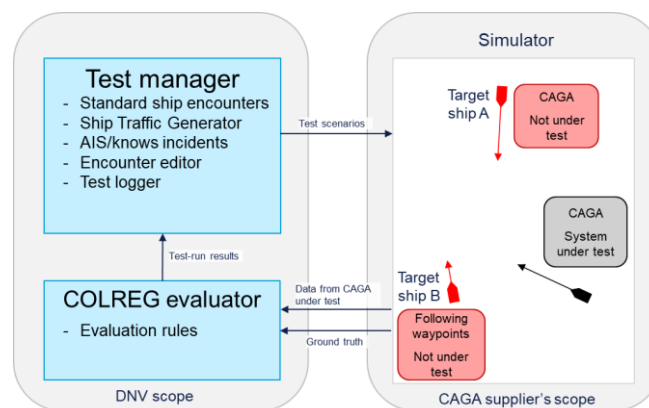

Figure 3: Proposed test setup for verification of CAGA system. Own ship navigated by the CAGA system under test in black and target ships in red. The test manager controls the ship encounters and the COLREG evaluator evaluates the results from the simulations.

COLREG evaluator are found on the left side of the figure. The test manager initiates the different ship encounters used when testing the CAGA system. The standard ship encounter set, as described in the next section, will normally be the first set of encounters used for this purpose. The next step of testing can then be performed using the ship traffic generator (see section 3.3). The test manager includes an encounter editor for manually inserting one or more encounters. A test logger can log all tests to be able to perform regression tests before installing updates to the CAGA system onboard the ships. The COLREG evaluator is used to evaluate how the CAGA system under test is handling the different encounters that have been simulated.

3. Set of ship encounters

3.1. Encounter classification

One-to-one encounters between the own ship and a target ship can be classified according to which rule in COLREG is most relevant. If a situation consists of several encounters, each encounter between own ship and a target ship is classified independently of the other encounters. Figure 4, which is based on [21], visualizes the different sectors used for classifying an encounter between two ships. Own ship is black and is encircled with orange, gray, purple, yellow, and blue sectors. The blue and purple sectors indicate overtaking (OT), either the own ship overtaking the blue target ship or being overtaken by the purple target ship. The yellow sector indicates the head-on (HO) sector if approaching the yellow target ship. The orange and gray sectors show the crossing (CR) sectors, indicating stand-on (SO) responsibility towards the orange target ship and give-way (GW) responsibility towards the gray target ship. The parameters θ_{crit}^{13}, θ_{crit}^{14}, θ_{crit}^{15}, θ_{min}^{15}, and θ_{max}^{15}, are used for the encounter classification and can be studied in Figure 4. As seen from the table, 3 parameters are tunable and may be changed if desired. For a more in-depth description of the parameters, the reader is directed to [23].

3.2. Standard ship encounter set

The set of standard ship encounters consists of 55 different situations and each situation consists of one, two, or three encounters triggering COLREG rules 2, 8, 13/16, 13/17, 14, 15/16, and 15/17. The standard ship encounter set is intended to be used for rapid "pre-screening" of the CAGA systems to verify that basic functionality is in place and that the system generally is able to operate in accordance with COLREG. If the CAGA software fails in handling any of the situations, software updates are required before a new round of testing is initiated using the standard ship encounter set. The situations

Table 1: Parameters used for encounter classification.

