

Création de différentes IA en utilisant une architecture I2A

Dans le contexte du jeu Sokoban

Thomas Guyomard - Jérémy Tremblay

Université du Littoral Côte d'Opale

Encadrant : Jérôme Buisine

2024/02/07

Sommaire

- 1 Références
- 2 Présentation du jeu Sokoban
- 3 Agents Augmentés par l'Imagination
- 4 IA Q-learning avec une représentation en tableau
- 5 Ajout d'un mécanisme de renforcement
- 6 Implémentation du Q-learning
- 7 Intégration du système de récompense
- 8 Résultats
- 9 Plan de Travail

Sébastien Racanière, Théophane Weber, David P. Reichert, Lars Buesing, Arthur Guez, Danilo Rezende, Adria Puigdomènech Badia, Oriol Vinyals, Nicolas Heess, Yujia Li, Razvan Pascanu, Peter Battaglia, Demis Hassabis, David Silver, Daan Wierstra, "Imagination-Augmented Agents for Deep Reinforcement Learning", DeepMind, arXiv:1707.06203v2 [cs.LG] 14 Feb 2018

Présentation du jeu Sokoban

Jeu de **réflexion**. Le but : le joueur doit ranger des caisses sur des cases cibles.

- Le personnage a la capacité d'effectuer des déplacements dans chacune des 4 directions possibles. Le niveau est validé une fois toutes les caisses rangées.

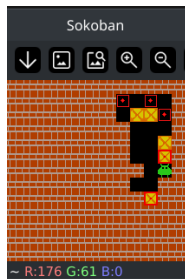


Figure: Capture d'écran du jeu Sokoban.

Definition

L'apprentissage par renforcement avec augmentation par l'imagination appelée I2A est une approche du domaine de l'intelligence artificielle permettant d'améliorer la prise de décision des agents d'apprentissage. Les agents combinent l'apprentissage automatique avec et sans modèles avec des générations de scénarios pour anticiper des situations inédites et proposer des solutions optimales.

IA Q-learning avec une représentation en tableau

Definitions

L'implémentation combine le Q-learning avec un modèle interne I2A pour résoudre le problème complexe du Sokoban, un jeu de réflexion.

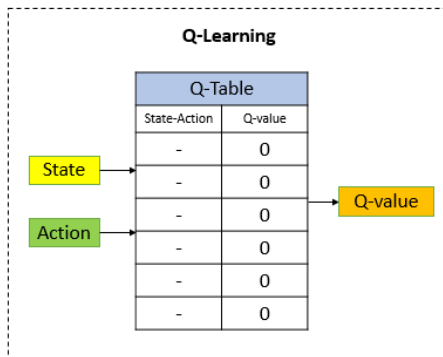


Figure: Schéma du fonctionnement du Q-Learning.

- **Application du renforcement au jeu du Sokoban**
 - L'agent (IA) apprend à jouer au Sokoban en interagissant avec l'environnement.
 - Les récompenses définies pour différentes actions guident l'agent vers des comportements souhaités.
- **Mise en application du Q-Learning**
 - La Q-table est mise à jour itérativement en fonction des récompenses obtenues.
 - L'agent utilise la Q-table pour choisir ses actions et optimiser sa politique de choix au fil du temps.

Implémentation du Q-learning

- **Q-table** Contient les états / actions du jeu, elle est initialisée avec des zéros.
- **Boucle d'apprentissage** Itérations sur des épisodes pour mettre à jour la Q-table.
- **Système de récompenses** Intégration des récompenses conformément à la politique définie.
- **Mise à jour de la Q-table** Prise en compte de la récompense immédiate et de l'estimation de la récompense future.
- **Exploration et Exploitation** Equilibre les connaissances actuelles et permet davantage d'apprentissage.

Intégration du système de récompense

Objectif

Introduire un système de récompenses permettant de guider l'agent vers les actions optimales grâce à des récompenses positives et des pénalités.

Action	Récompense
Effectuer une action	-0.1
Pousser une boîte sur la cible	1.0
Pousser une boîte hors de la cible	-1.0
Achever le niveau	10.0

Pour aller plus loin...

Pénalités en cas de caisses bloquées.

- **Un temps d'exécution élevé** Concernant la génération des niveaux.
- **Des performances faibles...** Valeurs des récompenses négatives à la fin de l'exécution et constantes: -300.
- **Des niveaux différents** Rendant l'apprentissage caduc.
- **Représentation de la partie** On cherche à récupérer un visuel du jeu afin de réaliser un I2A et d'améliorer le système de récompenses.

Plan de Travail

1	Résolution des problèmes de performances et apprentissage plus poussé.
2	Comparaison des performances avec des modèles comme MCTS pour évaluer l'efficacité relative.
3	Si possible, explorer la possibilité d'utiliser les représentations visuelles de Gym Sokoban pour enrichir la planification.
4	Tester l'agent sur différents niveaux afin d'évaluer sa capacité de généralisation.
5	Analyser les performances de chaque approche, en mettant en évidence les forces et les faiblesses dans le contexte de Sokoban.

Merci pour votre attention.