

Offre de stage M2

Apprentissage par renforcement observationnel

Pour toute candidature, envoyez votre CV ainsi qu'une lettre de motivation à `meritzell.vinyals@inrae.fr`, `leo.saulieres@inrae.fr` ou `regis.sabbadin@inrae.fr`.

Sujet

Nous considérons l'apprentissage comme l'amélioration des performances à travers l'expérience. Les expériences directes sont celles que l'apprenant acquiert par ses propres interactions avec l'environnement. Les expériences indirectes, quant à elles, proviennent de l'observation d'autres individus interagissant dans ce même environnement.

D'un côté, l'*apprentissage par observation* (AO) [BRH11] – apprendre en observant le comportement d'autres agents dans le même environnement – permet aux humains et aux animaux de découvrir des comportements utiles qui seraient difficiles à acquérir par l'exploration individuelle, et de s'adapter rapidement à de nouvelles circonstances. De l'autre, l'*apprentissage par renforcement* (AR) [SB18] est un paradigme dans lequel un agent apprend à interagir avec un environnement afin de maximiser un signal de récompense en élaborant une politique de décision. La combinaison de ces deux approches, appelée *apprentissage par renforcement observationnel* (ARO), constitue un champ de recherche émergent [BPMP17, NELJ21]. L'ARO vise à apprendre des politiques d'AR en exploitant les observations des actions d'autres agents dans l'environnement. Cependant, ce domaine manque encore de benchmarks standardisés permettant de tester et comparer les algorithmes existants.

L'objectif principal de ce stage est donc de concevoir un benchmark pour l'apprentissage par renforcement observationnel. Ce benchmark sera développé en Python avec l'API Gymnasium (<https://gymnasium.farama.org>), dédiée à la création d'environnements de simulation pour le développement et l'évaluation d'algorithmes d'apprentissage par renforcement. Un second objectif consistera à implémenter et évaluer des stratégies simples (heuristiques) ainsi que des algorithmes existants dans ce benchmark.

Objectifs du stage

- Réaliser une revue de la littérature sur l'apprentissage par renforcement observationnel.
- Concevoir et implémenter un benchmark dédié à l'ARO en utilisant l'API Gymnasium.
- Implémenter et évaluer des stratégies simples (heuristiques) et des algorithmes d'ARO existants dans ce benchmark.

Profil de candidat souhaité

- Connaissances en apprentissage par renforcement.
- Expérience en programmation Python.
- Autonomie et capacité de travail en équipe, de rédaction, de synthèse.
- Des compétences en développement collaboratif (utilisation de GIT, tests unitaires, documentation...) sont un plus, même si elles pourront être acquises pendant le stage.

Cadre de travail

Le stage, d'une durée de 5 à 6 mois et rémunéré, se déroulera au sein de l'unité MIAT (Mathématiques et Informatique Appliquées de Toulouse) de l'INRAE, située à Toulouse (24, chemin de Borde Rouge,

31320 Auzeville-Tolosane). Le Centre Inrae est facilement accessible en transports en commun (ou en vélo), dispose d'un restaurant d'entreprise, et d'une association propose de nombreuses activités sportives et autres.

Encadrement

Le stage sera encadré par Meritxell Vinyals, Léo Saulières et Régis Sabbadin (MIAT, INRAE, Toulouse), dont les sujets de recherche gravitent autour de l'apprentissage par renforcement.

Références

- [BPMP17] Diana Borsa, Bilal Piot, Rémi Munos, and Olivier Pietquin. Observational learning by reinforcement learning. *arXiv preprint arXiv :1706.06617*, 2017.
- [BRH11] Robert Boyd, Peter J Richerson, and Joseph Henrich. The cultural niche : Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, 108(supplement_2) :10918–10925, 2011.
- [NELJ21] Kamal K Ndousse, Douglas Eck, Sergey Levine, and Natasha Jaques. Emergent social learning via multi-agent reinforcement learning. In *International conference on machine learning*, pages 7991–8004. PMLR, 2021.
- [SB18] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning : An Introduction*. The MIT Press, second edition, 2018.