| Parameter	COLREG	Fixed/ tunable	Value	Parameter description
θ_{crit}^{13}	Rule 13	Tunable	67.5°	Tolerance for "*coming up with*" [8] relative bearing, used for classifying an overtaking encounter
θ_{crit}^{14}	Rule 14	Tunable	5°	Tolerance for "*reciprocal or nearly reciprocal courses*", "*when … in any doubt… assume… [head-on]*" [8], used for classifying a head-on encounter
θ_{crit}^{15}	Rule 15	Tunable	5°	Crossing aspect limit, used for classifying a crossing encounter
θ_{min}^{15}	Rule 13/15	Fixed	112.5°	22.5° aft of the beam, used for classifying a crossing and an overtaking encounter
θ_{max}^{15}	Rule 13/15	Fixed	247.5°	
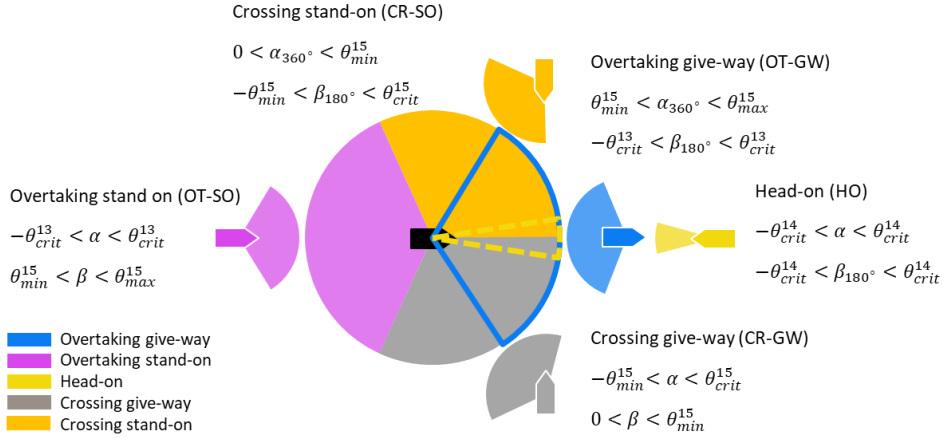
Figure 4: Encounter classification, based on [21]. An encounter is classified into head-on, overtaking give-way/stand-on and crossing give-way/stand-on.

in the standard ship encounter set are selected based on criteria stemming from an evaluation of historical traffic situations, interviews with experienced navigators, and an analysis of how different COLREG rules interact. In particular, the situations shall trigger the COLREG rules mentioned in section II [8]. In addition, it shall consist of both one to one and one to many ship encounters, and the traffic situations shall be realistic and probable to take place in real life.

An encounter is described as seen in Figure 5. Own ship (OS) is in the position (e_{OS_0}, n_{OS_0}) at some time t_0 with a course ψ_{OS_0}. After some time t, the own ship will be in the position (e_{OS_t}, n_{OS_t}) after traveling the distance r_{OS}. At the same time, a target ship (TS) is in the position (e_{TS_t}, n_{TS_t}) somewhere inside the encounter circle with a given circle and center point (e_{OS_t}, n_{OS_t}). For the standard ship encounters, it will be assumed that the radius of the encounter circle is 0 [nm] such that (e_{TS_t}, n_{TS_t}) and (e_{OS_t}, n_{OS_t}) are in the same point. It is now of interest to find the initial position of the target ship. By selecting the relative bearing, β, between the own ship and the target ship seen from the own ship, the relative speed of the target ship compared to own ship, and the travel time, it is possible to calculate the initial position of the target ship some distance s along the dashed line going from (e_{OS_0}, n_{OS_0}) with the angle β. The distance between (e_{TS_0}, n_{TS_0}) and (e_{TS_t}, n_{TS_t}) is r_{TS}.
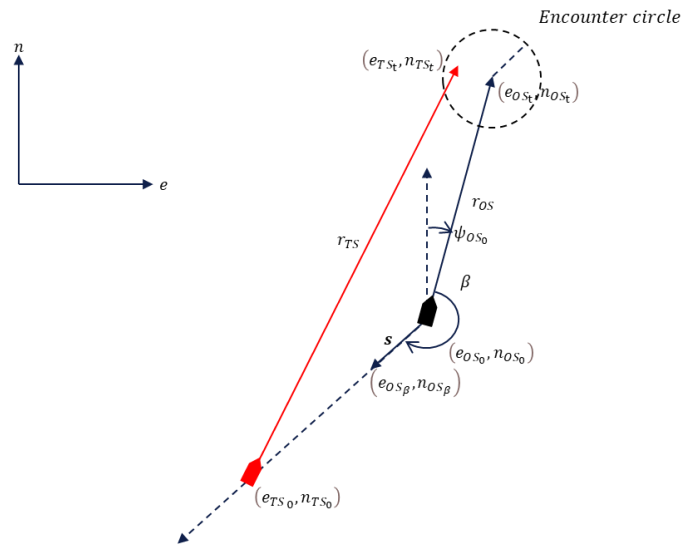

