2.2.4.3 – Verification of "deciding" requirements

Practical guidance – cross-domain

Authors: Dr Mario Gleirscher and Dr Radu Calinescu, University of York

This Body of Knowledge entry provides a guide to the analysis, design, synthesis, and assurance of safety controllers for use in *human-robot collaboration (HRC)* settings.[1] This guidance has been developed in the context of the manufacturing domain, however it is envisaged that the guidance would be applicable to other domains with similar characteristics.

An HRC setting typically comprises a mobile or stationary robot (the *Cobot*) collaborating with one or more human operators on *shared repetitive tasks (the Process)*. The goal is to combine the capabilities of humans and machines in order to improve quality and reduce cost. As an example of HRC in manufacturing, Figure 1 depicts the actors (blue), the important geometric features (boxes and arcs), and sensor-tracked safeguarded areas (red) of a manufacturing work cell. This work cell concept is inspired by a real-world setting in an actual industrial manufacturing company.

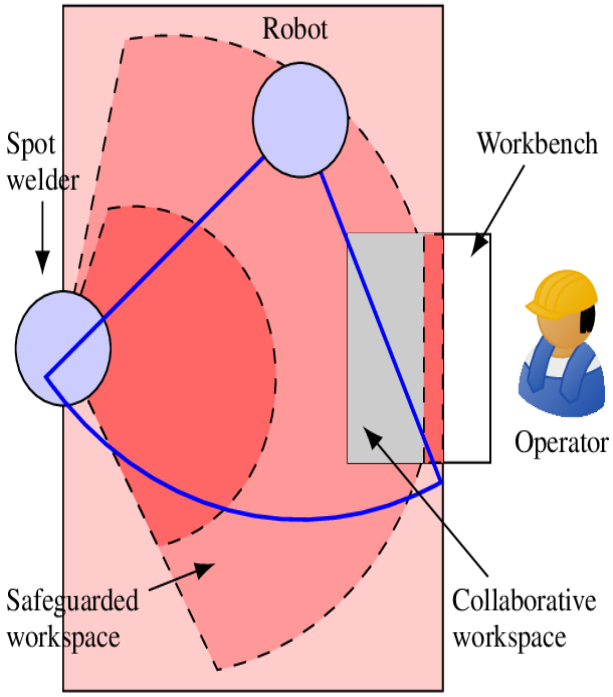

Figure 1: Exemplary HRC setting

An *automatic safety controller (ASC)* is used to improve occupational safety for the tasks performed in this work cell. The controller is responsible for handling *critical events (CEs)*. For example, when an operator enters the work cell if they are not supposed to then the ASC would (a) recognise this through the tracking system of the cell, (b) switch the Cobot and other machine elements into a contextually appropriate *safety mode* [1], (c) interact with the operator, and, (d) if the CE is mitigated, switch the work cell back to

1 This BoK guidance is a result of the AAIP demonstrator project CSI:Cobot

a mode where the obstructed task can be continued. The steps (a) to (d) suggest a decomposition of the system-level safety requirement derived from *hazard analysis and risk assessment* (HARA) of the HRC setting (see Objective 1.3).

The ASC is a critical Deciding element and, as such, a part of the overall Deciding element of the RAS architecture used in the HRC setting. A HARA of the HRC setting usually results in a list of *safety requirements* specifying and driving the design of the controller. This BoK entry focuses on the design of *discrete-event controllers* that can perform the steps (a) to (d) and it considers the main assurance evidence to be delivered by *probabilistic model checking*. This guidance is for verification engineers that want to model an HRC setting (i.e. actors, activities, actions) focusing the controller behaviour and verify this behaviour against safety requirements.

Stage Input and Output Artefacts

To follow the proposed approach, the following information is needed:

- A technical description of the HRC setting (i.e. the work areas, the shared tasks, the actors) (see Objective 1.1.2).
- The results of a (preliminary) HARA (i.e. a hazard log and a requirements list) (see Objective 1.1).

