CASAC Chair Internship Proposal:
Benchmarking DRL-based agents during a human-AI Collaborative task

Context

Since Alan Turing's introduction of artificial intelligence, technological prowess in AI has often been measured by its ability to outmaneuver humans in competitive games like Chess, Poker, or Go (Hassabis (2017); Silver et al. (2017)). The accomplishments of Artificial Agents (AAs) in these arenas emphasize the potency of sequential decision-making algorithms (Crandall et al. (2018)), and in particular Deep Reinforcement Learning (DRL) based agents. While many studies have honed DRL-agents for two-player games against other AAs or humans, the predominant focus on competitive arenas has often overlooked the invaluable realm of human-machine collaboration (Jaderberg et al. (2019); Silver et al. (2017); Vinyals et al. (2019)). Given that real-world situations frequently prioritize collaboration over competition, there's an imperative to cultivate AAs that can adeptly cooperate with humans. True synergy is not just about interaction but developing AAs that genuinely assist humans in mutual objectives (Carroll et al. (2019); European Commission et al. (2020)). Although games like Dota and Capture the Flag have shown potential in Human-Artificial Agent (HAA) collaboration, the focus remains more on individual AI prowess than collaborative efficacy (Carroll et al. (2019); Rosero et al. (2021); Yang et al. (2022)). Such limitations emphasize the need for training environments that offer more realistic human-inclusive scenarios. Video games, given their intricate yet controlled settings, emerge as a promising avenue for this endeavor.

Project & Environment descriptions

The game "Overcooked" epitomizes the potential of video games for delving into Human-AI collaboration, given its cooperative gameplay dynamics. As depicted in the adapted Overcooked environment above, it is tailored for the research theme HAICO (Human-AI COllaboration). While existing research has employed a Self Play agent, there's room to probe other methodologies, like those shown in the succeeding illustration.

In this context, this project aims to benchmarck different Deep Reinforcement Learning (DRL) based agents: (i) **Population-based training (PBT)** which evolves a group of agents by replicating successful hyperparameters;

(a) Overcooked environment developed with Unity
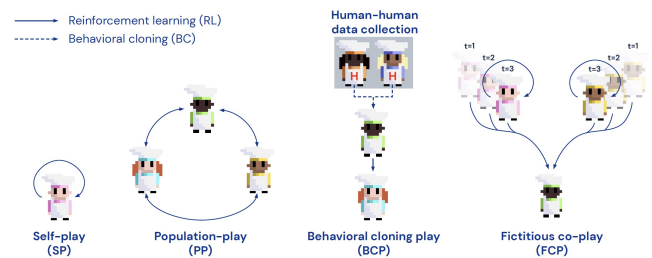
(b) DRL methods for collaborative Human-AA settings

Figure 1

(ii) **Behavioral cloning** which teaches AI to imitate human actions from observed data; and (iii) **Fictitious Co-play (FCP)** which trains agents to adapt to diverse strategies seen in collaborators, enabling flexible coordination.

Objectives

In detail, the objectives of the internship are:

- Review advancements in RL techniques geared for Human-Artificial Agent collaboration.

- Integrate varied methodologies within the Overcooked platform.

- Undertake agent training via different techniques.

- Formulate an extensive experimental methodology to assess the DRL techniques.

- Lead the recruitment of human participants for experimental purposes.

- Analyze and elucidate post-experiment data.

- Get introduced to the nuances of scientific article composition.

Profile of the Ideal Candidate:

Educational Background

An MSc or Engineering student with a robust background in Robotics, AI, and Reinforcement Learning is preferred. Familiarity with human behavior and cognitive sciences will be a notable plus.

Technical Skills

Candidates should showcase scientific aptitude, programming proficiency in languages like C#, Python, or C++, and familiarity with the Unity game engine. However, those keen to learn coding will also be considered.

Personal Attributes

A passion for cutting-edge research, the capability to work independently, adaptability, and exemplary communication skills are essential.

Administrative information

Laboratory location

DCAS Department
ISAE-SUPAERO (Institut Supérieur de l'Aéronautique et de l'Espace)
10 Av. Edouard Belin, 31400 Toulouse

Compensation

Interns will receive a stipend of 4.05€ net per hour, equating to approximately 560€ per month.

Duration

Five or six months, commencing in February or March 2024, with potential opportunities for a subsequent PhD.

Application to be sent to supervisors:

Prospective candidates should forward their CV and cover letter to:
Dr. Christophe Lounis at `christophe.lounis@isae-supaero.fr`
Dr. Caroline Chanel at `caroline.chanel@isae-supaero.fr`.

References

Carroll, M., Shah, R., Ho, M. K., Griffiths, T., Seshia, S., Abbeel, P., and Dragan, A. (2019). On the utility of learning about humans for human-ai coordination. *Advances in neural information processing systems*, 32.

Crandall, J. W., Oudah, M., Tennom, Ishowo-Oloko, F., Abdallah, S., Bonnefon, J.-F., Cebrian, M., Shariff, A., Goodrich, M. A., and Rahwan, I. (2018). Cooperating with machines. *Nature communications*, 9(1):233.

European Commission, A. et al. (2020). White paper on artificial intelligence: A european approach to excellence and trust. *COM*, 65:11–12.

Hassabis, D. (2017). Artificial intelligence: chess match of the century.

Jaderberg, M., Czarnecki, W. M., Dunning, I., Marris, L., Lever, G., Castaneda, A. G., Beattie, C., Rabinowitz, N. C., Morcos, A. S., Ruderman, A., et al. (2019). Human-level performance in 3d multiplayer games with population-based reinforcement learning. *Science*, 364(6443):859–865.

Rosero, A., Dinh, F., de Visser, E. J., Shaw, T., and Phillips, E. (2021). Two many cooks: Understanding dynamic human-agent team communication and perception using overcooked 2. *arXiv preprint arXiv:2110.03071*.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., et al. (2017). Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*.

Vinyals, O., Babuschkin, I., Chung, J., Mathieu, M., Jaderberg, M., Czarnecki, W. M., Dudzik, A., Huang, A., Georgiev, P., Powell, R., et al. (2019). Alphastar: Mastering the real-time strategy game starcraft ii. *DeepMind blog*, 2:20.

Yang, M., Carroll, M., and Dragan, A. (2022). Optimal behavior prior: Data-efficient human models for improved human-ai collaboration. *arXiv preprint arXiv:2211.01602*.