

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 : M. Jérôme Buisine

4 décembre 2023

Sommaire

- 1 Références
- 2 Présentation du jeu Sokoban
- 3 Environnement d'étude
- 4 Agents Augmentés par l'Imagination
- 5 Fonctionnement des Agents Augmentés par Imagination
- 6 Performances de différents modèles
- 7 Comparaison d'efficacité de l'imagination
- 8 Q-learning avec une représentation en tableau
- 9 Deep Q-Learning avec un réseau de neurones
- 10 Parallèle entre Sokoban et les I2A
- 11 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 joueur doit ranger des caisses sur des cases cibles.

- Le personnage peut réaliser neuf opérations :
 - 4 opérations concernent le déplacement dans une case adjacente à lui (droite, gauche, haut, bas).
 - 4 opérations concernent l'action de pousser une caisse sur une case adjacente à lui dans la direction de son mouvement (droite, haut, gauche, bas), sauf si celle-ci est bloquée par un obstacle quelconque.
 - La dernière action vide : ne rien faire.

Une fois toutes les caisses rangées le niveau est validé et ainsi de suite jusqu'à la fin de tout les niveaux.

Environnement de travail

- 1 Python
- 2 Gym
- 3 Plusieurs modes de jeu

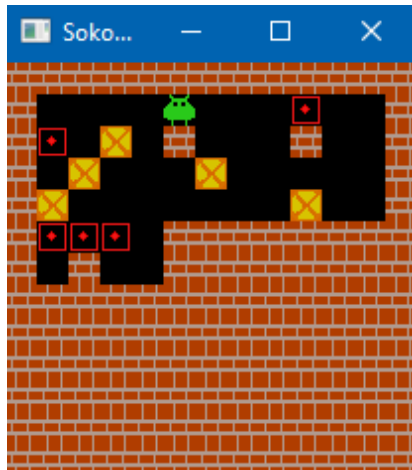


Figure – Capture d'écran du jeu Sokoban

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.

Fonctionnement des Agents Augmentés par Imagination

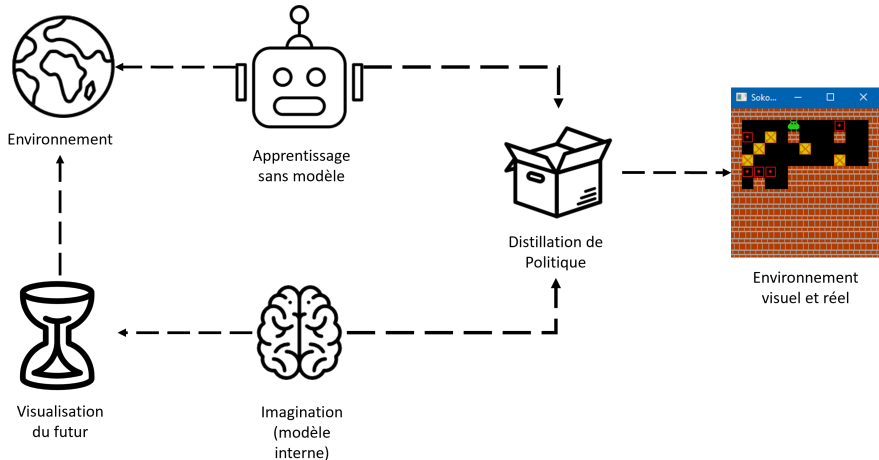


Figure – Schéma récapitulatif de l'approche I2A.

Performances de différents modèles

Informations

Les différents modèles ne performant pas de la même manière au jeu du Sokoban.

