Experience with a Model-based Safety Analysis Process for Autonomous Service Robot
Damien Martin-Guillerez, Jérémie Guiochet, David Powell

To cite this version:

HAL Id: hal-01285189
https://hal.archives-ouvertes.fr/hal-01285189
Submitted on 8 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Experience with a Model-based Safety Analysis Process for Autonomous Service Robot

Damien Martin-Guillerez∗†, Jérémie Guiochet∗†, and David Powell∗†

∗ CNRS ; LAAS ; 7 avenue du colonel Roche, F-31077 Toulouse, France
† Université de Toulouse ; UPS, INSA, INP, ISAE ; LAAS ; F-31077 Toulouse, France

Email: firstname.lastname@laas.fr

Abstract—Safety is a major concern for autonomous systems that physically interact with humans, such as service robots. However, modeling dynamics of such systems is hard so classical safety analysis methods need to be adapted. In this paper, we propose an approach based on a combination of well-known safety analysis techniques. We propose to describe scenarios of use with the common Unified Modeling Language. Risk analysis is then performed using a Preliminary Hazard Analysis, an adaptation of the HAZOP method and the classical Fault Tree Analysis. This paper explains the overall process and illustrates it through the example of the MIRAS projects which aims to develop a robotic strolling assistant that will help disabled persons to stand, sit and walk.

Safety assessment process, risk assessment, autonomous systems

I. INTRODUCTION

Safety is now a major concern in many computer-based systems and more particularly for autonomous systems such as service robots in physical contact with human. The traditional approach to analyze safety of such systems is to use methods such as Fault Tree Analysis (FTA) or Failure Mode, Effects, and Criticality Analysis (FMECA). Those methods are usually based on models of the systems such as block diagrams or functional decomposition and automata for dynamics. For autonomous systems it is impossible to represent dynamics with automata as they evolve in an unstructured environment, including humans. A functional decomposition of the decision architecture is also impossible. Moreover, the fact that environment is not structured means that the number of operating conditions is essentially infinite.

We propose an approach to cope with these issues through the combination and adaptation of several well-known techniques. We consider that the earliest models of a system usually describe scenarios of use of the system. We claim that the analysis of deviations of such scenarios allows the identification of major risks. We propose to describe scenarios of use with the common Unified Modeling Language (UML [1]), and to analyze risks using with a Preliminary Hazard Analysis (PHA), the guideword-based collaborative method HAZOP (HAZard OPerability) [2] and Fault Tree Analysis (FTA) [3]. A major advantage of using UML as input model is that it is now a de facto standard for system description, and non-experts can easily understand diagrams such as sequence and use-case diagrams. HAZOP analysis is also well-adapted to the initial steps of the development as it is easily understandable, and through the use of guidewords, it enables a systematic analysis.

This process has been successfully applied to robotic projects and we illustrate each section with the MIRAS project [4]. The objective of this project is to develop an assistive robot for standing up, sitting down and walking, and also capable of health state monitoring. It is designed to be used in elderly care centers by people suffering from gait and orientation problems. It is composed of a mobile base and a moving handlebar (Figure 1).

Fig. 1. Robuwalker – First prototype

This paper is structured as follows. We present our general process in Section II. This approach is detailed step by step with its application to the MIRAS project in Section III. In Section IV, we discuss the validity of our approach. Section V gives an overview of a tool we developed to support this approach. We present related work in Section VI. Section VII concludes this paper.

II. METHOD OVERVIEW

In safety critical systems, safety assessment is usually performed using a safety assessment process [6] where the objective is to reduce the risk to an acceptable level. This process and its terminology is now quite stable in industrial standards [7]. It is based on a decomposition of activities into risk analysis, risk evaluation and risk reduction [5]. Risk analysis aims to identify hazards and estimate the risk. Risk evaluation is a step for comparing the estimated risk against given risk criteria to determine the acceptability of the risk.
Risk reduction is a process in which decisions are made and measures implemented by which risks are reduced to, or maintained within, specified levels. In order to complete risk analysis many techniques have been developed (Fault Tree Analysis, Failure Mode Effects and Criticality Analysis, Event Tree, etc.) and applied in very different safety critical domains from nuclear power-plants to medical robots.

