Safety Risk Analysis in Multi-Agent Scenario

Michael Kieviet [1], Padma Iyenghar [2,3]

[1] *innotec GmbH, 49324 Melle, Germany*
[2] *innotec GmbH, 49324 Melle, Germany*
[3] *University of Osnabruck , 49069 Osnabruck, Germany*

Abstract

Functional safety is established as a requirement for product, process, and machinery safety. These areas are always in change and transform from single island application to more networked complex systems. This paper should contribute an approach to transform also the functional safety aspect for the new circumstance. Based on agent and multiagent models a safety agent model is proposed and explained in an AGV (Automated Guided Vehicle) application.

Keywords

Multi-agent-Systems, Automated risk control, Functional safety

1. Introduction

With the increasing bandwidth in communication networks even in wireless networks with the sense of IoT[1], machineries become mobile and flexible. This kind of flexibility offers the possibility to collaborate between the machines, so that a set of single machines partly work together to solve a job and after this, they diverge or rebuild in another constellation[5][11]. This kind of flexibility has an impact on the process of risk analysis and their resulting definition of safety functions. The meaning of risk in this context is the probability of harms to get dangerous in the combination with the severity of harm even for human beings.

Nowadays it is common state of the art in the industry to define the hazards and the resulting risks by predicting the environment and nearly all-possible future scenarios where the machine will interact with humans and could become dangerous[3][6][7][9]. This concludes three important steps by doing the risk analysis. First there is a need to define the boarder of the machinery including all their interfaces. The next step will be to identify all possible hazards depending on the different life phases and usage modes. And the third step will be the estimation of risk. The last step is of course the most difficult, because there are very seldom real probabilities available. Usually, it will be done in an empirical way and estimated by experts with domain know-how. Several methods are available to have a systematic approach for estimating the risk for example RPN (Risk Priority Numbers), Risk graph, ALARP (As Low As Reasonably Practicable) and much more. Depending from the risk level[2] it is necessary to define the mitigation measures as:

- inherent safe construction
- risk reduction by safeguarding or implementation of complementary protective measures
- risk mitigation measures by safety management or behaviour rules.

Proceedings CERCIRAS, September 09–02, 2021, Novi Sad, Serbia
EMAIL: michael.kieviet@innotecsafety.com (A. 1); piyengha@uos.de(A. 2);
ORCID: 0000-0003-1581-4130 (A. 1); 0000-0002-1765-3695 (A. 2)

[2] Risk level: The right definition of the level is always in discussion, we using the deducted values from MEM-Rate (Minimum Endogenous Mortality) which are influenced in several international safety standards[3].

This order shows also the priority of mitigation measures and recommend, whenever possible, to eliminate the hazard or as well the risk by a safe construction. The third point can be done if technical issues are not possible or if the effort is in no relation to the severity of harm. The second item in this list is addressed under the term functional safety and requires active safety functionality. This can mean simple reactive control functions realised in discrete wired technique or much more sophisticated controlled by E/E/PE Systems.[6] Based on the identified risk level, manly defined in a probability of risk r(p), the availability of the safety function will set against the risk probability. So that the residual risk probability should be lower or equal as the accepted residual risk.

This approach is easily applicable for machineries with low flexibility or better with relatively good predictable use cases, configurations, job definitions and environmental conditions. A safety function including their belonging safety level can be specified for each individual risk case. And the mitigation aspect is only given for this predicted case. That mean in conclusion, derived from the risky event, the safety function has to be specified. From the specification of the safety function an adequate technical safety chain has to be required, so that the next level of requirements define the software and/or the hardware requirements on the functional level and, depending from the required safety level, also on the integrity level. Further deriving cares about the realization or the development of the hardware and the software including all verification measures which are required, depending on the necessary quality requirements according to applicable functional safety standards.

