

Apprentissage par renforcement avec AFU

Mots clés : apprentissage par renforcement, BBRL, python

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

Nombre d'étudiants : 2

Résumé : L'objectif de ce projet est de développer une version de l'algorithme AFU en BBRL puis d'étudier certaines de ses propriétés.

Sujet développé :

Contexte

Actor-Free critic Updates (AFU) est un algorithme d'apprentissage par renforcement très récent et original par bien des aspects. L'auteur de l'algorithme, Nicolas Perrin-Gilbert, est membre de l'ISIR et il n'a pas encore étudié toutes les propriétés de son algorithme.

On dit d'un algorithme d'apprentissage par renforcement qu'il est « *off-policy* » s'il peut apprendre à partir de données qui n'ont pas été générées par l'agent en train d'apprendre. Un cas extrême pour un algorithme *off-policy* consiste à réussir à apprendre à partir de données générées avec des actions aléatoires. C'est le cas d'un des algorithmes les plus simples, Q-learning, dans le cas où les états et les actions sont discrets. Mais une telle propriété n'a été mise en évidence pour aucun algorithme d'apprentissage par renforcement profond travaillant avec des états et des actions continues.

Objectifs

Pour enseigner l'apprentissage par renforcement, l'encadrant de ce projet a développé une librairie à vocation pédagogique appelée BBRL. Le premier objectif de ce projet est de coder AFU dans la librairie BBRL. Le second objectif est d'étudier les propriétés d'AFU et, en particulier, de déterminer si AFU est capable d'atteindre une politique optimale à partir de données engendrées aléatoirement.

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 RLD du M2 DAC.
- ce sera l'occasion de développer une bonne maîtrise du codage d'algorithmes dans BBRL, utilisée en RLD, 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, à BBRL et à des aspects méthodologiques sera assurée durant les premières semaines.

Une forte motivation pour le *machine learning*, des bonnes bases en programmation python et une bonne maîtrise de github sont des prérequis.

Références

- Le dépôt de la librairie BBRL : <https://github.com/osigaud/BBRL>

- AFU: Actor-Free critic Updates in off-policy RL for continuous control
<https://arxiv.org/pdf/2404.16159>