In this context, we based our approach on the classical safety assessment process as described in Figure 2. The same cycle is repeated until the designed system achieves tolerable risk. When applying this process, it is important to:

- describe the target of evaluation at the right level of abstraction;
- facilitate communication and interaction between different stakeholders involved in the safety assessment process (e.g., in our case, the stakeholders are the patients, the medical staff, and the robotics experts);
- manage the combinatorics of risk analysis, which often results in an excessive number of documents and models;
- document safety analysis results and the assumptions on which these results depend to support reuse and maintenance.

We chose a subset of UML to describe the system, communicate with the stakeholders and organize safety analysis documents. UML is a standard general-purpose modeling language that includes a graphical notation enabling the representation of an abstract model of a system [1]. The UML model of a system is composed of different UML diagrams, each of which is a partial graphical representation of the system that concentrates on a particular viewpoint. Two diagrams are commonly used for description of the system usage: use case and sequence diagrams. Use cases represent intended use of the system and are linked with the actors that can trigger scenarios of the use case. Each use case is further documented by fields such as pre and post conditions. Each sequence diagram represents one particular scenario of one use case. Those two diagrams, as presented in the process view of Figure 2, are also the input for hazard identification step.

Hazard identification is then performed using a Preliminary Hazard Analysis (PHA [8]) and an adaptation of HAZOP and UML [9]. Through the HAZOP method, a system is analyzed by holding a review of the systematic generation of deviations defined by the conjunction of parameters of the system (e.g., pressure, temperature...) and guidewords (e.g., no, more, less...). We apply the PHA in a standard way but adapt the HAZOP method to apply it to UML use cases and sequence diagrams.

Risk estimation is then carried out after the HAZOP analysis. Quantitative risk estimation is however impossible to obtain, mainly because of the impossibility to evaluate probabilities of identified hazards. However, it is important to propose risk reduction means even if probability is not calculated. We cope with this problem by defining different objectives regarding the iterations of the development process. For instance, during the first iteration, only severity (see Table I) is considered, and when the cost (in term of design changes impact, development efforts) is acceptable by robotic experts, risk reduction is carried out. As a result of the first iteration, the document Hazards List is produced (see Figure 2). The next step is based on the PHA and HAZOP-UML hazards, which are analyzed with the Fault Tree Analysis (FTA). This leads to an estimation of the final risks (document Risks List in Figure 2) and also to new recommendations.

### III. Safety process steps

In this section, we detail the various steps of the safety process. These steps are illustrated by the application in the MIRAS project.
A. Definition of intended use

In the first step, the intended use of the system is modeled using UML use cases and scenarios with UML sequence diagrams. The use cases are completed by conditions of use: preconditions to be fulfilled before any action of the use case can be performed, postconditions to be satisfied at the end of the use case and, invariants that must hold during the use case. Interactions are represented by messages in sequence diagrams. Messages can be annotated (e.g., to specify whether the interaction is physical or cognitive).

These diagrams are often used for high-level representations of a system. They are easy to understand even by non-experts so they are suitable as a means for presenting the scenarios to the different stakeholders. Furthermore, they can be included as part of the documentation for the certification process. Sequence diagrams are highly expressive yet can remain quite simple when used to describe use scenarios. This simplicity makes them an attractive support for hazard identification by deviation analysis since it helps to keep the combinatorial aspects of such analysis under control. Throughout the modelling process, preliminary safety remarks and recommendations can be issued.

In the MIRAS project, we identified 11 use cases: Strolling (UC01), Standing up operation (UC02), Sitting down operation (UC03), Balance loss handling (UC04), Summoning and autonomous movement of the robot (UC05), End of use detection and movement to a waiting position (UC06), Positioning the robot by hand (UC07), Alarm handling (UC08), Patient profile programming (UC09), Patient profile learning (UC10), and Robot set-up (UC11).

By way of an example, Table II list the conditions put on use case UC02 (Standing up operation) and Figure 3 describes the nominal scenario for this use case. This sequence diagram depicts the interaction between the users and the robot as the potential source of harms. Robot decisions are represented as self-messages.

B. Hazard identification

To identify hazards that can arise from the use of the robot, we used two complementary techniques: Preliminary Hazard Analysis (PHA) and UML-HAZOP (UML - HAZard OPerability).

A standard PHA [8] is applied at an early stage of the design. The various stakeholders of the project meet together for several workshops (two in the MIRAS project). During the workshops, participants try to consider all the possible causes of hazards in the system (e.g., environmental, electrical, mechanical, hardware/software and human). For each cause, the participants identify the hazards that can arise. In the MIRAS project, the PHA led to the identification of 45 hazards (Tab. III).

