

8INF974 – ATELIER PRATIQUE EN INTELLIGENCE ARTIFICIELLE II

Projet #3 – Agent intelligent avec Gymnasium

Objectifs du projet :

- Explorer l'apprentissage par renforcement
- Acquérir de l'expérience sur Gymnasium, la plus importante librairie d'expérimentation pour le RL
- Comparer des approches de différents types
- Approfondir ses connaissances de manière autodidacte
- Acquérir de l'expertise dans le développement d'intelligences artificielles

Mise en situation :

Gymnasium est une reprise publique du projet abandonné d'OpenAI appelé Gym. Vous tenterez d'implémenter deux méthodes très populaires pour jouer à des jeux d'Atari. Vous ne serez pas noté sur les résultats, mais bel et bien sur vos efforts et la qualité de vos démarches.

Travail à effectuer :

Ce projet se divise en quatre grandes étapes. Premièrement, vous devez faire fonctionner Gymnasium et faire des tests avec différents environnements. Cependant, vous vous concentrerez d'abord sur un des *Zero-sum game* déterministe. Vous pouvez choisir *Video Chess*, *Othello*, *Video Checkers* ou *Tic Tac Toe 3D* sur Atari ([voir environnements ici](#)). Attention! Les jeux plus complexes seront plus difficiles à résoudre.

Votre objectif est d'implémenter ce qu'on appelle un *Monte Carlo Tree Search* une amélioration très puissante de *MiniMax* et *Alpha-Beta Pruning* en exploration locale. Il s'agit d'une technique qui a été exploitée la première fois avec UCB (problèmes de bandits, apprentissage par renforcement) dès les années 80. C'est toujours la méthode la plus populaire pour résoudre les problèmes de ce type.

Ensuite, vous devrez aussi implémenter directement une méthode Deep Q Learning et tenter d'entraîner une IA sur le ou les jeux Atari de votre choix. De plus, si vous souhaitez un plus gros défi, vous pouvez tenter de trouver et implémenter une approche récente sur le/les jeux d'Atari de votre choix. Je vous suggère de chercher une approche avec du code disponible tel que EfficientZero publié à NeurIPS 2021: <https://github.com/YeWR/EfficientZero>. Attention! Certaines approches seront trop gourmandes pour vos ordinateurs, soyez donc prudents dans vos tentatives d'entraînement.

Pour ce projet, on s'attend à l'équivalent de **30 heures de travail par personne** sur 5 semaines (2 et demi en semestre condensé d'été).

Rencontres avec le professeur :

Vous devez être préparés pour vos rencontres avec le professeur. Celles-ci devraient se dérouler de cette façon :

1. Récapitulatif d'où vous en étiez précédemment
2. Présentation de vos travaux respectifs
3. Présentation de votre plan et des problèmes auxquels vous faites face
4. Apprentissage fait depuis la dernière rencontre (le cas échéant)

Présentation de projet :

Pour vos présentations de projet, vous devriez être en mesure de succincement :

1. Présenter les points clés de votre implémentation

2. Effectuer une démonstration
3. Faire un bilan des bugs/défis/apprentissages (e.g., ce que vous feriez différemment)

Toutes les équipes doivent assister à toutes les présentations. Par la suite nous prendrons quelques minutes pour discuter tous/toutes ensemble.