Rule-based AI in Model-based Emergency Control

Avi Harel, Ergolight

ergolight@gmail.com

Challenges of emergency control

This presentation demonstrates the potential benefit of integrating AI processes in the design of emergency control. It is based on several Loss Of Control (LOC) case studies attributed to situational confusion. The study addresses the following engineering challenges:

1. A model of emergency control
2. A framework for Integrating AI for effective control
3. Generic rules enabling affordable implementation

Human-machine teaming

According to the traditional approach, each of the tasks should be allocated to either the human or the machine, by matching tasks to the strengths of humans and machines.

A key topic is the control dilemma: who should take the lead? In the AF 296 accident the automation took the lead, disabling the human operator. On the other hand, in the AF 447 accident the machine pushed the lead to the operators, who failed.

In emergency control, we deal with unexpected situations. Currently, automation cannot handle the unexpected. Apparently, this limitation is the single most important reason why we introduce human controllers in the system in the first place. Ironically, however, humans cannot handle unexpected situations under stress: in emergency situations, the human operators react according to their training, which is suites normal operation (Bainbridge, 1982, The Irony of Automation). The solution to this problem is by human machine collaboration in the emergency control. The design challenge is to define the way they should collaborate. Can the machine support the human needs to enable effective control?

Traditional machine support is sufficient for some of the tasks; yet other key control tasks require incorporating reasoning in the machine.

A model of emergency control

The model of emergency control describes a special case of a general model of system operation. The focus here is on avoiding emergencies and supporting the human operator when under stress. However, this same model may also describe emergency operation of Systems of Systems (SOS), in which one of the subsystems controls the behavior of the other subsystems.

An emergency may be defined as an exceptional situation in which the controller is under threat. It begins in a change from a normal to a risky situation. It is terminated by resuming the normal situation. The design goals are to shorten the emergency, to eliminate the risks of operating under threat, to prevent operator errors, and to support resilient operation. The model presents eight tasks in emergency control:

1. Detecting a risk
2. Hazard identification
3. Informing the human operator(s) about the hazard
4. Assessing the hazard risks
5. Proposing optional reactions
6. Evaluation of the optional reactions
7. Selecting a best option
8. Executing the selected option

In each task, the goal is to maximize efficiency and safety. These goals suggest a need for decision support, and for enforcing operation by rules.

Integrating AI in emergency control

AI processes may provide both the automation and the human operators with the reasoning required for protecting each of the tasks above.

The key to the collaboration design is the need to define the rules for constraining the operation, to eliminate emergency risks. Here are key AI activities in the tasks of the control model:

1. Detecting a risk: automation support, by risk indicators
2. Hazard identification: automated troubleshooting
3. Informing: human situation awareness, based on HCD principles
4. Assessing the hazard risks: providing preview information, by trend analysis
5. Proposing optional reactions (option generation)

6. Evaluation of the optional reactions (preview, by simulation)
7. Selecting a best option (Avoid potential mode error)
8. Executing the selected option (Setting proper conditions).

Implementation

The implementation of the collaboration is within the scope of integration engineering. The goal of integration engineering is to ensure that individual parts work together seamlessly to achieve a desired outcome. In the context of system operation, the individual parts are the human operator and the controlled machine. Normal operation may be specified in terms of project specific operational rules.

The engineering challenge is to facilitate system integration, to allow all projects to incorporate the design guidelines. To facilitate the design and testing, common operational rules may be expressed as generic rules, which may be customized to different kinds of emergencies. Here is a list of generic rules applicable to emergency control:

- Implementing scenario-based operation
- Diversion detection, on exception: automated detection and alerting
- Hazard detection, rebounding, and alerting,
- Decision support: situation preview, option evaluation, implemented by behavioral twins. These are digital twins optimized to support the human activities.

Vision

The history of accidents is saturated with examples of accidents that could have been prevented, had the industry found the ways to protect from human errors and to apply the investigation findings across different domains.

To protect from human errors, we need to develop affordable methods to oppose accountability biasing, to constrain operation by the rules, and to detect and alert about exceptions.

For cross-domain learning, we need to create a cross domain ontology, comprising standards that formulate the generic rules, and to enforce employing them by regulation.

References

Harel, A. 2024 (A) - Sensor Integration Verification, DOI: 10.13140/RG.2.2.27400.23044
Harel, A. 2024 (B) - Enforcing feature availability, DOI: 10.13140/RG.2.2.31237.36326
Harel, A. 2024 (C) - Configuration verification, DOI: 10.13140/RG.2.2.32397.35040