Figure 5: Encounter description. Own ship in black and target ship in red. Direction of the arrows indicates ship, while length of arrow indicates the distance each vessel will travel during a given time span. Position is given in north-east.
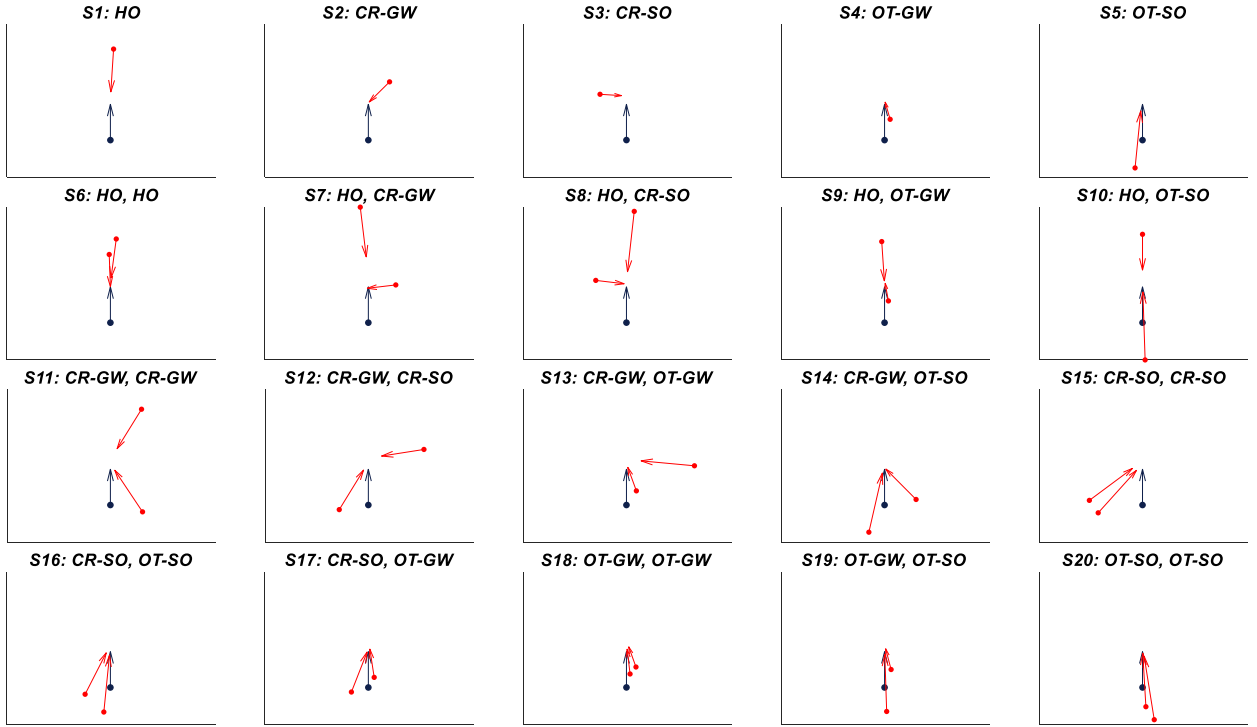
Figure 6: Baseline scenarios with own ship in black and one or two target ships in red. The initial position of the ships are indicated with the dots while direction and length of arrow indicates course and relative speed of the Ships.

Figure 6 shows 20 of the total 55 situations in the standard ship encounter set. Note that the situations with three target ships are not illustrated here, but they are following the same pattern. In the figure, the own ship is black while the target ships are red. The initial position of the ships are the dots, while the direction of the arrows indicates the navigational direction. The length of the arrows indicates the relative speed of the ships compared to each other. The corresponding values for the relative bearing between own ship and target ship seen from own ship, β, relative speed seen from target ship, v_{TS-OS}, and $TCPA$, can be found in Table 2 in Appendix.

The complete set of standard ship encounters including the encounter details are found in [19].

3.3. Ship Traffic Generator

As already mentioned, the set of standard ship encounters form the starting point of verifying the performance of a CAGA system. However, as failures sometimes can be triggered by minor changes to an encounter, a much larger set is needed to enable proper examination of the system behavior. A ship traffic generator can create an unlimited number of situations based on user input such as the number and type of encounters in a situation along with initial speed and course of own ship. The user may also specify the relative speed range between own ship and target ship(s) and the relative bearing range between own ship and target ship(s).

Figure 5 illustrates how the ship traffic generator defines an encounter. The user specifies the position of own ship together with speed and direction of own ship. Next, a time range for when own ship should meet a target ship is given together with the type of encounter. Based on the inputs from the user, the ship traffic generator calculates where own ship will be after some time t in the future. The position of the target ship will then be somewhere inside the encounter circle with a predefined radius with center point in own ship future position, see Figure 5. Using the type of encounter, the ship traffic generator will choose the relative bearing, β, and the relative speed, v_{TS-OS}, before calculating the initial position of the target ship. Relative bearing and speed may also be input from the user. The ship traffic generator is given as pseudocode in Algorithm 1 and 2.