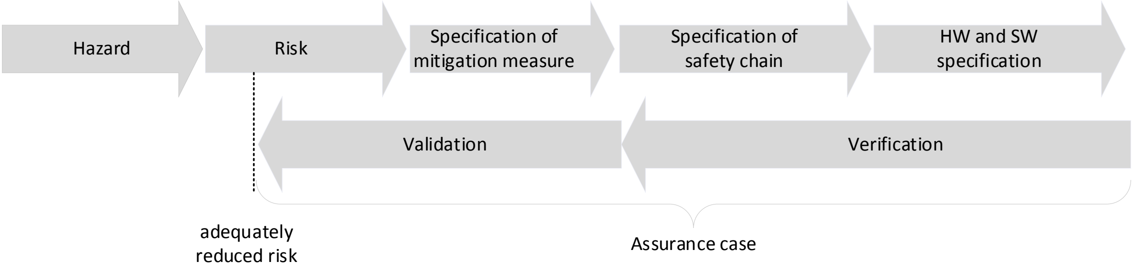

Figure 1: Safety development model[3][6]

This concept is based on the assumption that the system safety behaviour is reactive and static. The validation can be done by constructing reference test cases derived by the analysed use case. Outside the functional safety-world it becomes more and more popular to use AI controlled systems in the first step by imaging the surrounding (sensor) and in the next step by decision and controlling algorithms for perception and action. Such a solution provides additional safety functionalities which are declared as assistance systems but are currently not acceptable under functional safety criteria[2]. However, if such systems will be used in the functional safety domain it will be necessary to have new approaches to manage the hazards and risks.

Even the prediction of the risks is difficult for systems in a continuously changing environment, changing system behaviour and changing configuration.

The new approach should change the predicted risk analysis by a manual risk assessment to an ad-hoc automatic risk assessment done by the protagonist (agent) itself[4]. There are several advantages by doing this, e.g. the agent is able to adapt his safety behaviour depending on the environmental situation or the required task. It can also mean that the safety objective could change depending on the situation[2]. In current applications such an action requires a prediction of possible risks and a predefined safety function which can be activated by mode-selectors if needed. In future systems it could be possible to adapt the safety functions according to the situation. This would allow to justify or modify the safety function depending on the mission for the system. In a multi agent system (MAS) it will become a challenge to keep the risk below the acceptable level for the safety objectives.

That concludes that a safety behaviour of safety agents is depending on

- safety objectives / goals
- safety performance (implemented basic functions)
- safety adjusters (adaptive control parameter)
- safety risk mitigation capability (ASIL, SIL, PL)

and

- non-safety mission objectives

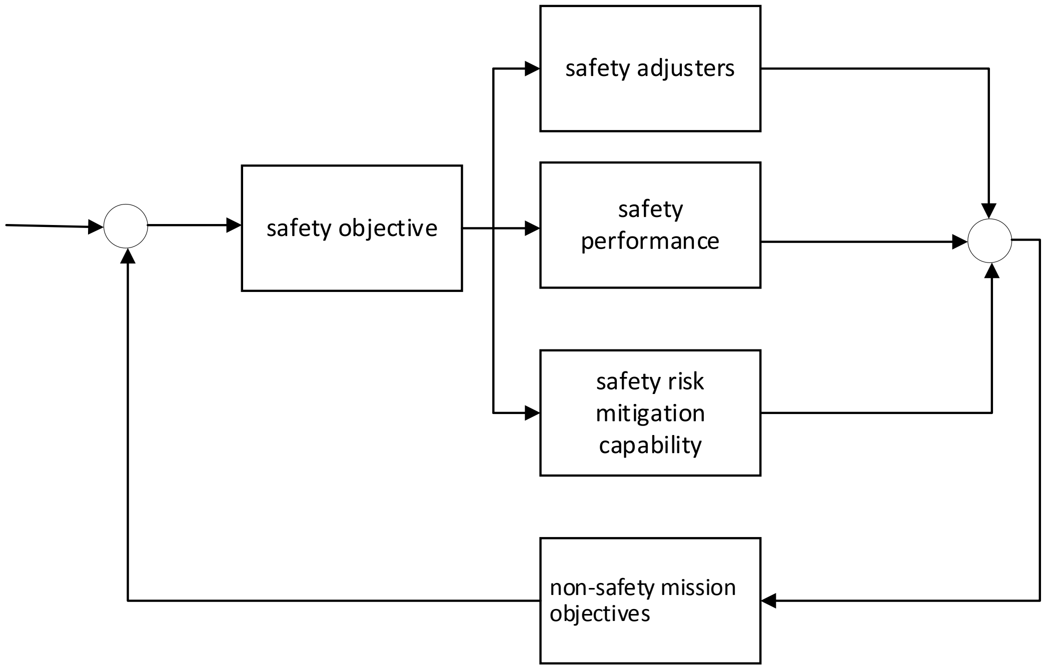
safety adjusters
safety objective
safety performance
safety risk mitigation capability
non-safety mission objectives

Figure 2: Safety agent model

Safety objectives represent the safety goals in a manner that the agent shall not injure a human being, avoiding material damages or protect against environmental pollutions.

