

M1 Informatique – UE Projet

Carnet de bord : les coulisses de la recherche documentaire

Les éléments que vous indiquez dans ce carnet donneront lieu à une notation

Noms, prénoms et spécialité :

CELLIER Roxane, ANDROIDE
FU Zhenyue, ANDROIDE
QIN Yi, ANDROIDE

Sujet :

Apprentissage conditionné par des buts en BBRL
--

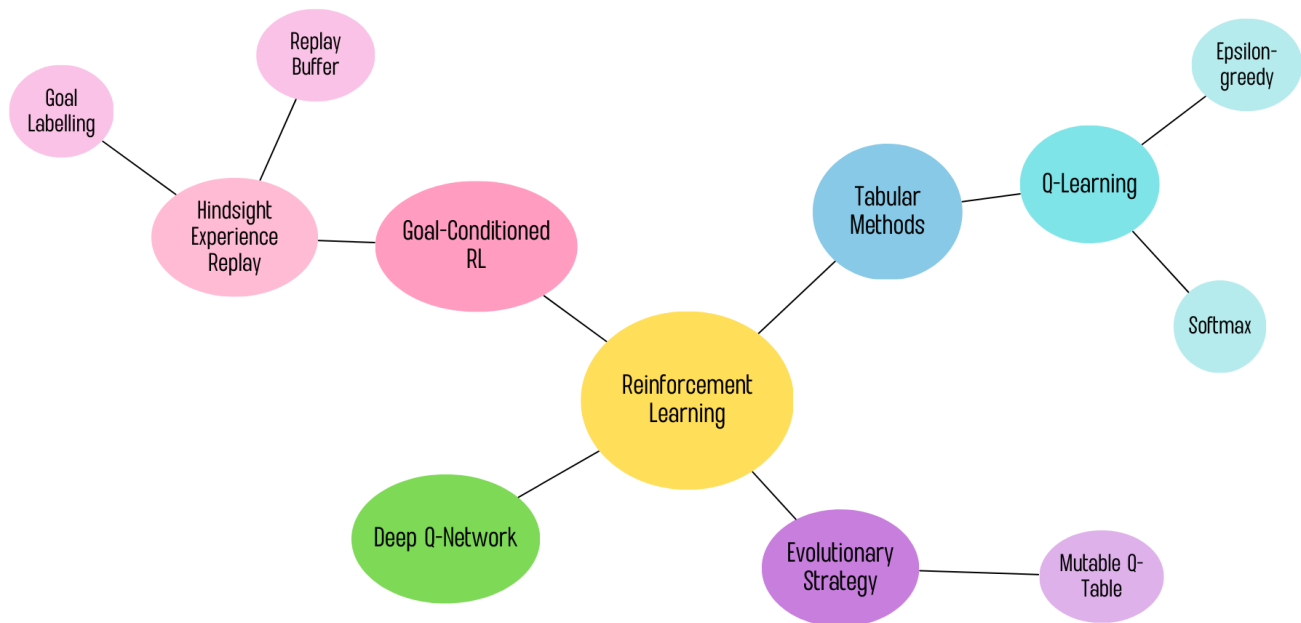
Consigne :

- 1 **Introduction (5-10 lignes max)** : Décrivez rapidement votre sujet de recherche, ses différents aspects et enjeux, ainsi que l'angle sous lequel vous avez décidé de le traiter.
- 2 **Les mots clés retenus** : Listez les mots-clés que vous avez utilisés pour votre recherche bibliographique. Organisez-les sous forme de carte heuristique.
- 3 **Descriptif de la recherche documentaire (10-15 lignes)** : Décrivez votre utilisation des différents outils de recherche (moteurs de recherche, base de donnée, catalogues, recherche par rebond etc.). Comparez ensuite les outils entre eux. A quelles sources vous ont-ils permis d'accéder ? Quelles sont leurs spécificités ? Quel est leur niveau de spécialisation ?
- 4 **Bibliographie produite dans le cadre du projet** : Utilisez la norme ACM.
- 5 **Evaluation des sources (5 lignes minimum par source)** : Choisissez 3 sources parmi votre bibliographie, décrivez la manière dont vous les avez trouvées et faites-en une évaluation critique en utilisant les critères vus sur les supports de TDs.