Algorithm 1 Generate a traffic situation

1: **procedure** Pseudocode for generating a traffic situation
2:   Set encounter criteria
3:   Set the own ship speed, course, and position
4:   **for** n number of encounters in the scenario
5:     generate_encounter(encounter_type)
6:   **end for**
7: **end procedure**

Algorithm 2 Generate an encounter

1: **procedure** Pseudocode for generating an encounter
2:   Generate (time) vector $\Rightarrow$ time between 10 and 20 minutes
3:   Calculate the future position of OS
4:   Generate the encounter circle around OS future position
5:   Set the TS future position $\Rightarrow$ somewhere inside the CPA circle
6:   **while** TS initial position not set
7:     Generate β if not inputted by user $\Rightarrow$ based on desired encounter type
8:     Assign the TS speed if not inputted by user $\Rightarrow$ dependent on encounter type
9:     Find the TS initial position
10:    Check generated encounter
11:    Check if TS path crosses land
12:    **if** generated encounter type == desired encounter type **and** TS path do not cross land
13:      TS's initial position is set
14:    **end if**
15: **end for**
16: **end procedure**

Each situation is built up using single encounters. Own ship will have the same initial position, speed, and course for each encounter, while a varying time vector and by that a varying encounter circle ensures that the different target ships will have different CPA and time to CPA towards own ship. The user may choose between overtaking give-way, overtaking stand-on, head-on, crossing give-way, and crossing stand-on encounters.

3.4. Examples of situation

Two situation examples are given for demonstrating the ship traffic generator.

The first example, seen in Figure 7, is built up by two head-on encounters triggering rules 8 and 14. The initial position of own ship is $(0,0)$[nm] with speed 10 knots, and the heading pointing north with 0 [deg]. The left and middle figure display how the situation is built with encounter 1 and 2, respectively, while the right figure shows the resulting situation. The initial position of the own ship and the target ships are illustrated by dark blue and red dots, respectively. The length of the dashed lines in the left and the middle figure is given by ship speed multiplied with time until the own ship meets a target ship within the dotted CPA_{limit} circle. The direction of the dashed lines indicates the course of the ships. The dark blue and red solid arrows in the right figure illustrate the relative speed of the own ship and the target ships. The course is given by the direction of the arrows. Figure 8 displays another situation with three encounters. The initial position and course of the own ship are $(0,0)$[nm] and 15 [deg] (north/east), respectively. In this case, rules 2, 8, 15/16, and 15/17 will be triggered. In addition, a shallow area polygon is also included in this scenario to increase the complexity.