Non-Safety mission objectives shall define the job definition of an agent, this is mainly the functionality which can be refined with attributes like availability, effort etc.

Safety performance defines the safeguarding and complementary protective measures e.g. (STO, SS1; SS2, SOS, SLS, SLI)[8].

Safety adjusters are the parameters in a safety controller as delay time of SS1 and SS2, or speed limit value in SLS[8].

Safety risk mitigation capability defines the safety metrics like degree hardware fault tolerance (HFT) safe failure fractions (SFF), internal diagnosis (DC), probability of dangerous failure per hour/per demand (PFD/PFH)[6][9].

The safety risk mitigation capability is usually determined for a system inside system boarders and considers the safety functions. An agent could have a set of independent safety functions. And a combination of safety functions in a multi agent environment has an impact on the achievable mitigation level for a MAS defined safety objective even if the safety agents combine their safety performance to fulfil their non-safety objectives.

A combination of safety functions between several agents (MAS) requires a cross communication possibility and an information exchange to negotiate the safety adjusters[11].

This paper introduces just a first proposal for such a model view and will be a beginning of a deeper research and implementation concept.

2. Application example

The following chapter should show an example of AGV (automated guided vehicle)[10]. Such an AGV has lots of safety functions included. Not all of them are responsible to mitigate the risk of the same hazard. One safety objective is not to hit a person during moving (avoid body contact with a critical level of moving energy). An independent different safety objective for example could be to avoid the destruction the Li-Ion battery, because it could also become dangerous.

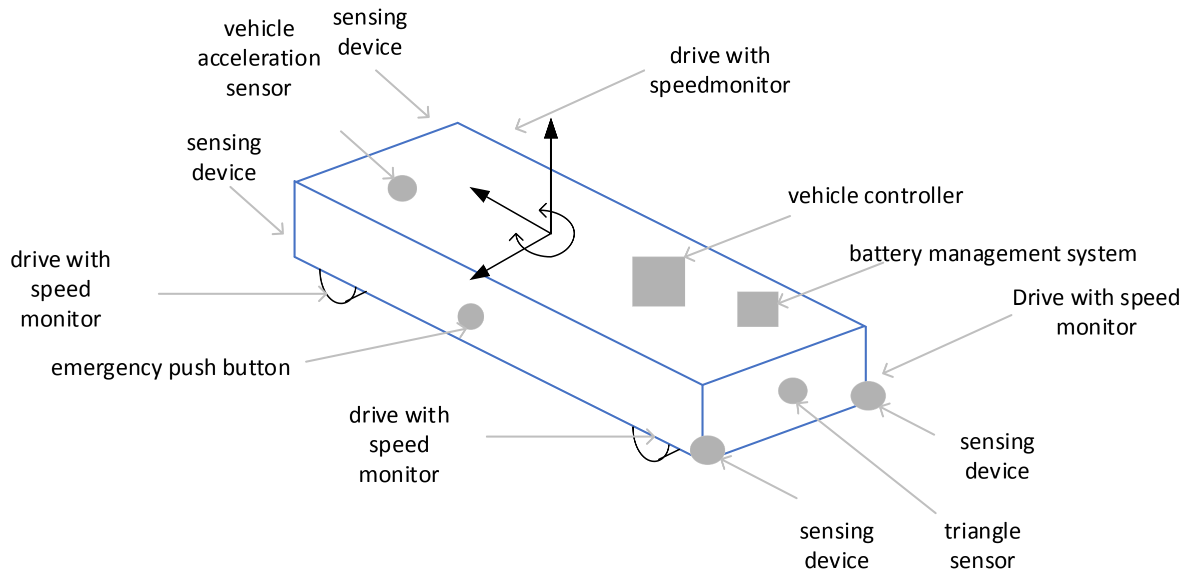

Figure 3: Safety agent model

$$AG1=SF\{sf1, sf2,\ldots sfn\} \qquad (1)$$

In the current example we would reduce the view only to the movement risks.

$$AG1=SF\{sf1\} \qquad (2)$$

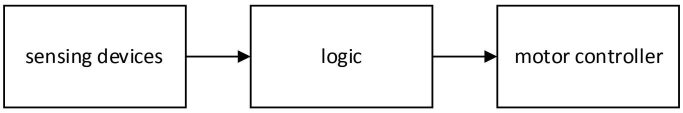

Figure 4: Safety chain of one separated safety function (sf) of a safety agent (AG).