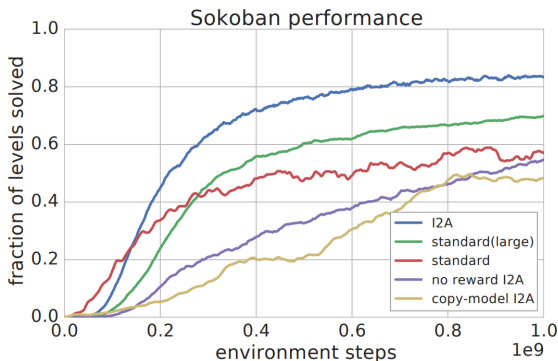


Figure – Schéma représentant les succès de différents modèles au jeu du

Comparaison d'efficacité de l'imagination

Boxes	1	2	3	4	5	6	7
I2A (%)	99.5	97	92	87	77	66	53
Standard (%)	97	87	72	60	47	32	23

Table – Efficacité de l'imagination pour différentes architectures.

I2A@87	~ 1400
I2A MC search @95	~ 4000
MCTS@87	~ 25000
MCTS@95	~ 100000
Random search	~ millions

Table – Généralisation de I2A à des environnements avec différents nombres de boîtes.

Q-learning avec une représentation en tableau

Definitions

Le Q-learning facilite les décisions séquentielles. Grâce à un tableau, chaque cellule stocke la valeur Q, qui est mise à jour lorsqu'un agent prend une action dans un état donné qui aide ainsi à l'amélioration petit à petit de la politique de prise de décision.

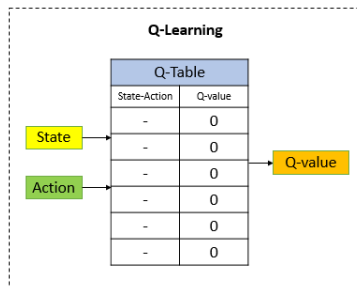


Figure – Schéma du fonctionnement du Q-Learning.

Deep Q-Learning avec un réseau de neurones

Definitions

Le **Deep Q-Learning** est une extension de Q-Learning qui utilise des réseaux de neurones profonds pour gérer des d'états et d'actions plus importants que le Q-Learning classique.

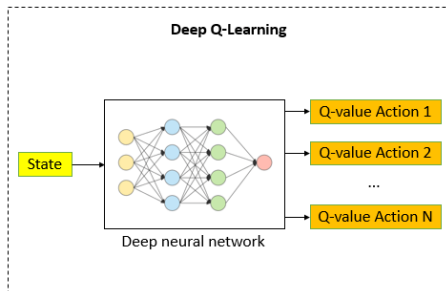


Figure – Schéma du fonctionnement du Deep Q-Learning.

Objectifs

Objectif : entraîner un agent étant en capacité de jouer et d'évoluer dans le jeu du Sokoban. Problématique : Comment l'I2A peut-elle être appliquée pour guider les actions dans un environnement complexe ?

- 9 actions mais une infinités de cartes de tailles différentes et parfois plusieurs solutions possibles.
- Il faut adapter les concepts de l'I2A au jeu du Sokoban afin d'obtenir un modèle performant.
- Plan d'action :
 - Distillation de la politique pour combiner apprentissage sans modèle et avec modèle.
 - Utilisation de l'imagination pour guider les actions dans l'environnement.
 - Ajout de visuels de l'environnement du jeu avec traitement d'image afin d'améliorer les résultats.

Plan de Travail

1	Compréhension et test de l'environnement Gym Sokoban.
2	Recherche et compréhension des différentes approches (Q-learning, MCTS, Deep Q-Learning).
3	Développement de différents modèles d'apprentissage automatique (Q-learning, Deep Q-learning...).
4	Utilisation de l'approche I2A dans notre jeu.
5	Comparaison des performances avec des modèles comme MCTS pour évaluer l'efficacité relative.
6	Si possible, explorer la possibilité d'utiliser les représentations visuelles de Gym Sokoban pour enrichir la planification.
7	Tester l'agent sur différents niveaux afin d'évaluer sa capacité de généralisation.
8	Analyser les performances de chaque approche, en mettant en évidence les forces et les faiblesses dans le contexte de Sokoban.

Merci pour votre attention.