Like the previous example, the three figures starting from the left side of Figure 8 showcase how the scenario is generated while the right figure displays the resulting scenario.
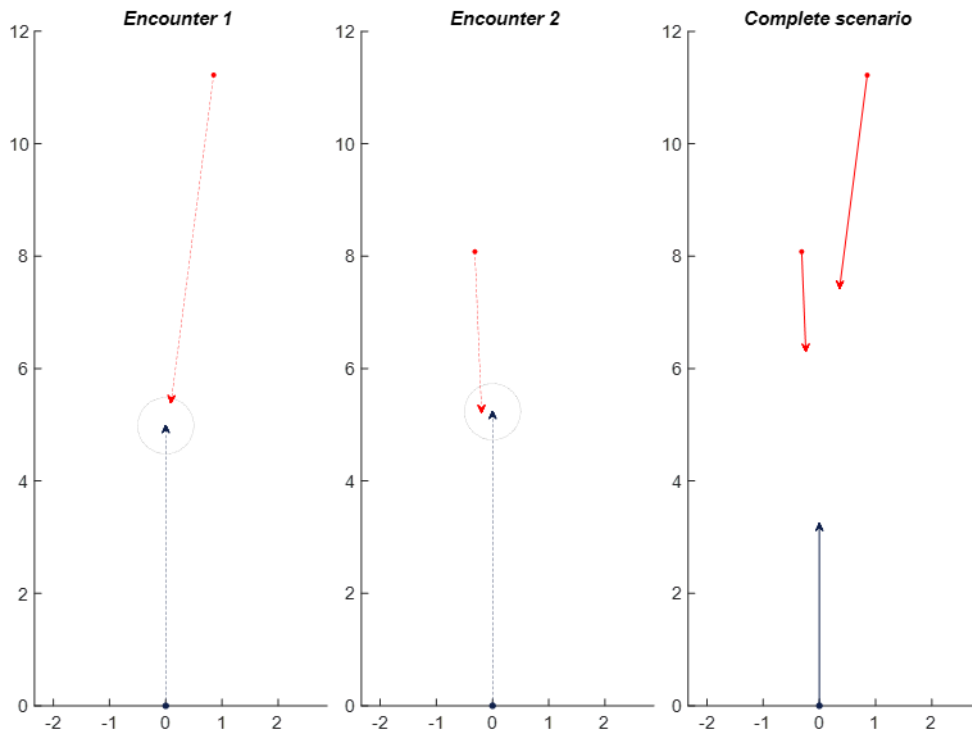
Figure 7: Constructing a head-on scenario where own ship (blue) is approaching two target ships (red). The left and middle figure display the construction of each head-on scenario while the right figure displays the resulting scenario. The length of the arrows in the right figure indicates relative speed between the ships while direction of the arrows indicates the course of the ships.
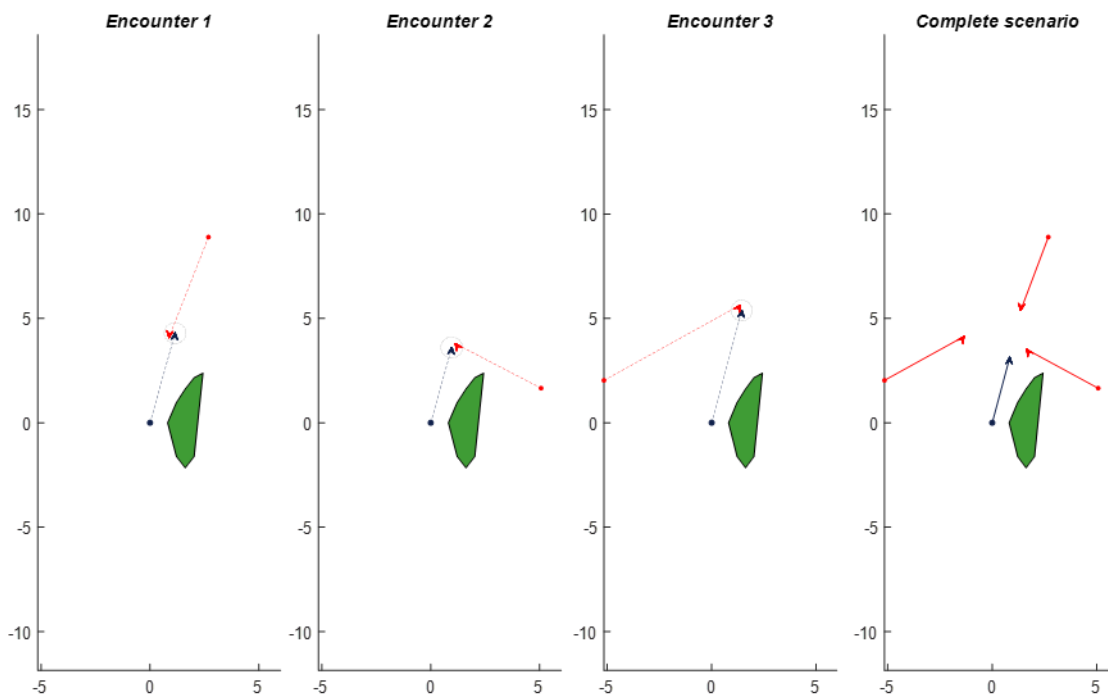

Figure 8: Constructing a combined head-on and crossing encounters including shallow area. The three figures to the left display the construction of the separate encounters while the figure to the right displays the complete scenario. The own ship is black while target ships are red. The length of the arrows in the right figure indicates relative speed between the ships while direction of the arrows indicates the course of the ships. Green area indicates shallow area.