1. Introduction

L'objectif de notre sujet est d'étudier différentes méthodes de Reinforcement Learning. En commençant par des méthodes tabulaires simples (Q-Learning par états-actions), nous devons étendre nos codes et algorithmes pour nous permettre d'implémenter et d'étudier des techniques plus poussées. Les tests de nos programmes et de nos algorithmes se feront par la recherche par un agent d'un chemin vers un (ou plusieurs) but(s) dans un labyrinthe. Nous cherchons ainsi à apprendre la meilleure politique de déplacement pour notre agent au sein de cet environnement.

2. Mots-clés



3. Recherche Documentaire

Wikipédia est une encyclopédie collaborative qui a pour but de centraliser les connaissances. Le rattachement des pages entre elles permet un rebond rapide entre différents articles, ce qui simplifie les recherches sous un même domaine. Les articles offrent souvent des définitions claires, cependant l'aspect collaboratif peut rendre certaines sources peu fiables, et les articles concernant des domaines pointus sont souvent très incomplets. Ils fournissent par contre de nombreuses sources sur lesquelles nous pouvons étendre notre lecture. Nous avons utilisé Wikipédia comme une première approche pour certains concepts au début de notre projet.

Google Scholar est un moteur de recherche en ligne qui nous permet de trouver des articles scientifiques, des thèses, des livres, des rapports et des documents, notamment par auteur ou par thème. Puisqu'il s'agit d'articles publiés et vérifiés, son niveau de spécialisation est assez élevé. La recherche se fait par mot-clé ou concept, et le moteur renvoie les articles de la base de données les plus cités et les plus

pertinents. À la manière de Google, le moteur nous renvoie ensuite vers le site hébergeant l'article en question.

ArXiv est un site pour trouver des articles scientifiques en pré-publication. Les articles publiés dans ArXiv sont de grande qualité et sont souvent utilisés comme références pour des recherches ultérieures. Le site nous permet d'orienter nos recherches selon le domaine voulu (Physique, Sciences Informatiques). Il est également possible d'affiner par sous-domaine (Machine Learning, Robotique). Le site nous offre aussi la possibilité de chercher par titre, auteur ou mot-clé, nous proposant les résultats selon la date de pré-publication et la pertinence de l'article.

Papers with Code est un site communautaire réunissant de nombreuses publications scientifiques principalement sur le domaine du Machine Learning, permettant facilement la recherche par thème. Les publications peuvent être rattachées à des dépôts github, permettant également l'étude de codes implémentant certaines approches proposées dans ces articles. Cette plateforme est supervisée par Meta AI Research, ce qui peut attester d'une certaine légitimité.

