

ÉVALUATION ET AMÉLIORATION D'UNE POLITIQUE

1 Présentation du problème

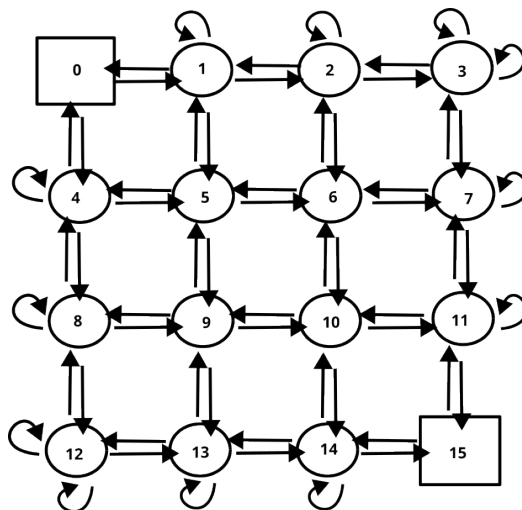
Nous allons appliquer les algorithmes d'*évaluation de la politique*, d'*amélioration de la politique* et d'*amélioration de la valeur* vus lors de la séance de cours précédente. Pour cela, vous disposez d'un fichier, permettant de générer des exemples de grilles semblables à l'exemple du cours, où un agent se déplace dans une grille carré jusqu'à atteindre l'un des 2 coins noté *TS* (*terminal state*). Cette grille pourra prendre des tailles variées.

Ainsi pour une grille de taille 4, il y aura :

- 14 états non-terminaux
- 2 états terminaux

Pour rappel

- Actions qui quittent la grille reviennent au même état
- Récompense de -1 pour toute arrivée dans un état non-terminal, 0 sinon
- Politique uniforme dans 4 directions



2 Objectifs

Votre objectif est d'implémenter ces 3 algorithmes et de les tester sur des tailles de grille de plus en plus importantes. Parallèlement, vous devrez proposer des méthodes de vérifications des résultats donnés pour des grilles de faibles tailles (3,4,..., 10) aussi bien en terme de *value function* associée à une politique que de politique optimale. Ces méthodes faisant doublon avec les algorithmes (*évaluation de la politique*, *amélioration de la politique*), vous devrez expliquer quand on devra faire usage des unes ou des autres.

Pour ce TP, vous disposez d'un environnement codé sous forme de dictionnaire. Au besoin, vous pourrez rassembler vos méthodes sous la forme d'une classe.

3 Rendus

Dans un fichier .py, le code des méthodes suivantes

- d'évaluation de la politique (*Policy Evaluation*)
- d'amélioration de la politique (*Policy Iteration*)
- d'amélioration de la valeur (*Value Iteration*)

Ainsi qu'un compte rendu en notebook.