Towards an Experimentation Platform for Hybrid Human-AI Sequential Decision-Making

Hüseyin AYDIN [a,1], Kevin GODIN-DUBOIS [b 1], Libio GONCALVEZ BRAZ [a 1], Floris DEN HENGST [b,2 1], Kim BARAKA [b], Mustafa Mert ÇELIKOK [c], Andreas SAUTER [b], Shihan WANG [a] and Frans A. OLIEHOEK [c]

[a] *Utrecht University*
[b] *Vrije Universiteit Amsterdam*
[c] *TU Delft*

Abstract. We present SHARPIE (**S**hared **H**uman-**A**I **R**einforcement Learning **P**latform for **I**nteractive **E**xperiments), a generic framework to support experiments with RL agents and humans. It consists of a versatile wrapper for RL environments and algorithm libraries, a participant-facing web interface, logging utilities, and deployment on popular cloud and participant recruitment platforms. It empowers researchers to study a wide variety of research questions related to the interaction between humans and RL agents and aims to standardize the field of study on RL in human contexts.

Keywords. reinforcement learning, hybrid AI, experimentation, software

1. Introduction

Reinforcement learning (RL) refers to a family of algorithms in which agents learn from the interactions with an environment to maximize the long-term reward obtained by making sequential decisions [1]. RL agents interact with humans in a wide variety of ways, highlighting the need to incorporate humans in the training and evaluation of RL agents [2]. Currently, these interaction patterns remain largely limited and unidirectional. In contrast, hybrid human-AI intelligence (HI), as envisioned by [3] emphasizes the need for rich, diverse, and dynamic interaction patterns between humans and RL agents in order to effectively address problems that neither humans nor agents can solve independently.

Popular RL libraries provide a framework to train RL agents in established benchmark and novel training domains [4,5,6,7,8,9,10]. These libraries, however, do not offer rich, diverse and dynamic interactions among humans and RL agents such as multimodal communications for shared observations, delegated actions and goals that we believe to be required for HI.

[1] Equal contribution
[2] Corresponding Author: f.den.hengst@vu.nl.

We present SHARPIE (**S**hared **H**uman-**AI** **R**einforcement Learning **P**latform for **I**nteractive **E**xperiments), a platform for studying how multiple humans and RL agents interact and collaborate effectively. The platform [11]: (i) provides a versatile wrapper for (multi-agent, and multi-objective) popular RL environments and algorithms, (ii) supports configurable communication channels between the human and the RL agents with various modalities, and (iii) offers logging services, deployment utilities, and participant recruitment platforms integration.

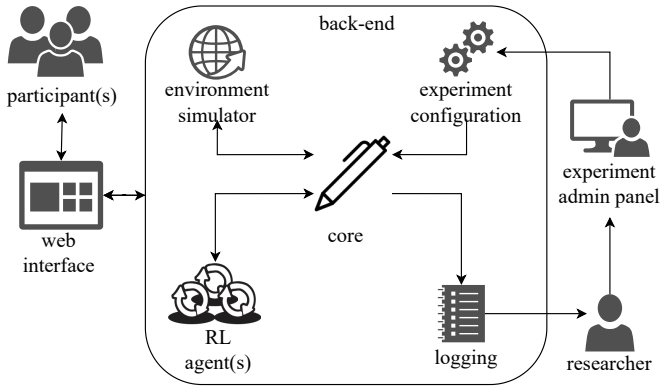

Figure 1.: High-level SHARPIE architecture.

SHARPIE can be used to facilitate experiments between RL agent and humans, and aims to introduce a standard for RL-based human-agent interactions in multi-agent settings.

2. The SHARPIE Framework

The SHARPIE library[3] is a Python-based web framework, and can encapsulate any existing environment that implements the Gymnasium API [4,5] which encompasses most of the existing RL platforms to date.