4. Conclusion

In this paper, the need for simulation-based testing to close the verification-gap for the CAGA system has been discussed as it is evident that full scale and pilot operation alone will not be able to bring forward all the failures that may be in a CAGA system. The set of standard ship encounters combined with the ship traffic generator presented in this paper will therefore aid in structured and rigorous initial testing of a CAGA system. This is achieved by generating situations where own ship encounters one or more target ship(s) entering different variations of angles and speeds triggering the COLREG rules 2, 8, 13/16, 13/17, 15/16, and/or 15/17. The complexity may be increased by testing with maneuvering obstructions such as shallow areas. In addition, introducing target ships not approaching the own ship, but instead limiting the maneuvering area for the own ship, will increase the complexity of the situation even further.

The intended use of the generated traffic situations is to be combined with smart testing to achieve exploratory testing. In general, it is not possible to foresee where the system under test will fail, and the type of failure may be different from system to system. Hence, the set of standard ship encounters combined with the Ship traffic generator should be used for generating initial tests, both single and more complex multi-target situations. The results from these initial tests should then be used to guide where to test more or discover areas where the system is deficient. The same approach may be used when testing systems for lower level of autonomy such as decision support systems. Suppose the human operator will strictly follow what is suggested by the system, then the decision support system may be tested using the same set of situations as a fully autonomous ship.

The set of standard ship encounters and the tool ship traffic generator is made freely available, see Appendix A, and may be used by anyone that is working with autonomous functionalities such as vendors, students, or any other actors involved in the development or assurance of automated collision avoidance systems.

Reference

[1] NTNU, 09 2022. [Online]. Available: https://forskning.no/data-miljoteknologi-ntnu/verdens-forste-provedrift-av-selvkjorende-passasjerferge-i-by-starter-opp-i-trondheim/2081327. [Accessed 03 2023].

[2] Torghatten, 06 2023. [Online]. Available: https://kommunikasjon.ntb.no/pressemelding/torghatten-og-zeabuz-har-sjosatt-verdens-forste-kommersielle-utslippsfrie-autonome-passasjerferje-midt-i-stockholm?publisherId=17848259&releaseId=17972971. [Accessed 06 2023].

[3] NordJyske, 08 2022. [Online]. Available: https://nordjyske.dk/nyheder/aalborg/foererloes-faerge-jeg-doeber-dig-greenhopper/3348072. [Accessed 03 2023].

[4] Yara, 04 2022. [Online]. Available: https://www.yara.com/corporate-releases/crown-prince-and-youths-christen-worlds-first-emission-free-container-ship/. [Accessed 03 2023].

[5] O. Energy, 09 2022. [Online]. Available: https://www.offshore-energy.biz/asko-maritime-names-two-fully-electric-autonomous-sea-drones/. [Accessed 03 2023].

[6] Orca AI, "https://www.orca-ai.io/," 03 2023. [Online].

[7] Sea Machines, 03 2023. [Online]. Available: https://sea-machines.com/collision-avoidance/.

[8] IMO, "COLREG - Convention on the International Regulations for Preventing Collisions at SEA, 1972," 2003.

[9] C. Chauvin og S. Lardjane, «Descision making and strategies in an interaction situation: Collision avoidance at sea,» *Elsevier, Science Direct,* nr. Transportation Research Part F 11, pp. 259-269, 2008.

[10] T. A. Pedersen, J. A. Glomsrud, E.-L. Ruud, A. Simonsen, J. Sandrib and B.-O. H. Eriksen, "Towards

simulation-based verification of autonomous navigation systems," *Safety Science,* 2020.

[11] A. Basu, Software Quality Assurance, Testing and Metrivs, PHI Learning, 2015.

[12] M. Wood, P. Robbel, M. Maass, R. D. Tebbens, m. Meijs, M. Harb, J. Reach, D. Wittmann, T. Srivastava, M. E. Bouzouraa and M. Löhning, "Safety First for Automated Driving," 2019.

[13] A. Bakdi, I. K. Glad, E. Vanem and Ø. Engelhardtsen, "AIS-Based Multiple Vessel Collision and Grounding Risk Identification based on Adaptive Safety Domain," *Journal of Marine Science and Engineering,* 2019.

[14] R. J. Bye og P. G. Almklov, «Normalization of maritime accident data using AIS,» *Elsevier, Marine Policy,* vol. 109, 2019.

[15] R. J. Bye og A. L. Aalberg, «Maritime navigation accidents and risk indicators: An exploratory statistical analysis using AIS data and accident reports,» *Elsevier, Reliability Engineering & System Safety,* vol. 176, 2018.

[16] H. Imazu, "Research on collision avoidance manoeuvre, PhD Thesis," 1987.