The proposed approach makes further *assumptions*:

i. The sensing and tracking system dependably captures the Deciding information sent to the controller (e.g. the occupancy of safeguarded areas) (see Objective 2.2.4.2)

ii. The environment (e.g. the trained operator) behaves within reasonable limits (e.g. no outrageous or miraculous operator behaviour)

iii. The control signals to the Process are accurately implemented by the actuators (e.g. a stop signal to the Cobot is reacted to within milliseconds) (see Objective 2.2.4.5)

A successful application of the proposed approach will result in:

1. A behavioural *model* of the discrete-event safety controller capable of monitoring the listed hazards and responding to any occurrence of these hazards with appropriate actions (e.g. safety mode switches) to mitigate risks from these hazards.
2. A *formalisation* of the requirements suitable to the vocabulary of the controller model (1.).
3. *Verification evidence* that the controller model (1.) fulfils the requirements (2.) under the assumptions stated above.

Activities/Procedure

The work steps of the approach are depicted in Figure 2:

(1) *Process Modelling:* This step aims at capturing the actors (i.e. Cobot, operator, other machines) working on the shared tasks in the HRC setting in a behavioural model, particularly, a *Markov decision process* (MDP).

(2) *Safety Analysis:* The goal of this step is to identify and model relevant hazards and to specify the ASC requirements.

(3) *Mitigation Design:* In this step, the design of the safety controller is derived from the controller requirements and added to the process model from step (1). The result is a stochastic model of the HRC setting (e.g. a welding machine, a robot arm, a human operator) including the safety controller, particularly, the behaviours for the steps (a) to (d) from above.

(4) *Verified Controller Synthesis:* The extended model (3) is parametrised (e.g. situational probabilities, mitigation options with different properties) to form a design space one can select a controller from that optimises criteria expressed over these parameters. This step combines the verification of ASC requirements (2) with the selection of an optimal ASC. Further more specific requirements can then be verified of the selected ASC. The controller can then be tested in the demonstrator for validation (e.g. to verify the aforementioned assumptions).