We select an easy assumption for the safety function. If the sensor device detects an obstacle the logic will decide to reduce the speed and the motor controller monitors not to exceed the speed limit. Based on the safety agent model of figure 2 it has the following definition:

Safety objective: Avoid hitting the obstacle.
Non-safety objective: keep on moving if possible
Safety adjuster: adapt the speed parameter or the trajectory (we concentrate to the speed)
Safety performance: Safely limited speed (SLS)[8]
Safety risk mitigation capability: SIL 3, PFH= 20 FIT, SFF>90%, HFT=1

Now it will be possible that agents shall co-operate in a scenario and that they get new non-safety objective, for example: agent 1 and agent 2 shall frontal move close together as fast as possible. The safety objective is: without to box in somebody. That mean, the speed limit between agents must be negotiated.

The safety performances have to be compared. If both agents provide the same performance, it is simple to divide the speed and to set the safety adjuster. If the agents provide different performance, it will be necessary to negotiate the adjuster values depending on the super positioning of the two unequal safety functions. For example, agent one performs SOS and agent 2 only performs SS1.

Additional safety performance needs to be considered in both agents, because the number of elements which are known in the responsibility to mitigate the risk of the new safety objective is now increased with two agents. This constellation can happen regarding the rules of modelling safety metrics for example with RBDs (reliability block diagrams). In the worst case all safety functions of the agents are in a chain. That mean in the worst case for the

$$\text{PFH}=\sum PFH_{Agent\ n} \qquad (3)$$

and this has to be below the limits of required safety level for the given safety objective.

3. Conclusion

We saw that the current way to manage functional safety risk analysis in dynamic complex systems are not practicable regarding the need to foreseeability of risk scenarios. In the next step we propose to change the view from a static analysis to a safety agent model. Based on this model approach we will show for a simple example how this can be realized.

In further steps we want to refine the model and validate the possibilities of the negotiation processes. So that such systems can be programmed and approved on the agent level and that they can do their self-validation process on the multi agent level.

4. References

[1] S. K. Sharma, I. Woungang, A. Anpalagan, S. Chatzinotas, Toward Tactile Internet in Beyond 5G Era: Recent Advances, Current Issues, and Future Directions, IEEE Access 8 (2020)

[2] ISO, ISO/PAS 21448, Road vehicles – Safety of the intended functionality, Vernier, Geneva, (2019)

[3] DIN, DIN EN 50126, Railway applications – The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS), German version EN 50126:1999

[4] M. Wooldrige, An Introduction to MultiAgent Systems: 2nd Edition, WILEY & Sons Ltd, West Sussex, United Kingdom, (2009)

[5] P.G. Balaji, An Introduction to Multi-Agent Systems. July 2010, DOI:10.1007/978-3-642-14435-6_1, URL:https://www.researchgate.net/publication/226165258_An_Introduction_to_Multi-Agent_Systems , Springer-Verlag, Berlin, Heidelberg Germany, (2010)

[6] IEC, IEC 61508 Functional safety of electrical/electronic/programmable electronic safety-related systems, Geneva, Switzerland, (2010)

[7] ISO, ISO 12100, Safety of machinery-General principles for design- Risk assessment and risk reduction, Vernier, Geneva, (2010)

[8] IEC, IEC 61800-5-1 Adjustable speed electrical power drive systems – Part 5-2: Safety requirements – Functional, Geneva, Switzerland, (2016)

[9] ISO, ISO/DIS 13849-1.2, Safety of machinery-Safety-related parts of control systems, Vernier, Geneva, (2021)

[10] VDI, Fahrerlose Transportsysteme Leitfaden Sicherheit, VDI Statusreport,(2020)

[11] SmartFactory^{KL}, Safety an modularen Maschinen, Whitepaper SF-3.1: Technologie-Initative Smart Factorie KL e.V., Kaiserslautern, Germany, (2008)