The front-end provides various utilities to aid experimental processes: from (a)synchronous evaluations of a learning agent to scheduling and management on long-term data storage. SHARPIE is designed to support multimodal communication channels to allow bidirectional communication between (RL or human) agents. These channels may be used, for example, for coordination, teaching, or to override actions in scenarios with multiple humans and multiple RL agents [12]. While initially designed for short semantic content and text, such as predefined sets of symbols, these communication mediums will be expanded upon to include any relevant format (video, audio, etc.). Finally, the library contains utilities for deploying to a cloud server, or a private (local) machine.

We present an initial implementation of SHARPIE and a selection of illustrative use-cases, with which, we aim to provide an easily integrable framework that researchers can use to painlessly set up experiments involving both human and AI agents to solve sequential decision-making tasks. Our hope is that in turn such an architecture lays the foundation for a standard for the interaction between human and such agents.

We plan to expand the scope of possible human-agent interactions by incorporating additional modalities [13,14]. Finally, we envision a hosted version of SHARPIE that can be used for outreach, education, and user literacy purposes.

[3] `https://github.com/libgoncalv/SHARPIE` and `https://youtu.be/J9iF7K-MGkM`

References

[1] Sutton RS, Barto AG. Reinforcement Learning: An Introduction. MIT Press; 1998.

[2] Den Hengst F, Grua EM, el Hassouni A, Hoogendoorn M. Reinforcement learning for personalization: A systematic literature review. Data Science. 2020;3(2):107-47.

[3] Akata Z, Balliet D, De Rijke M, Dignum F, Dignum V, Eiben G, et al. A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. Computer. 2020;53(8):18-28.

[4] Brockman G. OpenAI Gym. arXiv preprint arXiv:160601540. 2016.

[5] Towers M, Kwiatkowski A, Terry J, Balis JU, De Cola G, Deleu T, et al. Gymnasium: A standard interface for reinforcement learning environments. arXiv preprint arXiv:240717032. 2024.

[6] Terry J, Black B, Grammel N, Jayakumar M, Hari A, Sullivan R, et al. Pettingzoo: Gym for multi-agent reinforcement learning. Advances in Neural Information Processing Systems. 2021;34:15032-43.

[7] Rutherford A, Ellis B, Gallici M, Cook J, Lupu A, Ingvarsson G, et al. JaxMARL: Multi-Agent RL Environments and Algorithms in JAX. In: Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems. AAMAS '24. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems; 2024. p. 2444–2446.

[8] Alegre LN, Felten F, Talbi EG, Danoy G, Nowé A, Bazzan AL, et al. MO-Gym: A library of multi-objective reinforcement learning environments. In: Proceedings of the 34th Benelux Conference on Artificial Intelligence BNAIC/Benelearn. vol. 2022; 2022. p. 2.

[9] Taylor ME, Nissen N, Wang Y, Navidi N. Improving reinforcement learning with human assistance: an argument for human subject studies with HIPPO Gym. Neural Computing and Applications. 2023;35(32):23429-39.

[10] Godin-Dubois K, Miras K, Kononova AV. AMaze: A Benchmark Generator for Sighted Maze-Navigating Agents. Journal of Open Source Software. 2024:in press.

[11] Aydın H, Godin-Dubois K, Goncalves Braz L, den Hengst F, Baraka K, Çelikok MM, et al. SHARPIE: A Modular Framework for Reinforcement Learning and Human-AI Interaction Experiments. In: AAAI Bridge Program Workshop on Collaborative AI and Modeling of Humans. Philadelphia, Pennsylvania, USA; 2025. .

[12] Zhu C, Dastani M, Wang S. A survey of multi-agent deep reinforcement learning with communication. Autonomous Agents and Multi-Agent Systems. 2024;38(1):4.

[13] Christofi K, Baraka K. Uncovering Patterns in Humans that Teach Robots through Demonstrations and Feedback. In: Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction; 2024. p. 332-6.

[14] Knierim M, Jain S, Aydoğan MH, Mitra KD, Desai K, Saran A, et al. Leveraging Prosody as an Informative Teaching Signal for Agent Learning: Exploratory Studies and Algorithmic Implications. In: Proceedings of the 26th International Conference on Multimodal Interaction; 2024. p. 95-123.