Human-System Interaction in Autonomy Method – a Structured Approach to Risk Monitoring

Marilia A. Ramos(*); Christoph Thieme

Department of Marine Technology, Norwegian University of Science and Technology - NTNU, Norway
(*) marilia.a.ramos@ntnu.no

Introduction

Research and development projects on autonomous systems have faced increasing interest, and some are currently in a testing phase. Autonomous systems' operation may be safer than traditional manned systems, since human error may be a contributing factor to many accidents. Nevertheless, a fully autonomous systems with no supervision and/or interference from humans are not expected soon. The operation will thus rely on a human-autonomous system (H-AS) collaboration. This interaction may not be constantly the same and the role and tasks of the operator may change. Then the autonomous system is designed with a dynamic Level of Autonomy (LoA), i.e., the LoA may change during operation depending on certain conditions.

As humans will still be involved in the operation at some level, human error may still occur [1–3]. In addition to human error, autonomous systems create new challenges, such as increased cyber security threats, detection of unforeseen conditions and actions from other people or the possibility of losing communication with other partners. Hence, risk assessments of operation are important [4]. They face two main challenges: i) the strong reliance on H-AS collaboration during the operation, and ii) the possibility of a dynamic LoA.

Few publications address topics related to hazards and risks associated with autonomous systems' operation. A recent review [4] of risk models aiming conventional and maritime autonomous surface ships (MASS) revealed that current approaches do not sufficiently model the functions carried out by software-based systems and that human operators are often treated superficially. Different operational modes of vessels are only covered to a limited extent. The current literature concerning autonomous systems does not model and analyse the H-AS interaction as potential contributor to the risk of operation, nor does it reflect the dynamic LoA of the operation. The Human-System interaction in

Autonomy (H-SIA) method intends to fill this gap. The method, although being developed foremost for MASS, is generic in nature, reproducible and structured.

This paper summarizes the H-SIA method, its background advantages and current limitations. More detailed information on the method and a case application can be found in the full article [5].

Methodology

The H-SIA method, presented in this Section, is initially composed of two elements: (i) an event sequence diagram (ESD), and (ii) a concurrent task analysis (CoTA). The method was specifically developed for and applied to collision scenarios between an autonomous ship and another vessel or object. Nevertheless, it is expected to have general applicability for autonomous systems.

Figure 1 presents the three main steps in the H-SIA method. Steps 2 and 3 are described in more detail in the following sub-sections. The general approach comprises familiarization (Step 1) to ensure that the analyst can apply the flowchart for the ESD development. The ESD development is the second step, where the ESD is built by answering design related questions of the autonomous system and the LoA of its operation. The developed ESD can be further analyzed with the CoTA (Step 3).

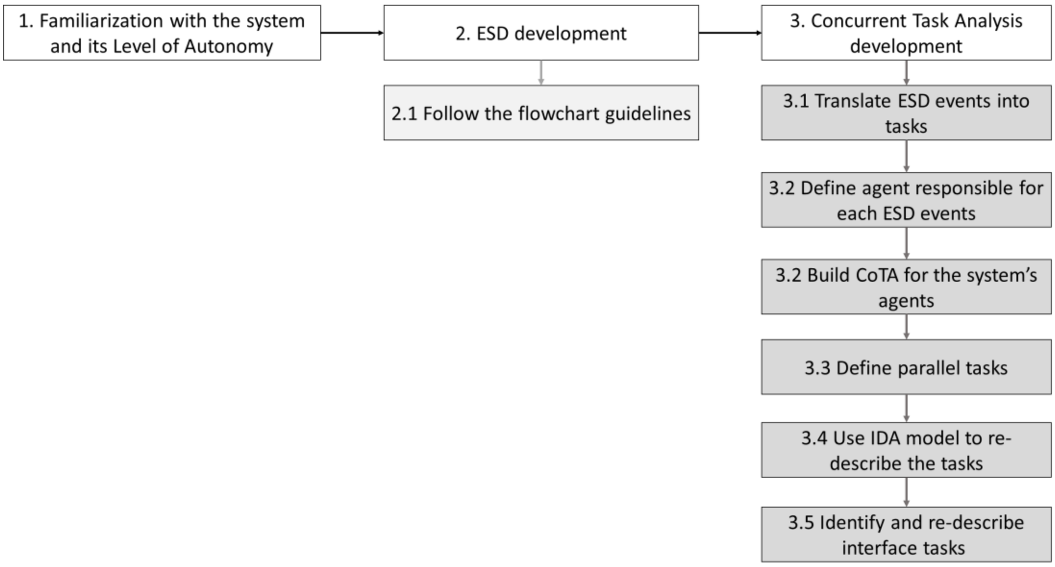
1. Familiarization with the system and its Level of Autonomy
2. ESD development
2.1 Follow the flowchart guidelines
3. Concurrent Task Analysis development
3.1 Translate ESD events into tasks
3.2 Define agent responsible for each ESD events
3.2 Build CoTA for the system's agents
3.3 Define parallel tasks
3.4 Use IDA model to re-describe the tasks
3.5 Identify and re-describe interface tasks

