

Apprentissage par renforcement et diversité des comportements appris par les agents interactifs

Laboratoire d'accueil : laboratoire CEDRIC, Conservatoire national des arts et métiers (Paris 3^e)

Durée du stage : 5 à 6 mois, démarrage entre mars et juin 2021

Rémunération : environ 600 euros/mois

Encadrants : Nicolas Audebert, Guillaume Levieux, Clément Rambour



À gauche, deux agents tentent de marquer un but contre un troisième qui défend les cages. À droite, un agent tente de détruire un robot en tirant dessus.

1 Contexte

Une force des médias interactifs est de générer des situations qui attirent la curiosité. Un jeu vidéo est généralement plus apprécié s'il propose une grande variété de situations et parvient à surprendre les joueurs et joueuses, les incitant à varier leurs stratégies de victoire. Il est ainsi souhaitable de pouvoir évaluer si un environnement de jeu (règles, placement des obstacles, comportements des adversaires...) est propice ou non à encourager des approches nouvelles. Ce stage s'intéresse à produire des comportements de jeux variés et à quantifier cette variété de sorte à comprendre comment modifier l'environnement pour encourager de nouveaux comportements.

En parallèle, une piste largement explorée dans la littérature consiste à exploiter l'apprentissage par renforcement pour produire des agents virtuels capables d'explorer et d'interagir avec l'environnement de jeu. De tels agents servent généralement à simuler des adversaires « intelligents » mais peut aussi servir à simuler le comportement d'un joueur virtuel. Dans une phase de prototypage, il devient alors envisageable d'utiliser ces comportements simulés afin d'obtenir des métriques de diversité des actions entreprises, plutôt que de devoir acquérir des données réelles. Les récentes publications en *machine learning* tendent à montrer qu'il est possible pour des agents d'apprendre à jouer à de nombreux jeux, allant de jeux simples et fortement contraints (Space Invaders, Breakout...) [5] à des jeux à environnements ouverts et aux contrôles plus libres tels que des FPS [4, 3]. Toutefois, ces méthodes d'apprentissages tendent à converger vers des solutions optimales souvent identiques. Les agents présentent ainsi des comportements répétitifs, brisant l'illusion de réalité s'ils doivent interagir avec un joueur humain. Apprendre des comportements sous-optimaux mais plus variés demeure donc un défi important [7].

2 Objectifs

L'objectif central de ce stage est de construire une intelligence artificielle capable de favoriser la diversité de l'expérience du joueur en modifiant à la volée son environnement.

Ce stage se découpe en trois grandes étapes :

1. Définir des métriques caractérisant la diversité des comportements d'un joueur (réel ou simulé) en exploitant les historiques d'actions sur différentes parties.
2. Implémenter et comparer différents algorithmes d'apprentissage par renforcement dans deux environnements de jeux de référence.
3. Concevoir et implémenter des variantes de ces algorithmes produisant des agents aux stratégies diversifiées [1, 6].

La plateforme expérimentale se basera sur Unity ML-Agents [2]¹ afin de créer des environnements de jeux variés instrumentés pour l'apprentissage par renforcement.

3 Profil

Nous recherchons pour ce stage un-e candidat-e de niveau M2 ou dernière année d'école d'ingénieur avec une formation en apprentissage automatique ou en informatique. Le ou la candidat-e idéal-e a une appétence pour la recherche et des bases en apprentissage par renforcement. Sans être indispensable, un intérêt pour les médias interactifs est un point positif pour ce stage.

Une connaissance de la programmation avec Python est préférable, il est toutefois envisageable pour un-e candidat-e connaissant un autre langage de programmation de se former au cours du stage. Une première expérience avec une bibliothèque d'apprentissage profond (TensorFlow ou PyTorch) ou la programmation de jeux (Unity) est la bienvenue.

4 Organisation

Cette offre de stage porte sur un stage de 5 à 6 mois avec un début au printemps 2021. Le stage se déroulera au centre de recherche et d'études en informatique et en communications (CEDRIC) du Conservatoire national des arts et métiers (Cnam) à Paris, 3^e arrondissement.

Le stage sera co-encadré par Nicolas Audebert, Guillaume Levieux et Clément Rambour, enseignants-chercheurs au Cnam.

5 Candidater

Envoyer une candidature (CV + brève explication de votre motivation) par email à :

nicolas.audebert@cnam.fr, guillaume.levieux@lecnam.net, clement.rambour@lecnam.net.

Références

- [1] Z.-W. Hong, T.-Y. Shann, S.-Y. Su, Y.-H. Chang, T.-J. Fu, and C.-Y. Lee. Diversity-driven exploration strategy for deep reinforcement learning. In *NeurIPS*, page 12.
- [2] A. Juliani, V.-P. Berges, E. Teng, A. Cohen, J. Harper, C. Elion, C. Goy, Y. Gao, H. Henry, M. Mattar, and D. Lange. Unity : A general platform for intelligent agents.
- [3] A. Khan et al. Playing first-person shooter games with machine learning techniques and methods using the VizDoom game-AI research platform. 34:100357.
- [4] G. Lample and D. S. Chaplot. Playing FPS games with deep reinforcement learning. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, AAAI'17*. AAAI Press.
- [5] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller. Playing atari with deep reinforcement learning. In *NeurIPS Deep Learning Workshop*, page 9.
- [6] R. Shen et al. Generating behavior-diverse game AIs with evolutionary multi-objective deep reinforcement learning. volume 4, pages 3371–3377.
- [7] I. Szita, M. Ponsen, and P. Spronck. Effective and diverse adaptive game AI. 1(1):16–27. Conference Name : IEEE Transactions on Computational Intelligence and AI in Games.

1. <https://github.com/Unity-Technologies/ml-agents>