Mini Project - Learning from Human Feedback

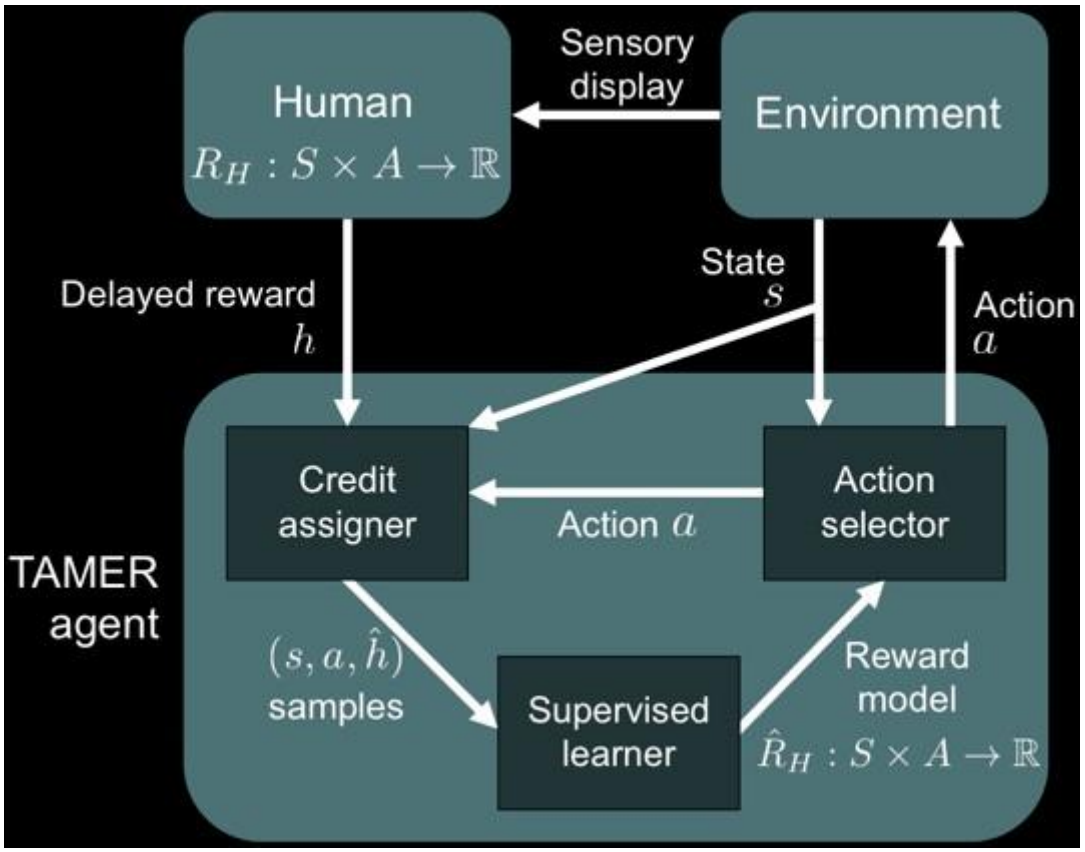
Assistant Professor: Silvia Tulli

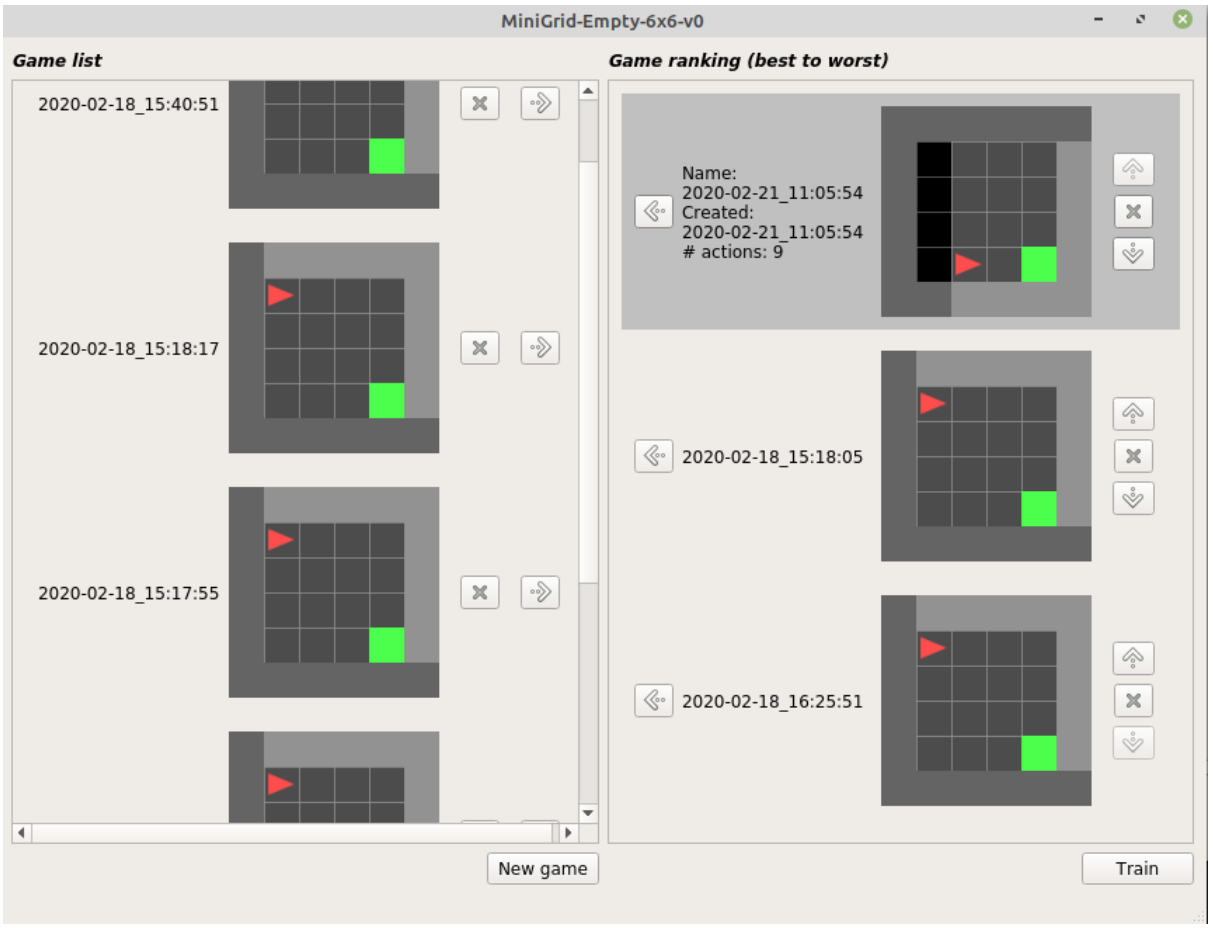
Course of Social Robotics Coordinated by Mohamed Chetouani

Several approaches have been used to integrate human feedback into machine learning (Argal et al. 2009, Christiano et al. 2017, Sumers et al. 2020, Caselles-Dupré et al. 2022, Celemin et al. 2022). The objective of the micro-project is to obtain **better performance** with a human-in-the-loop approach. You can use any learning algorithm as a baseline to evaluate the integration of human feedback.

You can employ existing environments from openAI gym or linked libraries, such as gym-minigrid, LunaLander, and test different types of human feedback (e.g., evaluative, imperative, and descriptive) to train your agent in the selected environment.

The way you incorporate the feedback of the human can vary from keyboard commands to speech, text, and images. Your project will be evaluated in terms of the performance obtained with respect to the adopted baseline. The complexity of the environment and integrated feedback will also be considered.





1. Do a benchmark of similar existing projects/approaches (4/20)
2. Choose one of the existing projects and try to replicate the results obtained by the authors (4/20)
3. Define your alternative approaches (at least 2) and modify the existing project allowing for the integration of human feedback (4/20) - first delivery
4. Plot the results of a baseline algorithm compared to the human-in-the-loop approach you implemented (4/20)
5. Discuss your methods and results, highlighting the advantages and limitations of your approach. (4/20) - second delivery

Format

- Your code should be uploaded on GitHub and include the necessary documentation to allow others to run it on other machines (i.e., requirements for setting up the environment etc.).
- A report of 6 pages maximum should be provided. Please find here the template for the report.
- The report should include five sections, each detailing the points mentioned above.