Figure 1: H-SIA method application steps (from *[5]*)

Figure 2 presents a general view of the H-SIA method results. The CoTA is success-oriented; it describes the tasks involved in the success paths of the events of the ESD. The interactions between the interface tasks of the agents are indicated with circles: a circle with an arrow exiting the event indicates that the task results in an output necessary to the accomplishment of a specific task of the other agent. Similarly, an arrow entering the event indicates a task that receives input from a specific task from the other agent. Interactions are identified by following the rules for task re-description and the CoTA stop rules. The events in the ESD cover either events related to the human operator or the autonomous ship. Some events may be related to both entities.

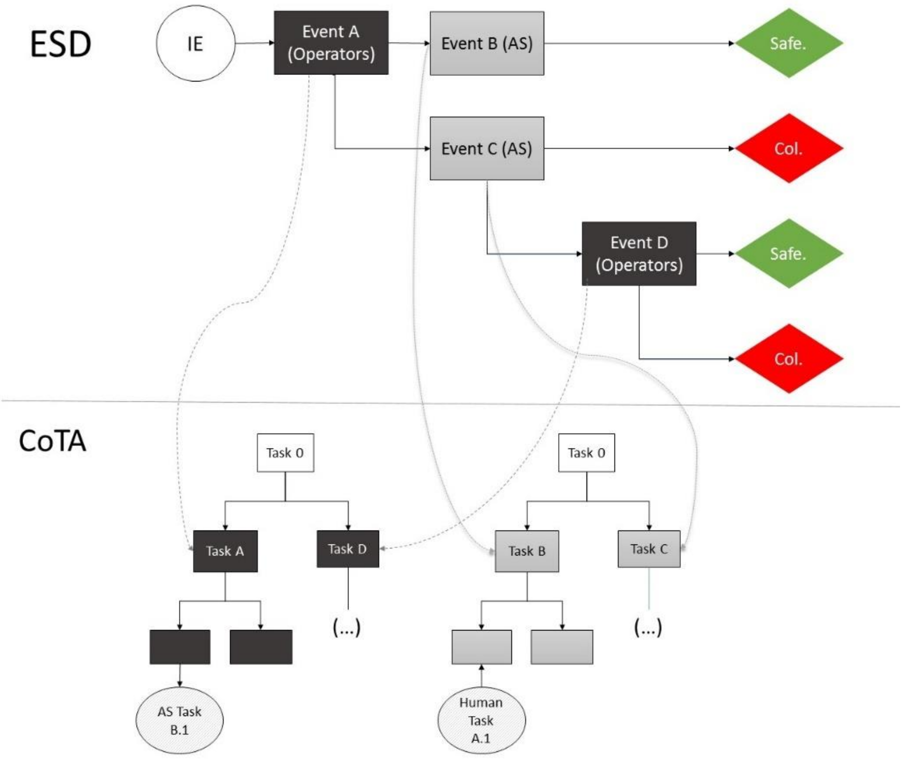

Figure 2: Simplified example of H-SIA method elements. (Adapted from *[5]*) Abbreviation: AS - Autonomous ship

Event sequence diagram and flowchart for development

ESDs are a generalized form of event trees. The ESD framework is flexible in modeling the behavior of key processes and hardware and operator state changes. The timing aspect is considered through the order of events. Thus, it is a more literal representation of a system state than event trees [6]. ESD are used, e.g., in the Phoenix Human Reliability methodology, which makes use of a flowchart approach to build a Crew Response Tree [7–9]. This is encouraging to apply the ESD framework and flowcharts for their development.

