

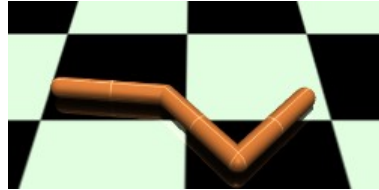
Apprentissage par renforcement sur Swimmer partiellement observable

Mots clés : apprentissage par renforcement, observabilité partielle, mémoire et plans, débog

Encadrant : Olivier Sigaud, ISIR, Sorbonne Université (Olivier.Sigaud at isir.upmc.fr)

Nombre d'étudiants : 2

Résumé : L'objectif de ce projet est d'étudier l'application d'algorithmes d'apprentissage par renforcement profond à l'environnement Swimmer, dans le cas où l'agent est privé de certains capteurs.



Sujet développé :

L'apprentissage par renforcement permet à un agent d'apprendre à contrôler un système de façon à maximiser un signal de récompense. En général, l'agent dispose d'un accès complet à l'état du système grâce à ses capteurs, ce qui lui permet de choisir au mieux l'action à réaliser à chaque pas de temps. Lorsque cet accès est incomplet, on parle d'observabilité partielle et les algorithmes d'apprentissage par renforcement fonctionnent moins bien. Pour compenser l'information manquante, on peut les doter d'une mémoire des états précédents ou d'une capacité à planifier des séquences d'actions. L'objet de ce projet est d'étudier l'impact de ces mécanismes dans Swimmer, un environnement standard (*benchmark*) d'apprentissage par renforcement dans lequel une créature articulée doit nager le plus vite possible dans un fluide (cf. image).

Le projet sera réalisé dans un premier temps à l'aide de la librairie locale BBRL, puis les résultats obtenus seront confrontés finement à ceux que l'on peut obtenir avec d'autres librairies d'apprentissage par renforcement validées par une large communauté, comme Stable Baselines 3. L'objectif secondaire derrière cette comparaison sera la détection et la résolution d'éventuels bugs profonds dans la librairie BBRL.

Pour les étudiant.e.s qui seront choisi.e.s, les bénéfices de ce projet seront multiples :

- ce sera l'occasion de se familiariser avec les concepts et outils de l'apprentissage par renforcement, qui jouent un rôle significatif dans les progrès actuels de l'Intelligence Artificielle. Les étudiant.e.s verront en avance une partie du contenu de l'UE RL du M2 MIND.
- ce sera l'occasion de développer une bonne maîtrise du codage d'algorithmes dans BBRL et d'apprendre une méthodologie rigoureuse pour l'évaluation d'algorithmes.

Ce projet faisant appel à de nombreuses notions nouvelles, une formation initiale à l'apprentissage par renforcement et à des aspects méthodologiques sera assurée durant les premières semaines.

Une forte motivation pour l'apprentissage par renforcement, des bases très solides en programmation python et une bonne maîtrise de github sont des prérequis.

Références

- BBRL : <https://github.com/osigaud/BBRL>
- Stable Baselines 3 : <https://stable-baselines3.readthedocs.io/en/master/>
- Swimmer : <https://gymnasium.farama.org/environments/mujoco/swimmer/>