[17] K. L. Woerner og M. R. Benjamin, «Real-Time Automated Evaluation of COLREGS-constrained Interactions Between Autonomous Surface Vessels and Human Operated VEssels in Collaborative Human-Machine Partnering Missions,» *OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO),* 2018.

[18] Y. Cai og K. Hasagawa, «Evalluating of Mainre Traffic Simulation System through Imazu Problem,» *The Japan Society of Naval Architects and Ocean Engineers,* 2013.

[19] DNV, "Ship Traffic Generator," DNV, 2023. [Online]. Available: https://github.com/dnv-opensource/ship-traffic-generator.

[20] T. A. Pedersen, Å. Neverlien, J. A. Glomsrud, I. Ibrahim, S. M. Mo, M. Rindarøy, T. Torben and B. Rokseth, "Evolution of Safety in Marine Systems: From System-Theoretic Process Analysis to Automated Test Scenario Generation," in *The 4th International Conference on Maritime Autonomous Surface Ships (ICMASS)*, Singapore, 2022.

[21] I. B. Hagen, Topics on Marine Collision Avoidance, Doctoral theses at NTNU, 2022:311, 2022.

[22] J. Adeojo, "The Extreme Power of Long Tailed Distributions," January 2021. [Online]. Available: https://towardsdatascience.com/the-power-of-long-tailed-distributions-bd46f8856039. [Accessed July 2023].

[23] K. Woerner, Multi-Contact Protocol-Constrained Collision Avoidance for Autonomous Marine Vehicles, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 2016.

Appendix A

The set of standard ship encounters and the Ship traffic generator tool are made freely available at [19].