H-SIA provides a flowchart for the ESD development. The questions guide the building of the ESD and assist in including only relevant issues in the ESD that appear in the logic order of the questions. The use of the flowchart ensures traceability and reproducibility of the analysis. Furthermore, it provides the flexibility for assessing in the ESD development for different LoAs and system designs – from a LoA as low as remote control to high as fully autonomous. The flowchart and guidelines can be seen at [5].

Concurrent task analysis

The CoTA developed for the H-SIA method is built over Task Analysis (TA) theory and methods, and expanded to explicitly include the interactions between different parts or agents of the systems. TA was developed in the 1960s [10] and had the initial focus of analyzing human performance. Task analysis is "the collective noun used in the field of ergonomics, which includes HCI, for all the methods of collecting, classifying, and interpreting data on the performance of systems that include at least one person as a system component". Different forms to develop a TA exist, such as Hierarchal TA, Tabular TA, and Cognitive TA [11].

TA allows analyzing complex tasks through the decomposition of goals into sub-goals, so called re-description. The goals and sub-goals are organized in HTA through plans [10]. Plans state the order of the sub-goals to achieve the main goal. From a systems perspective, the HTA should focus on the analysis of the task to understand how the system is supposed to behave and how it may fail. An important element of HTA are the stop rules that determine when to end the re-description. In this work the stop rules are based on the Information Decision Action (IDA) framework.

The IDA model was initially developed as a human behavior model for the operation of nuclear powerplants [12]. It consists of the cognitive phases I (Information collection and pre-processing); D (decision making and situation assessment); and A (action taking). The IDA model has been developed and extended further in recent years [12–16]. It is possible to adapt IDA to different agents of a system. Since the H-SIA method analyzes the interaction between two or more agents, it is beneficial to use a similar model that allows for decomposing functions into the same low-level unit of analysis. In the H-SIA method, thus, IDA model was extended to describe phases and categorize tasks of the autonomous ship as well.

The CoTA consists of several TAs, in which the tasks described as the events in the ESD are re-described until the tasks correspond to one of the IDA phases and the relationship between the sub-task and another agents' task can be established, if this exists. In addition, the CoTA includes a new type of task named "parallel task". Parallel tasks are supporting tasks, i.e., they are necessary for the execution of the other tasks and the interaction between the agents but not explicitly included in the ESD. Parallel tasks are related to the normal operation of the system being executed continuously, not following a specific order in a plan, i.e., they are executed at the same time with the other tasks. The parallel tasks are normally the ones related to gathering data, monitoring, or communication between the agents.

The CoTA is based on the ESD developed in step 2. The events from the ESD translate into tasks that are performed by the agents. Hence, the ESD presents *what* can happen, and the CoTA further details *how* these events may occur. The CoTA is a success-oriented method that enables the analyst to understand better each agent's tasks that needs to be accomplished for the events of the ESD to take place.

For instance, an event in the ESD may be "Detection of the collision candidate by the autonomous ship". This event is translated into the task "Detect the collision candidate" in the AS' Task Analysis. This task is then re-described using the CoTA stop-rules. The re-description details the sub-level tasks that must be accomplished for the AS to successfully detect the object as a collision candidate, e.g.: gathering and processing data, apply relevant norms, among others.

There are two main approaches when using the CoTA: Analyze the tasks involved in all events of the ESD (i), or to (ii) analyze a specific sequence of events in the ESD scenario. When developed for all the events of the ESD (alternative i), the CoTA provides a detailed overview of how the agents should act to be successful in the possible events of the ESD. The scenario specific CoTA (ii), presents the tasks that should be performed for a success outcome in a specific sequence of events.