We then apply the UML-HAZOP method. We adapt the HAZOP [2] method to analyze deviations of the UML use cases and sequence diagrams. According to the Defence Standard 00-58 [11], HAZOP analysis is the systematic identification of every deviation of every attribute of every entity. Each deviation is a potential hazard that can lead to a harmful event. We adapted the guideword lists to apply them to attributes of use cases and sequence diagrams. The guideword list we use for the use case entity is given in Table V and an extract of the analysis of UC02 (Standing up operation) is presented in Table VII. All guidewords are applied to generate deviation. The analyst then establishes the effect at the use case level, and the result in the real world. The other columns of the table guide the analyst to establish a severity level, to deduce requirements and otherwise make remarks on that deviation. The complete method is presented in [9].

In the MIRAS project, the analysis of 297 deviations led to the identification of 13 hazard classes (Table IV). This table presents the main hazardous situations of the system.

<table>
<thead>
<tr>
<th>Num.</th>
<th>Severity</th>
<th>Type of injury</th>
<th>SIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>None</td>
<td>None</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>Minor</td>
<td>Superficial injury</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Moderate</td>
<td>Recoverable</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Serious</td>
<td>Possibly recoverable</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Severe</td>
<td>Not fully recoverable without care</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Critical</td>
<td>Not fully recoverable with care</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>Fatal</td>
<td>Not survivable</td>
<td>4</td>
</tr>
</tbody>
</table>

---

**TABLE I**

**Severity list used in the MIRAS project, derived from the abbreviated injury scale of [10])**

**TABLE II**

**UC02 “Standing up operation”**

<table>
<thead>
<tr>
<th>Use case name</th>
<th>UC02. Standing up operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>The patient stands up with the help of the robot</td>
</tr>
<tr>
<td>Precondition</td>
<td>The patient is sitting down</td>
</tr>
<tr>
<td>Postcondition</td>
<td>The patient is standing up</td>
</tr>
<tr>
<td>Invariant</td>
<td>The robot holds both handles of the robot</td>
</tr>
</tbody>
</table>

---

**FIG. 3. Sequence diagram UC02.SD01 giving main scenario of UC02 “Standing up operation”**
In the HAZOP analysis, each deviation that potentially leads to a hazard class is labeled with the corresponding number (column “Hazard Number”, Table VII). Table VI) gives an extract of the list of recommendations resulting from application of the UML-HAZOP method. This list is derived from the “new safety requirements” column of the UML-HAZOP tables.

C. Risk estimation

During the first iteration of the process, a qualitative approach is followed to estimate the risk associated with each hazard class (severity column of Table IV). Safety Integrity Level (SIL [12]) requirements are estimated when
performing HAZOP. To do this, an indicative SIL is assigned to each severity level (cf. Table I, SIL column) and thereby to safety-related components associated with each HAZOP deviation (cf. Table IV, integrity level requirements column) according to the most severe hazard induced by this deviation (e.g. the patient position detection system should be SIL1).

During the second iteration of the safety assessment process, we extract FTA top events from the hazard class list of Table IV. The FTA gives top event occurrences and enables the computation of quantitative risk. Fault tree analysis was carried out in the usual way so, given the space limitation, we do not detail it there. The only point we want to stress is the link between FTA and the other artifacts of the safety analysis process.

**D. Risk evaluation**

As mentioned before, risk evaluation consists in comparing the estimated risk with given risk criteria to determine the acceptability of the risk. Even if some proposals have been made for acceptability criteria [13] there is no set of generally accepted risk acceptance principles for service robotics. This is an important issue but it deals with political and ethical concerns. Nevertheless, criteria are needed for engineers to determine which risks have to be reduced. This step is again an iterative process. In the MIRAS study, after the first iteration (coming after a preliminary risk estimation only considering hazard severity), each hazard and possible recommendations were proposed to the robotics experts, and on the basis of our risk analysis results they proposed a classification of risk acceptance according to three versions of the robot: development version for using the robot in the laboratory, evaluation version for the prototype used in hospitals for clinical evaluation, and final version for operational life.

In a second iteration of this activity, the final risks as presented in Table IV with a severity and an occurrence estimation are classified using the ALARP (As Low As Reasonably Practicable) principle. Three risk level zones are defined: intolerable region, the ALARP region, and the broadly acceptable region. A risk can stay in the ALARP region only if further risk reduction is impracticable or if its cost is disproportionate to the improvement gained. For each risk, the corresponding sources of deviations and hazards, and the associated recommendations are evaluated for judging acceptability. The final result is an argumentation for the final acceptable risk of the system.