Appendix B

Table 2: Baseline scenarios

ID	Number of encounters	Encounter type	Rules	Parameters settings
S01	One	HO	R8, R14	$\beta_1 = 2[deg]$, $v_{TS-OS_1} = 1.2[-]$, $TCPA_1 = 15[min]$, range $= 11[nm]$
S02	One	CR-GW	R8, R15/16	$\beta_1 = 20[deg]$, $v_{TS-OS_1} = 0.8[-]$, $TCPA_1 = 12[min]$, range $= 7.5[nm]$
S03	One	CR-SO	(R8), R15/17	$\beta_1 = 330[deg]$, $v_{TS-OS_1} = 0.6[-]$, $TCPA_1 = 17[min]$, range $= 6.4[nm]$
S04	One	OT-GW	R8, R13/16	$\beta_1 = 15[deg]$, $v_{TS-OS_1} = 0.5[-]$, $TCPA_1 = 14[min]$, range $= 2.6[nm]$
S05	One	OT-SO	(R8), R13/17	$\beta_1 = 195[deg]$, $v_{TS-OS_1} = 1.6[-]$, $TCPA_1 = 19[min]$, range $= 3.5[nm]$
S06	Two	HO & HO	R8, R14	$\beta_1 = 359[deg]$, $v_{TS-OS_1} = 0.9[-]$, $TCPA_1 = 11[min]$, range$= 8.2$ [nm] $\beta_2 = 4[deg]$, $v_{TS-OS_2} = 1.1[-]$, $TCPA_2 = 14[min]$, range$= 6.3$ [nm]
S07	Two	HO & CR GW	R8, R14, R15/16	$\beta_1 = 356[deg]$, $v_{TS-OS_1} = 1.4[-]$, $TCPA_1 = 20[min]$, range$= 14.0$ [nm] $\beta_2 = 36[deg]$, $v_{TS-OS_2} = 0.8[-]$, $TCPA_2 = 10[min]$, range$= 5.6$ [nm]
S08	Two	HO & CR SO	R2, R8, R14, R15/17	$\beta_1 = 4$, $v_{TS-OS_1} = 1.7[-]$, $TCPA_1 = 15[min]$, range$= 13.5$ [nm] $\beta_2 = 324$, $v_{TS-OS_2} = 0.8[-]$, $TCPA_2 = 13[min]$, range$= 6.3$ [nm]
S09	Two	HO & OT GW	R8, R13/16, R14	$\beta_1 = 358$, $v_{TS-OS_1} = 1.1$, $TCPA_1 = 13[min]$, range$= 9.8$ [nm] $\beta_2 = 10$, $v_{TS-OS_2} = 0.5$, $TCPA_2 = 16[min]$, range$= 2.7$ [nm]
S10	Two	HO & OT SO	R2, R8, R13/17, R14	$\beta_1 = 0[deg]$, $v_{TS-OS_1} = 1[-]$, $TCPA_1 = 17[min]$, range$= 10.7$ [nm] $\beta_2 = 176[deg]$, $v_{TS-OS_2} = 1.9[-]$, $TCPA_2 = 15[min]$, range$= 4.5$ [nm]
S11	Two	CR GW & CR GW	R8, R15/16	$\beta_1 = 18[deg]$, $v_{TS-OS_1} = 1.3[-]$, $TCPA_1 = 19[min]$, range$= 12.2$ [nm] $\beta_2 = 102[deg]$, $v_{TS-OS_2} = 1.4[-]$, $TCPA_2 = 15[min]$, range$= 4.2$ [nm]
S12	Two	CR GW & CR SO	R2, R8, R15/16, R15/17	$\beta_1 = 45[deg]$, $v_{TS-OS_1} = 1.2[-]$, $TCPA_1 = 19[min]$, range$= 9.5$ [nm] $\beta_2 = 261[deg]$, $v_{TS-OS_2} = 1.3[-]$, $TCPA_2 = 16[min]$, range$= 3.6$ [nm]
S13	Two	CR GW & OT GW	R8, 13/16, R15/16	$\beta_1 = 60[deg]$, $v_{TS-OS_1} = 1.5$, $TCPA_1 = 18[min]$, range$= 9.5$ [nm] $\beta_2 = 35[deg]$, $v_{TS-OS_2} = 0.7[-]$, $TCPA_2 = 15[min]$, range$= 2.1$ [nm]
S14	Two	CR GW & OT SO	R2, R8, R13/17, R15/16	$\beta_1 = 80[deg]$, $v_{TS-OS_1} = 1.2[-]$, $TCPA_1 = 12[min]$, range$= 3.9$ [nm] $\beta_2 = 210[deg]$, $v_{TS-OS_2} = 1.7[-]$, $TCPA_2 = 15[min]$, range$= 3.8$ [nm]
S15	Two	CR SO & CR SO	(R8), 15/17	$\beta_1 = 275[deg]$, $v_{TS-OS_1} = 1.5[-]$, $TCPA_1 = 17[min]$, range$= 6.4$ [nm] $\beta_2 = 260[deg]$, $v_{TS-OS_2} = 1.6[-]$, $TCPA_2 = 15[min]$, range$= 5.4$ [nm]
S16	Two	CR SO & OT SO	(R8), 13/17, R15/17	$\beta_1 = 255[deg]$, $v_{TS-OS_1} = 1.3[-]$, $TCPA_1 = 16[min]$, range$= 3.2$ [nm] $\beta_2 = 195[deg]$, $v_{TS-OS_2} = 1.6[-]$, $TCPA_2 = 15[min]$, range$= 3.1$ [nm]
S17	Two	CR SO & OT GW	R2, R8, 13/16, R15/17	$\beta_1 = 255[deg]$, $v_{TS-OS_1} = 1.2[-]$, $TCPA_1 = 13[min]$, range$= 2.1$ [nm] $\beta_2 = 30[deg]$, $v_{TS-OS_2} = 0.8[-]$, $TCPA_2 = 20[min]$, range$= 1.4$ [nm]
S18	Two	OT GW & OT GW	R8, R13/16	$\beta_1 = 15[deg]$, $v_{TS-OS_1} = 0.7[-]$, $TCPA_1 = 15[min]$, range$= 1.7$ [nm] $\beta_2 = 25[deg]$, $v_{TS-OS_2} = 0.6[-]$, $TCPA_2 = 19[min]$, range$= 2.7$ [nm]
S19	Two	OT GW & OT SO	R2, R8, 13/16, R13/17	$\beta_1 = 15[deg]$, $v_{TS-OS_1} = 0.6[-]$, $TCPA_1 = 20[min]$, range$= 2.3$ [nm] $\beta_2 = 175[deg]$, $v_{TS-OS_2} = 1.6[-]$, $TCPA_2 = 14[min]$, range$= 2.9$ [nm]
S20	Two	OT SO & OT SO	(R8), R13/17	$\beta_1 = 160[deg]$, $v_{TS-OS_1} = 1.8[-]$, $TCPA_1 = 15[min]$, range$= 4.1$ [nm] $\beta_2 = 170[deg]$, $v_{TS-OS_2} = 1.5[-]$, $TCPA_2 = 13[min]$, range$= 2.4$ [nm]