The CoTA adopts and expand the HTA plans described in [17]. The CoTA plans describe the order of sub-tasks in order to achieve a successful main task. The CoTA plan may determine for instance a sequence (e.g. 1→2→3 – the tasks 1, 2, and 3 must be performed in this order); or a decision (e.g., Task 1 is performed and, if a condition is satisfied, task 2 is performed; if no, task 3 is performed). In addition, it contains the parallel tasks, and a scenario-specific plan.

The CoTA can be developed from the ESD following the steps below, the relationship between CoTA and ESD is highlighted in Figure 2:

1. Definition of agents to be analyzed, each of the agents will have an HTA;
2. Definition of Task 0, this may be to avoid collision and recover successfully from the initiating event;
3. Definition of agents that are mainly acting in each event agents;

4. Definition of high-level tasks: each event of the ESD translates into a high-level task in each of the respective HTAs. It is recommended to develop a table for correspondence between the event from the ESD and the Task ID in the CoTA;
5. Identification of parallel tasks;
6. Re-description of tasks until stop rules are satisfied. The first rule always must be satisfied, whereas the second may not be satisfied.
   i) The task is associated with only one of the I-D-A phases and, for the dependent tasks;
   ii) The task represents the interaction with another agent.
7. Identification and highlighting of interface tasks.

The CoTA can be used for multiple purposes, such as development of procedures, identification of specific subsystems and components that are necessary for a successful task, identification of failure sources of the human operator or the autonomous system identification of tasks that need to be accomplished for a certain outcome, identification of interface tasks, and analysis of failure propagation.

The scenario specific CoTA

As stated previously, the CoTA may be used for analyzing a specific sequence of events instead of all events of the ESD. This may be achieved from the complete CoTA or directly from the specific sequence of events. In both cases, the development of the scenario specific CoTA starts with the identification of the events involved in the desired ESD path. To make use of the complete CoTA, the analysts identify and selects the tasks of each agent's TA that belong to that sequence. This process may be assisted by the table developed in Step 4. When developing the CoTA from the sequence of events, the analysist follows all the steps outlined above, just for these specific tasks. An example of a scenario specific CoTA is shown in Figure 3.

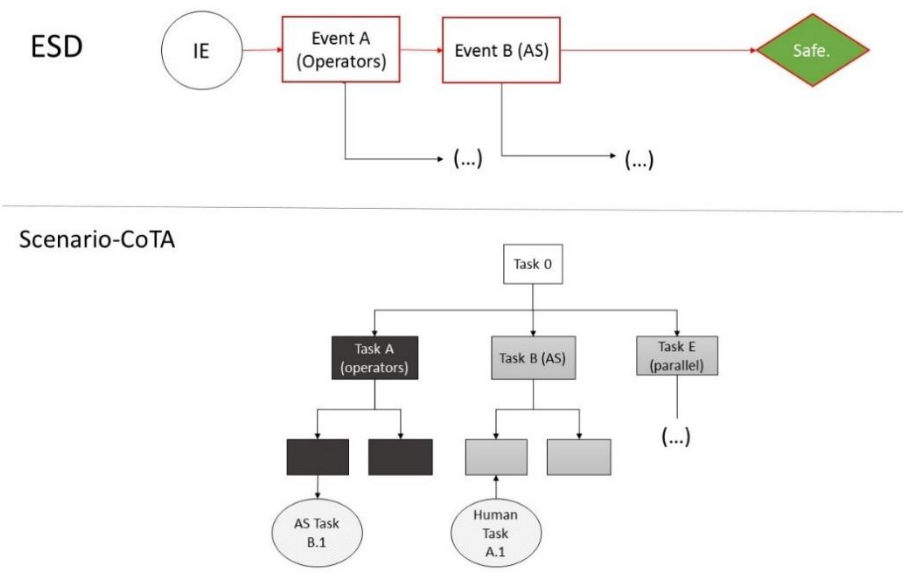

Figure 3: Scenario-specific CoTA example (Adapted from [5])

Discussion and conclusion

In the H-SIA method an autonomous system is analyzed as whole, rather than focusing on each component separately. The process may assist in the comparison of different concepts and designs of an autonomous system. The use of a generic flowchart and generally valid principles produces results that are comparable, reproducible and traceable. An additional benefit of the H-SIA method is the identification and tracking of interdepend tasks of different agents in a system.