**E. Risk reduction**

Recommendations were issued as a result of each step of the safety analysis process. UML modeling, PHA and UML-HAZOP give rises to general recommendations. Integrity level requirements are issued during UML-HAZOP analysis. A specific integrity level requirement leads to specific safety recommendations given by IEC 61508 [12]. These recommendations are applied to reduce risks if their current level is not tolerable. Of course, design recommendations should be applied first, then protective devices then, if no other solution exists, information for users [6].

Some risk reduction techniques will reduce the severity of the corresponding hazard (e.g., a bumper added to the robot will reduce the severity of a collision) while others will reduce the probability of occurrence (e.g., infrared sensors to detect the patient’s feet will decrease the probability of a collision between the patient and the robot). The next version of the MIRAS robot (Fig. 4) is now under validation. In addition to various ergonomic corrections, the MIRAS robotics experts\(^1\) took into account the risk reduction recommendations resulting from the first cycle of our safety analysis.

\(^1\)ISIR - Institut des Systèmes Intelligents et de Robotique (http://www.isir.fr) and ROBOSOFT (http://www.robosoft.fr)
The new version reduces the severity of hazard classes 8, 9 and 10 and reduces the occurrence rate of hazard classes 1, 6, 7, 8, 9 and 11 (Tab. IV, columns Severity and Occurrence rate). This new version especially applies the recommendations and meets the requirements asked for the evaluation version of the robot (i.e. a robot that will be used in the laboratory and used for clinical evaluation).

IV. EVALUATION OF THE METHOD

We evaluate our approach from four different perspectives: integrability into the development process, and usability, validity and applicability of the method.

Integrability: in our approach, UML design models are shared with the development process. Deviation analysis can be carried out at the same time as design refinement or testing/coding by the development team. The results of risk assessment and of testing can be used to generate another cycle (and modifying either the design or the implementation) or to accept the prototype as a final version. Another noticeable point is that the early integration of the safety assessment process enables the design to be modified as a result of identified risk reduction recommendations.

Usability: the overhead of using this method in the overall process is quite low. For instance, in our first iteration of the cycle, apart from the design that was shared between stakeholders, the only critical path overhead that was induced by the meetings specific to the risk assessment process: two workshops of two hours each for the Preliminary Hazard Analysis and two sessions of two hours to present the outcomes of the UML-HAZOP Analysis and the assessed risks. The time devoted to the safety analysis itself (not to the critical path) consisted of the time to prepare guidelines for PHA workshops (a few hours), the time to carry out the UML-HAZOP analysis (2 weeks) and the time to format the results (a few weeks). UML-HAZOP already proved to be usable even by hand but would be even more so with a tool to assist book-keeping and result formatting.

Validity: Our approach relies on the UML-HAZOP method that identifies a large set of operational hazards. Other hazards, especially environmental hazards should be covered by PHA. Of course, all hazardous situations cannot be foreseen however much effort is put into safety analysis. Nevertheless, using these techniques, all classical hazards of robots were identified and several new hazards were discovered (e.g., incorrect position of the patient) together with the corresponding misuse or system failure. This gives use confidence that the coverage of UML-HAZOP and PHA are sufficient to identify most hazards of service robots.

Applicability: the first iteration of our risk assessment process led to conclusive evidence that the first prototype of the MIRAS robot was unsafe. The hazard list and the recommendations issued were accepted by the robotic experts in the MIRAS project. Our recommendations were taken into account in the design of the next prototype (which also includes several ergonomic changes).

V. CASE TOOL DESCRIPTION

Following two case studies where we use our method to analyse safety of service robots (PHRIENDS [14] and MIRAS [4] projects) we developed a Computer-Aided Software Engineering tool (c.f. Figure 5) to support the method, with the following motivating features:

- Support for UML modelling (Use Case and Sequence Diagrams)
- Partial automatic generation of HAZOP tables for systematic analysis
- Management of the combinatorial aspects of the HAZOP
- Guaranteed consistency between UML model and risk analysis tables
- Support for building a safety argumentation for risk acceptability
- Profiles allowing project-dependent configuration of severity levels, HAZOP table columns and guideword lists

