

Compte Rendu de Réunion 1

PRÉSENTATION DU PROJET

Le projet s'articule autour de l'étude de l'algorithme AFU (*Actor-Free Updates*), récemment proposé par Nicolas Perrin. Bien que l'article original soit mathématiquement complexe, l'algorithme présente un potentiel significatif qui mérite d'être exploré et mis en valeur.

OBJECTIF PRINCIPAL

L'objectif principal est d'étudier les propriétés *off-policy* d'AFU. L'hypothèse est qu'AFU pourrait être véritablement *off-policy*, contrairement à de nombreux algorithmes de deep RL actuels qui, bien que classés comme *off-policy*, montrent des limitations significatives avec des données éloignées de leur politique courante. Pour tester cette hypothèse, il faut utiliser des données générées uniformément dans l'espace état-action, un scénario où le Q-Learning classique converge mais où la plupart des algorithmes de deep RL échouent.

Si le temps le permet, il est possible également explorer un second axe : la capacité d'AFU à gérer efficacement la transition entre apprentissage *offline* et *online*, où la plupart des algorithmes actuels montrent une dégradation temporaire des performances.

APPROCHE MÉTHODOLOGIQUE

L'étude débutera avec des environnements simples de *Gymnasium*, notamment *CartPole* en version continue, *MountainCar* ou *Pendulum*. Ces environnements permettront une première validation des propriétés d'AFU. La difficulté technique principale résidera dans la génération d'états uniformes, l'état représentant l'interaction entre l'agent et l'environnement. Une modification de l'API des environnements sera nécessaire pour permettre la configuration directe des états.

L'implémentation se fera dans le framework BBRL, qui offre une base solide pour le développement d'algorithmes de RL.

ASPECTS ADMINISTRATIFS ET ORGANISATION

Le projet s'organise autour de réunions hebdomadaires, avec possibilité de passage à un rythme bihebdomadaire à mi-semestre selon les besoins. Chaque réunion nécessite l'envoi d'un ordre du jour 24h à l'avance, le mardi matin et est suivie d'un compte-rendu.

Les livrables attendus comprennent :

- Un cahier des charges à mi-semestre
- Un travail bibliographique documenté
- Un rapport final
- Une soutenance devant un jury, incluant M. Sigaud

L'évaluation sera collective et prendra en compte l'avis de l'encadrant. La note finale sera basée sur l'ensemble des livrables et la présentation orale.

PROCHAINE ÉTAPE

La priorité est de se familiariser avec les concepts fondamentaux du RL et du deep RL, en commençant par le cours sur le RL tabulaire prévu pour vendredi et BBRL. La mise en place du repository Github, des libraries et d'un environnement de test automatisé est aussi prévue pour cette semaine.