

# Reinforcement Learning : Projet

December 8, 2024

## Introduction

Dans le cadre de ce projet, vous allez mettre en oeuvre les notions et algorithmes de Reinforcement Learning (RL).

Un robot est en charge de nettoyer un environnement (dangereux pour les humains). Chaque jour, le robot commence par récupérer le matériel nécessaire à l'accomplissement de sa tâche, lequel est entreposé dans une première pièce (un hangar ( $H$ )). Ensuite, il se rend dans la seconde pièce (un entrepôt ( $E$ )) pour la nettoyer. A la fin de la journée, il se rend dans une troisième pièce (un garage ( $G$ )) pour se recharger.

Chaque matin, très tôt, des employés réapprovisionnent le hangar en matériel mais ne remettent pas le matériel au même endroit que la veille. Chaque matin, le robot doit donc apprendre à trouver le matériel dans le hangar. Il en va de même avec le garage où les stations de recharge sont déplacées chaque jour.

Le but de ce projet est de mettre en place un environnement simulé pour modéliser cette situation et de tester différents algorithmes de RL pour résoudre ce problème. Le robot sera modélisé par un agent et les pièces par des grid worlds.

## Environnement

L'environnement simulé est composé de trois grid worlds  $H$ ,  $E$  et  $G$  de tailles potentiellement différentes :

- Le premier grid world  $H$  est un environnement simple. Les cellules peuvent être des obstacles, des cellules vides ou contenir du matériel. Une cellule particulière permet de sortir de l'environnement.
- Le deuxième grid world  $E$  est un environnement plus complexe. Les cellules peuvent être des obstacles, des cellules vides ou contenir de la saleté. Une cellule particulière permet de sortir de l'environnement.
- Le troisième grid world  $G$  est un environnement simple. Les cellules peuvent être des obstacles, des cellules vides ou contenir des stations de recharge.

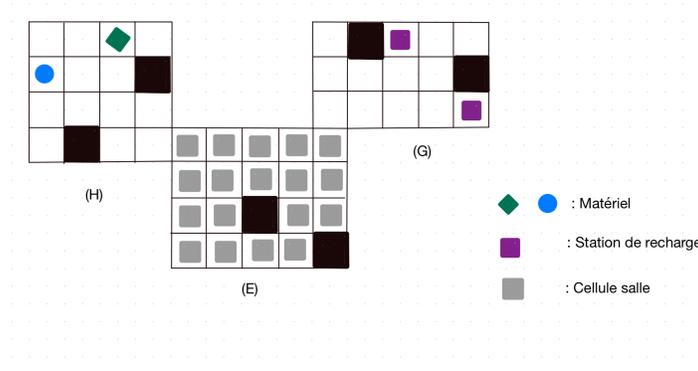


Figure 1: Environnement (au début de la journée).

## Agents

Le robot est modélisé par un agent qui évolue dans les environnements  $H$ ,  $E$  et  $G$ . Ses tâches sont les suivantes :

- Trouver le matériel dans le hangar  $H$ ,
- Nettoyer l'entrepôt  $E$ .
- Se rendre dans le garage  $G$  et y trouver une station pour se recharger.

La figure 1 illustre un exemple d'environnement simulé. Les cellules noires représentent des obstacles, les cellules blanches sont vides, Les cellules grisées dans l'entrepôts contiennent de la saleté. Pour simplifier, on considère que tout l'entrepot est sale.

La figure 2 illustre un exemple d'environnement simulé après que le robot ait trouvé le matériel dans le hangar et nettoyé l'entrepôt. Les cellules vertes dans l'entrepôt représentent les cellules nettoyées.

## Travail à réaliser

- Définir les grid worlds  $H$ ,  $E$  et  $G$ .
- Définir l'agent qui évolue dans les environnements  $H$ ,  $E$  et  $G$ .
- Définir les tâches de l'agent.
- Définir les récompenses associées à chaque tâche.
- Définir les politiques des agents.
- Implémenter un/des algorithm(e)s de RL pour résoudre le problème.

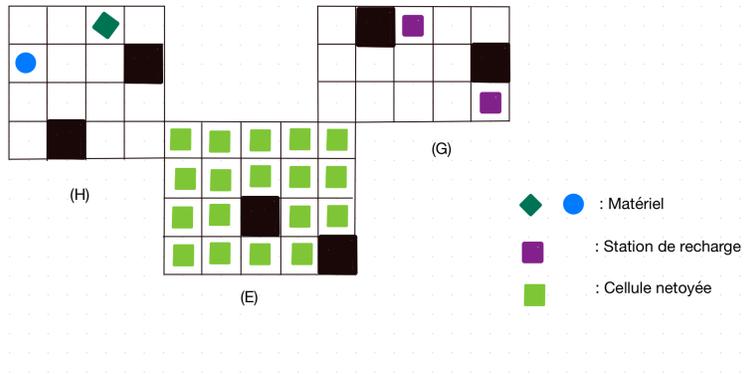


Figure 2: Environnement (après que le robot ait trouvé le matériel dans le hangar et nettoyé l'entrepôt).

Vos solutions doivent proposer différentes approches : model-based, model-free, etc.

## Rendu attendu

- Un rapport détaillant votre travail. Les explications concernant la démarche suivie et les expériences réalisées sont tout aussi importantes, voire plus, que le résultat final. Il est donc possible d'obtenir une bonne note même si les agents ne sont pas parfaits à la fin.
- Un code source propre et bien documenté.

Le projet est à réaliser en binôme. Vous avez jusqu'au 05 janvier 2025 pour rendre votre travail. Vous devez envoyer une archive (Nom1Prenom1-Nom2Prenom2.zip) contenant le rapport et le code source sur à [akka.zemmari@u-bordeaux.fr](mailto:akka.zemmari@u-bordeaux.fr)