The tool is built as an Eclipse plugin (www.eclipse.org) using the Graphical Modelling Framework (GMF). A first prototype has been released implementing most of the steps of our method. Users can draw and document UML use case and sequence diagrams. Based on configurable guideword listings and on UML models, HAZOP tables are partially filled. The user can then define severity levels, and enter all data required for the analysis. Final documents (Hazard List, Recommendations and Required Integrity Levels) are then extracted from HAZOP or PHA Tables, and traceability links are displayed to link hazards to causes. The tool is easy to use because of its simplicity and integration to a common environment (Eclipse). The method and guideword list can be adapted thanks to HAZOP Table templates. Furthermore, the analysis can be exported in CSV to reuse the results outside the tool. However, rich formats like HTML or Excel are not yet available for exportation, which currently limits the integration of our tool with other software. A second prototype is under development integrating visual improvements and a more user-friendly production of documentation.

VI. RELATED WORK

There have been several previous studies aimed at linking model-based development with risk analysis. For instance, in the CORAS project [15], [16], developed a framework, exploiting risk analysis and object oriented modelling concepts, for risk assessment of security critical systems. In our case we focus on safety and not on security, but the objectives of our study are quite similar to CORAS. Nevertheless, we do not have the same claims in terms of UML diagrams (we only focus on use case and sequence diagrams) and risk analysis techniques (we only focus on HAZOP and FTA). A major difference is that we strongly interconnect UML models and techniques such as HAZOP whereas that is not the case in CORAS. For instance, they use HAZOP without any real link with UML models (their HAZOP guidewords are not applicable on UML elements). Actually, they identify critical
assets and analyse for instance any deliberate/unintentional manipulation of this asset [17]. In safety, assets are not elements of the system but they are the system itself or system users. Hence, their approach is hardly applicable for safety analysis [16].

Our risk analysis approach is based on a re-interpretation of the guidewords for Hazard and Operability Studies (HAZOP) in the context of different UML models. The proposal in [18] followed by a more systematic study in [19], also considers a HAZOP guideword interpretation for the deviations of UML elements such as class, association, classifier role, message, etc. A similar approach was followed in [20] and [21], which also present a statistical analysis of the usability of this method. The guideword interpretation for the static UML diagrams in those studies aims to inspect the model to determine development faults and not to identify operational deviations. Nevertheless, for the UML dynamic diagrams (use case, sequence, activity, and statechart diagrams) many guideword interpretations can be used for exploring deviations during operational life. This is the case in studies presented in [22] and more formally in [23], which focus on use cases. The latter study led to a method that has been successfully used in [24] and [25].

This work on use cases also inspired a similar approach for security where new interpretations of guidewords have been proposed [26]. Even if this work is more oriented towards malicious behavior of actors, several interpretations can be applied in safety-critical systems with human-machine interactions. We combine and extend the results of those studies, but focus only on use case and sequence diagrams in order to explore deviations during operational life. We also give a particular attention to the integration of HAZOP-like human error analysis techniques as presented in [27]. Indeed, human factors methods [28] are a major issue in safety-critical systems but their analysis is often uncorrelated with preliminary system modeling activities. On the contrary, a key point of our approach is to consider human factors from the outset, by including them in model based risk analysis.

VII. CONCLUSION

To tackle safety of autonomous systems, appropriate analysis methods are needed especially when the system physically interacts with humans. Standard risk assessment methods are however limited to simple systems and usual model-based risk assessment methods do not enable to model dynamics. Thus we adapted the classical process with a new
model based approach for autonomous systems in physical contact with humans. We model the system using a subset of the well-known standard format UML. We apply PHA and our adaptation of HAZOP to identify hazards. A qualitative method is used to evaluate the risk on the preliminary design allowing the safety process to be integrated early in the development process. FTA is used to evaluate the risk on the other iterations of the safety process. A tool was developed to support the process.

We applied the process to the robot assistant developed in the MIRAS project. The first iteration of the safety process in that project confirmed the needs for high-level design analysis. Furthermore, starting the safety assessment process at the very first step of the design is helpful. In MIRAS, we obtained several results on the first iteration and the recommendations issued in that process enabled integration of safety constraints in the design of the second prototype. Assistant robotic is lacking of standard especially regarding the modeling of humans in the safety assessment process. We are currently checking the design for the second prototype with our partners in the MIRAS project and plan to apply a second iteration of the safety assessment process, including a quantitative risk estimation using FTA.

Acknowledgements

This work was partially supported by the MIRAS Research Project, funded under the TecSan (Technologies for Healthcare) program of the French National Research Agency (French ANR).

References