

Minigrid Dynamic Obstacles

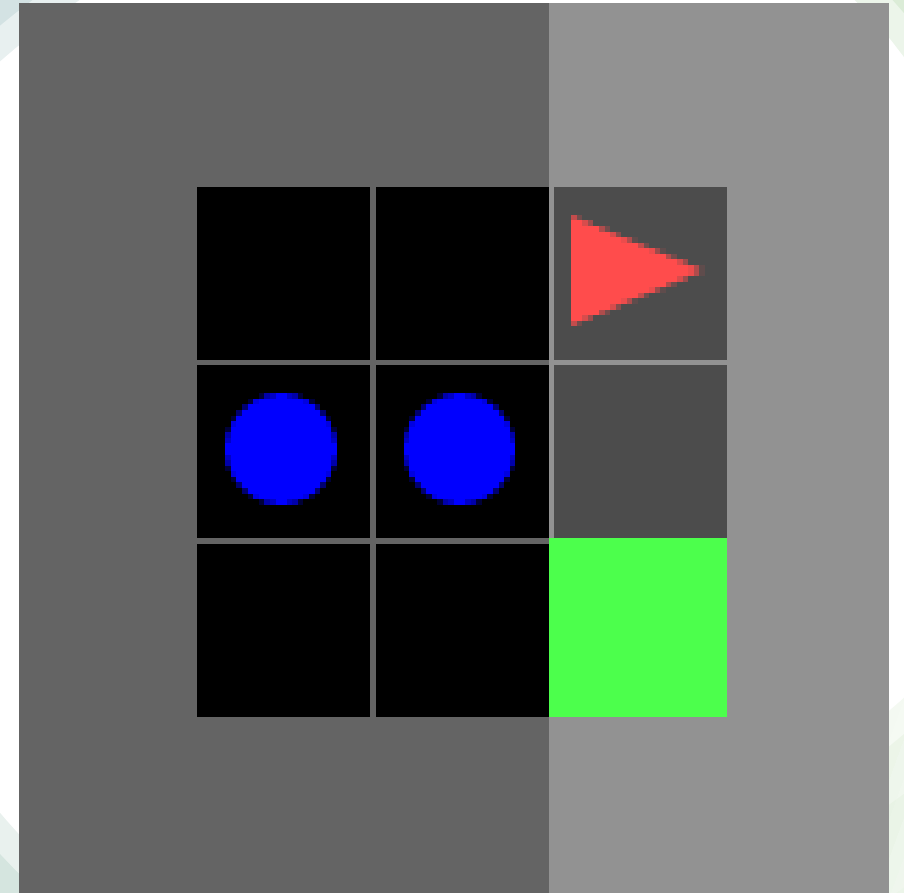
Julien Bertrand n°:20221132
Baptiste Coeffier n°:20202553

https://github.com/baptiste-coeffier/Minigrid_moving_Obstacles/tree/main



Le problème

- Environnement
- But
- Action Space
- Terminaisons
 - But atteint
 - Collision
 - Timeout
- Récompense
 - Succès: $1 - 0.9 * (\text{step_count} / \text{max_steps})$
 - Collision: -1
 - Echec: 0



Modélisation

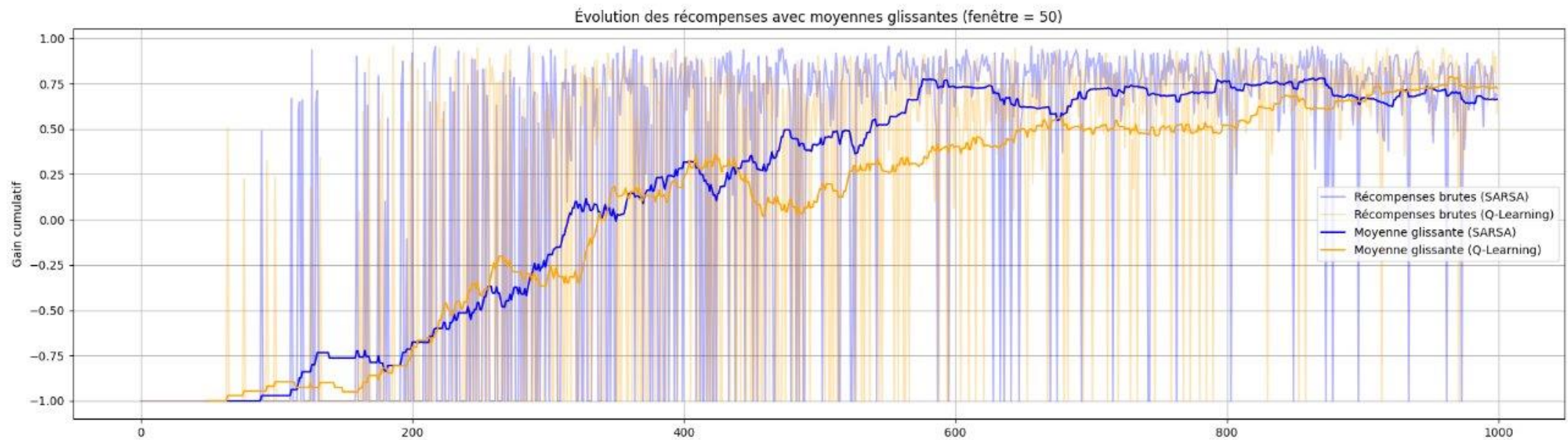
SARSA

- Initialisation
 - $Q(s,a)$
- Boucle sur les épisodes
 - Réinitialisation env
 - Choix première action (epsilon-greedy)
- Boucle sur les transitions
 - Exécute l'action, observe la récompense r , l'état s' , et l'action suivante a'
- Adaptation dynamique
 - Diminution epsilon -> moins d'exploration au fil du temps
 - Décroissance optionnelle de alpha -> stabilisation apprentissage

Q-Learning

- Initialisation
 - $Q(s,a)$
- Boucle sur les épisodes
 - Réinitialisation env
 - Choix première action (epsilon-greedy)
- Boucle sur les transitions
 - Exécute l'action, observe la récompense r , l'état s'
 - Mise à jour de $Q(s,a)$ en utilisant la meilleure action de s'
- Adaptation dynamique
 - Diminution epsilon -> moins d'exploration au fil du temps
 - Décroissance optionnelle de alpha -> stabilisation apprentissage

Résultat





Conclusion

Bon apprentissage sur le problème 5*5

Augmentation de la taille de la grille fait chuter l'apprentissage

Apprentissage non modulaire