The features of the ESD and CoTA makes the H-SIA method a valuable technique for analysis of safety of autonomous systems' operations. It may be used in the design phase, to develop procedures and to derive specifications, for failure events identification, and the results can be further integrated into risk assessments.

Some limitations of the methods are that although the CoTA is developed using clear guidelines and stop-rules, the identification of parallel tasks and the re-description depends also on the analyst. This may lead to different CoTAs when the H-SIA method is used by different analysts. This variability is, in one sense, a limitation of the method. On the other hand, it offers flexibility for the CoTA to be developed and detailed according to the purpose of the analysis.

Future work includes the detailing of the failure events, through e.g., the development of fault trees and BBNs, in a hybrid causal logic model. The method can benefit from validation through applications to existing autonomous systems and projects, as well as through feedback from experts use.

References

[1] Rødseth ØJ, Tjora A. A risk based approach to the design of unmanned ship control systems. Proceeding Conf Marit Technol 2014:153–62.

[2] Ramos M, Utne IB, Vinnem JE, Mosleh A. Accounting for human failure in autonomous ships operations, Taylor & Francis Group; 2018, p. 355–63.

[3] Ramos M, Utne IB, Mosleh A. Collision avoidance on maritime autonomous surface ships: operators' tasks and human failure events. Saf Sci 2018.

[4] Thieme CA, Utne IB, Haugen S. Assessing ship risk model applicability to Marine Autonomous Surface Ships. Ocean Eng 2018;165:140–54. doi:10.1016/j.oceaneng.2018.07.040.

[5] Ramos M, Thieme CA, Utne IB, Mosleh A. Human-System Concurrent Task Analysis for Maritime Autonomous Surface Ship Operation and Safety. *Submitted to* Reliab Eng &andSystems Saf

[6] Swaminathan S, Smidts C. The Event Sequence Diagram framework for dynamic Probabilistic Risk Assessment. Reliab Eng &andSystems Saf 1999;63:73–90.

[7] Ekanem NJ, Mosleh A, Shen S-H. Phoenix–A model-based Human reliability analysis methodology: Qualitative analysis procedure. Reliab Eng Syst Saf 2015;145:1–15. doi:10.1016/j.ress.2015.07.009.

[8] Ekanem NJ, Mosleh A. Phoenix – A Model-Based Human Reliability Analysis Methodology : Quantitative Analysis Procedure and Data Base. Proc. to Probabilistic Saf. Assess. Manag. PSAM 12, Hawaii: 2014.

[9] Ramos MA. A Methodology for Human Reliability Analysis of Oil Refineries and Petrochemical Plants. Federal University of Pernambuco, 2017.

[10] Shepherd A. Hierarchical Task Analysis. London: Taylor & Francis; 2001.

[11] Annett J, Stanton N. Tasks Analysis. 2000.

[12] Smidts C, Shen SH, Mosleh A. The IDA cognitive model for the analysis of nuclear power plant operator response under accident conditions. Part I: problem solving and decision making model. Reliab Eng Syst Saf 1997;55:51–71. doi:http://dx.doi.org/10.1016/S0951-8320(96)00104-4.

[13] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents Part 3: IDAC operator response model. Reliab Eng Syst Saf 2007;92:1041–60. doi:10.1016/j.ress.2006.05.013.

[14] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents . Part 2 : IDAC performance influencing factors model 2007;92:1014–40. doi:10.1016/j.ress.2006.05.010.

[15] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents Part 5: Dynamic probabilistic simulation of the IDAC model. Reliab Eng Syst Saf 2007;92:1076–

101. doi:10.1016/j.ress.2006.05.012.

[16] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 4: IDAC causal model of operator problem-solving response. Reliab Eng Syst Saf 2007;92:1061–75. doi:10.1016/j.ress.2006.05.011.

[17] Annett J. Hierarchical Task analysis. In: Diaper D, Stanton NA, editors. Handb. Task Anal. Human-Computer Interact., London: Lawrence Erlbaum Associates; 2008, p. 67–82.