

Enseignant(s)

VIDAL Nicolas

Email(s)

nvidal@myges.fr

2021-4A-IABD - (Deep) Reinforcement Learning P1

1 Matières, formations et groupes

Matière liée au projet :

Formations : -

Nombre d'étudiant
par groupe :**3**Règles de constitution des groupes: **Libre**Charge de travail
estimée par étudiant : **30,00 h**

2 Sujet(s) du projet

Type de sujet : **Imposé**

3 Détails du projet

Objectif du projet (à la fin du projet les étudiants sauront réaliser un...)

Implémenter et comprendre quand utiliser les algorithmes classiques de l'apprentissage par renforcement, qu'ils soient issus de la programmation dynamique ou de techniques à base de simulations Monte Carlo, comprendre l'impact de l'utilisation de techniques d'approximation de fonctions dans ce contexte.

Descriptif détaillé

Dans ce projet, les étudiants devront en premier lieu implémenter les algorithmes suivants et s'assurer de leur bonne implémentation sur des cas de tests bien précis :

Partie 1

- Dynamic Programming
 - Création des contrats liés à un MDP
 - Création d'un environnement simple : Line World
 - Implémentation de l'algorithme "Policy Evaluation"
 - Implémentation de l'algorithme "Policy Iteration"
 - Implémentation de l'algorithme "Value Iteration"
 - Tests et vérifications de ces algorithmes sur l'environnement Line World
 - Création d'un environnement un peu plus complexe : Grid World
 - Tests et vérifications de ces algorithmes sur l'environnement Grid World
 - Evaluation sur l'environnement *secret 1*
- Méthodes Monte Carlo
 - Implémentation d'un TicTacToe 1 (joueur VS Random)
 - Implémentation de l'algorithme "Monte Carlo ES"
 - Implémentation de l'algorithme "On-policy first visit Monte Carlo Control"
 - Implémentation de l'algorithme "Off-policy Monte Carlo Control"
 - Tests et vérifications de ces algorithmes sur l'environnement TicTacToe 1
 - Evaluation sur l'environnement *secret 2*
- Temporal Difference Learning
 - Implémentation de l'algorithme Sarsa
 - Implémentation de l'algorithme Q-Learning
 - Implémentation de l'algorithme Expected Sarsa
 - Tests et vérifications de ces algorithmes sur l'environnement TicTacToe 1
 - Optionnel : Implémentation de l'algorithme "n-step Q-Learning"
 - Evaluation sur l'environnement *secret 3*
- Planning
 - Optionnel : Implémentation de l'algorithme "Dyna-Q"
 - Optionnel : Implémentation de l'algorithme "Dyna-Q+"
 - Evaluation sur les précédents environnements

Partie 2

- Deep Reinforcement Learning
 - Création d'un environnement beaucoup plus complexe : PacMan
 - Implémentation de l'algorithme "Episodic Semi-gradient Sarsa"
 - Implémentation de l'algorithme "Deep Q Learning"
 - Tests et vérifications de ces algorithmes sur l'environnement TicTacToe 1
 - Tests et vérifications de ces algorithmes sur l'environnement PacMan
 - Optionnel : Implémentation de l'algorithme "Double Deep Q Learning"
 - Evaluation sur l'environnement *secret 5*
- Policy gradient Methods
 - Implémentation de l'algorithme "REINFORCE"
 - Implémentation de l'algorithme "REINFORCE with Baseline"
 - Tests et vérifications de ces algorithmes sur l'environnement TicTacToe 1
 - Tests et vérifications de ces algorithmes sur l'environnement PacMan
 - Evaluation sur l'environnement *secret 5*

Il sera également nécessaire de présenter une interface graphique permettant de regarder jouer chaque agent et également de mettre à disposition un agent 'humain' (au pire en ligne de commande).

L'interface graphique de l'environnement *secret* ne sera fournie qu'en fin de cours.

Pour chaque environnement, les étudiants devront étudier les performances de l'agent et retranscrire leur résultats.

Les étudiants devront fournir l'intégralité du code leur ayant permis d'obtenir leurs résultats ainsi que les

modèles (keras ou pytorch ou torch-rs) entraînés et sauvegardés prêts à être exécutés pour confirmer les résultats présentés.

Les étudiants devront présenter ces résultats lors d'une soutenance. Dans cette dernière, les étudiants devront faire valoir leur méthodologie de choix d'hyperparamètres, et proposer leur interprétation des résultats obtenus

Ouvrages de référence (livres, articles, revues, sites web...)

Reinforcement Learning: An Introduction :
<http://incompleteideas.net/book/the-book.html>
Richard S. Sutton
and Andrew G. Barto

Outils informatiques à installer

Unity / Unreal Engine / PyGame / ... ?
Tensorflow / Keras / pyTorch / torch-rs pour DQN et REINFORCE

4 Livrables et étapes de suivi

1

Rendu final

Présentation des résultats et du protocole suivi pour les environnements de Deep Reinforcement Learning
Livrables :
- rapport sous la forme de Notebook Jupyter + pdf
- présentation (slides) de 10 minutes
- sources complètes du projet
- démonstration

jeudi
29/07/2021
10h00

5 Soutenance

Durée de présentation
par groupe :

15 min

Audience : **A huis clos**

Type de présentation :

Présentation / PowerPoint - Démonstration

Précisions :