4. Bibliographie

- [1] OpenAI, Marcin Andrychowicz et al. 2017. Hindsight Experience Replay. *NeurIPS*. Retrieved from <http://papers.neurips.cc/paper/7090-hindsight-experience-replay.pdf>
- [2] DeepMind Technologies, Volodymyr Mnih et al. 2013. Playing Atari with Deep Reinforcement Learning. *University of Toronto*. Retrieved from <https://www.cs.toronto.edu/~vmnih/docs/dqn.pdf>
- [3] Shangdong Zhang and Richard S. Sutton. 2018. A Deeper Look at Experience Replay. *arXiv*. Retrieved from <https://arxiv.org/abs/1712.01275>
- [4] Ludovic Denoyer, Alfredo de la Fuente et al. 2021. SaLina: Sequential Learning of Agents. *arXiv*. Retrieved from <https://arxiv.org/pdf/2110.07910.pdf>
- [5] Melrose Roderick, James MacGlashan et Stefanie Tellex. 2017. Implementing the Deep Q-Network. *arXiv*. Retrieved from <https://arxiv.org/pdf/1711.07478.pdf>
- [6] Meng Fang, Cheng Zhou, et al. 2019. DHER : Hindsight Experience Replay for Dynamic Goals. *Papers with Code*. Retrieved from <https://openreview.net/pdf?id=Byf5-30qFX>
- [7] Microsoft, Zichuan Lin, Tianqi Zhao, et al. 2018. Episodic Memory Deep Q-Networks. *arXiv*. Retrieved from <https://arxiv.org/pdf/1805.07603>
- [8] Minghuan Liu, Menghui Zhu et Weinan Zhang. 2022 (pré-publication). Goal-Conditioned Reinforcement Learning: Problems and Solutions. *arXiv*. Retrieved from <https://arxiv.org/pdf/2201.08299>
- [9] Devin, C., Gupta, A., Darrell, T., Abbeel, P., and Levine, S. 2016. Learning modular neural network policies for multi-task and multi-robot transfer. *arXiv*. Retrieved from <https://arxiv.org/pdf/1609.07088.pdf>
- [10] Schaul, T., Quan, J., Antonoglou, I., and Silver, D. 2015. Prioritized experience replay. *arXiv*. Retrieved from <https://arxiv.org/pdf/1511.05952.pdf>
- [11] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and

Wojciech Zaremba. Openai gym, 2016. *arXiv*. Retrieved from <https://arxiv.org/pdf/1606.01540.pdf>

[12] Wang, Z., Bapst, V., Heess, N., Mnih, V., Munos, R., Kavukcuoglu, K., and de Freitas, N. 2016. Sample efficient actor-critic with experience replay. *arXiv*. Retrieved from <https://arxiv.org/pdf/1611.01224.pdf>

5. Évaluation des sources

Hindsight Experience Replay a été écrit par un groupe d'auteurs de l'organisation OpenAI, une entreprise spécialisée dans l'intelligence artificielle et souhaitant promouvoir le développement dans ce domaine. C'est une publication scientifique qui a été présentée lors de la Conférence NeurIPS (Neural Information Processing Systems) 2017, et qui nous a été conseillée par notre encadrant comme étant un bon moyen de découvrir l'algorithme HER. Depuis sa publication, elle a été citée quelque 1.954 fois selon la base de données Google Scholar. Cet article présente un équilibre judicieux entre les preuves et l'analyse. Les auteurs fournissent des preuves solides de la validité de leur approche, tout en expliquant en détail le fonctionnement de l'algorithme du même nom. L'article répond à toutes les exigences et est très fiable de par sa provenance.

Playing Atari with Deep Reinforcement Learning a principalement été écrit par Volodymyr Mnih, avec d'autres co-auteurs, tels que Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra et Martin Riedmiller. L'objectif est de présenter des algorithmes d'apprentissage par renforcement, utilisant principalement du Q-Learning. Nous l'utilisons pour nous apporter une meilleure compréhension des algorithmes et de leurs implémentations. Volodymyr Mnih est un chercheur scientifique chez Google DeepMind. Ce papier a été initialement présenté au NIPS Deep Learning Workshop en 2013, et nous l'avons trouvé par le moteur de recherche Google Scholar. Selon Google Scholar, il a été cité environ 11.107 fois. L'article est très empirique en termes d'équilibre entre la preuve et l'analyse, décrivant les expériences et les résultats obtenus en utilisant leur méthodologie. L'article répond à toutes les exigences et est très fiable.

A Deeper Look at Experience Replay a été écrit par Shangdong Zhang et Richard S. Sutton. Shangdong Zhang est professeur adjoint au département d'informatique de l'université de Virginie, et Richard S. Sutton est un informaticien canadien, enseignant en informatique à l'université de l'Alberta. Cette publication a été initialement présentée au NIPS 2017 Deep Reinforcement Learning Symposium, et nous l'avons retrouvée grâce au site arXiv. Elle explique plus en détail le principe des algorithmes utilisant l'Experience Replay, et présente de nombreux résultats sur différentes utilisations et implémentations. Selon Google Scholar, elle a été citée environ 250 fois. L'article répond à toutes les exigences et est très fiable.