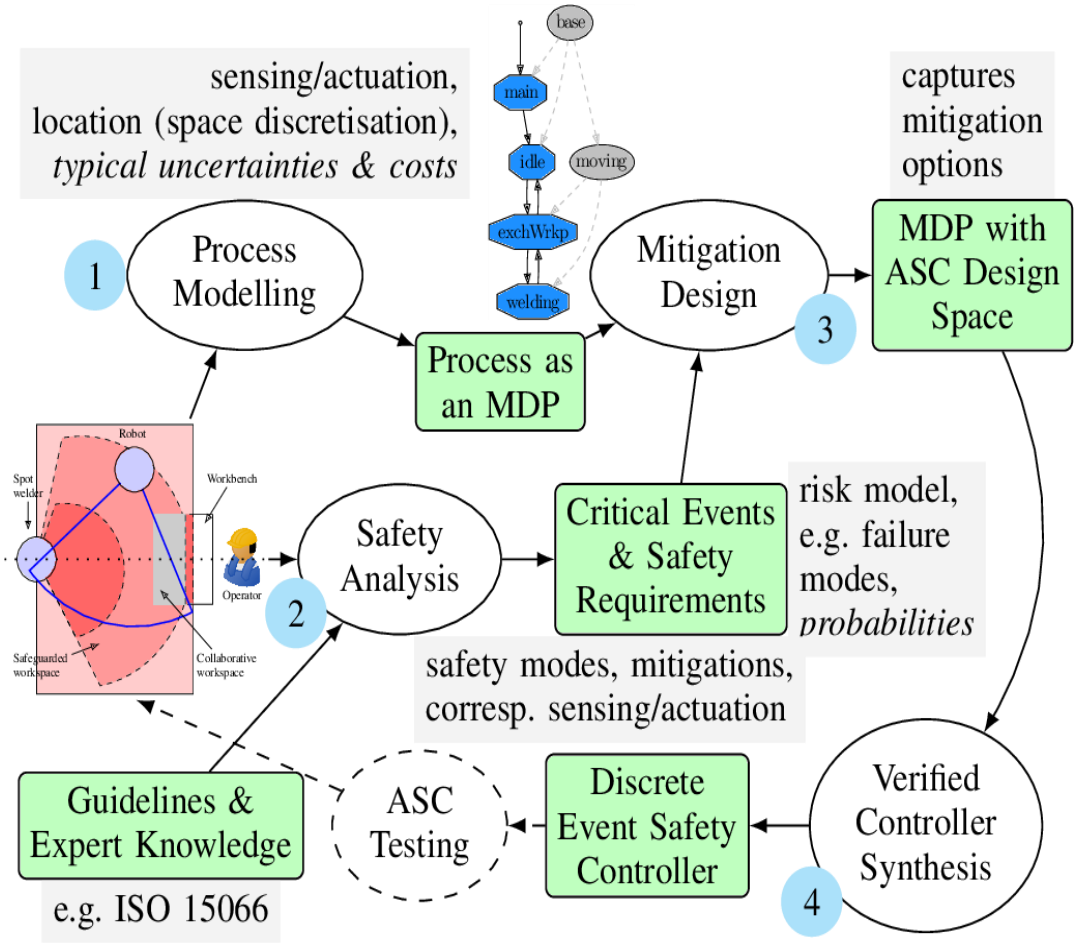

Figure 2: Overview of the proposed approach

Desired Assurance Properties

Examples of controller requirements verifiable in the formal model of the Process:

- If a CE is detected, the controller correctly carries out an appropriate mitigation procedure.
- The controller maximizes the probability of accident freedom under probabilistic unexpected operator behaviour, that is, behaviour of the operator deviating in a random way from what is expected of them.
- The controller minimises the probability of unnecessary task obstructions.

- The Process model is well-formed (e.g. nominal behaviours finish the HRC task). Well-formedness strengthens verification evidence and can simplify model debugging.

Methods

- The proposed approach is *agnostic of the chosen HARA* approach as long as the result is a hazard log and the corresponding safety requirements for those hazards are to be addressed by an automatic controller. (This should be true of most HARA approaches.)
- The tool YAP [2] is recommended to be used to bridge the gap between the hazard log, risk modelling, and controller design.
- The PRISM [3] stochastic model checker is recommended to be used for design space verification, exploration, and optimal discrete-event controller synthesis.

Further Details about the Approach

The first results of the proposed approach are published in [4], including an evaluation of several Process models derived from hazard logs of increasing size (up to seven concurrent hazards). A hands-on guide to the proposed approach is provided in [2].

Advantages of the Approach

- Flexible abstraction: With stochastic models, one can capture probabilistic or uncertain phenomena, such as sensor faults, human errors, and actuator perturbations.
- Exhaustive requirements validation: Model checking allows full model exploration for reasonably small models.
- Rapid rigorous prototyping.

Challenges and Limitations of the Approach

- The proposed approach requires both experience in probabilistic model checking and in abstracting the application domain into the corresponding modelling and property languages. Strong abstraction is necessary to keep the model practically small.
- The controller resulting from the synthesis step is symbolic and following a "detection → response" or "if X then Y" rule pattern. Detailing this pattern by corresponding monitoring predicates and response or mitigation actions is necessary but can be automated. The controller can then be tested and validated in the actual HRC setting.

References

- [1] ISO/TS 15066, Robots and robotic devices – collaborative robots, Robotic Industries Association (RIA), Standard, 2016.
- [2] Gleirscher, M.: Yap: Tool Support for Deriving Safety Controllers from Hazard Analysis and Risk Assessments. Luckuck, M. & Farrell, M. (Eds.) *Formal Methods for Autonomous Systems (FMAS), 2nd Workshop*, *EPTCS, 329, pp. 3*1-47, 2020. doi:10.4204/EPTCS.329.4 .

- [3] See www.prismmodelchecker.org
- [4] Gleirscher, M. & Calinescu, R.: Safety Controller Synthesis for Collaborative Robots. *Engineering of Complex Computer Systems, 25th International Conference, 28 - 31 October 2020, Singapore,* 2020. https://arxiv.org/abs/2007.03340 .