Engineering Risk Analysis Group
Prof. Dr Daniel Straub

Who saves the system? Learning to negotiate responsibility in complex engineering systems via multi-agent reinforcement learning

Motivation and Objectives

In engineering risk analysis, redundancy is a central design principle for ensuring safety and reliability. Whether through simple k-out-of-n logic, complex topological connectivity, or functional overlap, redundancy allows systems to tolerate local failures without catastrophic consequences [1]. However, in Multi-Agent Reinforcement Learning (MARL) [2], redundancy presents a fundamental paradox.

While engineers design redundancy for robustness, reinforcement learning agents often view it as an opportunity for "free-riding". In a redundant system, where the successful operation of the whole does not strictly require the effort of every individual, agents often face the "lazy agent" problem [3]. The ambiguity in credit assignment (i.e., who was responsible for the success?) can lead to policies where agents rely excessively on their peers, creating "brittle" cooperation that collapses when the system is stressed [4].

The objective of this MSc thesis is to investigate how different forms of redundancy influence cooperative behavior in MARL under a centralized-training–decentralized-execution (CTDE) paradigm. The specific goals are to:

- Analyze the impact of redundancy: Investigate how varying degrees and types of redundancy (e.g., voting systems vs. structural load-sharing) affect convergence, credit assignment, and the emergence of free-riding behaviors.
- Study failure modes: Identify "hidden risks" where agents learn policies that appear successful during normal operation but lack the necessary coordination to handle component failures.
- Develop communication strategies: Propose and test communication mechanisms that allow agents to "negotiate responsibility," ensuring that redundancy is preserved as a safety buffer rather than exploited for cost minimization.

Methodology

The project will follow a systematic approach:

- Literature review: Review the intersection of MARL (e.g., CTDE methods like QMIX or MAPPO), reliability engineering, and cooperative game theory.
- Problem formulation and scoping: The student will select a specific redundancy paradigm to investigate. Options include, but are not limited to:
  - K-out-of-n systems: Voting logic or parallel reliability blocks.
  - Topological redundancy: Graph-based connectivity (e.g., power grid or logistics networks).
  - Structural redundancy: Load-sharing systems (e.g., trusses or shared resources).
- Environment implementation: Develop a stylized multi-agent environment that captures the chosen redundancy dynamics and allows for tunable "criticality" (how essential an agent is to the system).
- Baseline analysis: Train standard MARL baselines without communication to quantify the extent of the credit assignment/free-riding problem.
- Communication and robustness (optional): Implement communication protocols (e.g., shared signals, learned messages, or bottlenecked attention) and evaluate if they enable agents to maintain robust cooperation in the face of simulated component failures.

Requirements

- Background: Solid understanding of probability, stochastic processes, or control theory.
- Technical skills: Proficiency in Python; familiarity and/or interest in deep reinforcement learning frameworks (implemented in PyTorch/JAX).
- Interests: Curiosity for the interface between machine learning (optimization/learning dynamics) and engineering (risk and reliability).

Starting date

Flexible.

Supervised by

This work will be supervised by Pablo G. Morato (Engineering Risk Analysis Group, TUM). If you are interested in this topic, please send an email to pablo.morato@tum.de .

References

[1] Albrecht, S.V., Christianos, F. and Schäfer, L., 2024. *Multi-agent reinforcement learning: Foundations and modern approaches*. MIT Press.

[2] Zio, E., 2007. *An introduction to the basics of reliability and risk analysis* (Vol. 13). World scientific.

[3] Sunehag, P., Lever, G., Gruslys, A., Czarnecki, W.M., Zambaldi, V., Jaderberg, M., Lanctot, M., Sonnerat, N., Leibo, J.Z., Tuyls, K. and Graepel, T., 2017. Value-decomposition networks for cooperative multi-agent learning. *arXiv preprint arXiv:1706.05296*.

[4] Vinitsky, E., Du, Y., Parvate, K., Jang, K., Abbeel, P. and Bayen, A., 2020. Robust reinforcement learning using adversarial populations. *arXiv preprint arXiv:2008.01825*.