

# Navigation d'un robot

Le but de ce projet est de mieux comprendre les processus décisionnels markoviens (PDM en anglais) et l'algorithme d'itération sur les valeurs au travers de la recherche d'une politique optimale pour un robot dans un labyrinthe.

## 1 Présentation du problème

Un agent-robot se trouve en un point A sur une grille (figure 1), muni de quatre actions de déplacement (Haut, Bas, Gauche et Droite). Il doit aller au point B.

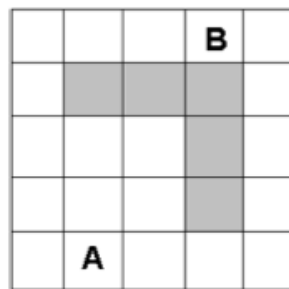


FIGURE 1 – Le robot est positionné en A. Il souhaite atteindre la position B. Les "marécages" sont matérialisés en gris.

Rappelons qu'un MDP est défini par le tuple  $\langle S, A, P, R, T \rangle$  :

- S : ensemble des états de l'agent ;
- A : ensemble des actions ;
- P : matrice de transition entre les différents états. Il y a beaucoup de vent par conséquent lorsque l'agent décide de se déplacer dans une direction il y a 80% de chance pour qu'il se retrouve dans l'état visé et 20% de chance pour qu'il se retrouve d'un côté (10%) ou de l'autre (10%) de l'état visé. Cas particulier, si l'agent est sur un bord et qu'il prévoit de se déplacer parallèlement au bord, il a 90% de chance de se retrouver dans l'état visé et 10% de chance de se retrouver à coté de l'état visé ;
- R : fonction de récompense. Elle est définie comme suit :
  - R = +5 : lorsque l'agent atteint le point B.
  - R = -1 : lorsque l'agent passe sur un marécage.
- T : horizon, considéré comme infini.

Pour faciliter la mise en place de notre MDP, nous considérerons que les politiques sont déterministes, c'est-à-dire que pour chaque état il n'y a qu'une seule action possible.

## 2 Travail à réaliser

### 2.1 Réalisation

Vous devez implémenter l'algorithme d'itération sur les valeurs afin de comparer l'effet du coefficient  $\gamma$  sur le choix de la décision. Pour démarrer vous pouvez prendre  $\gamma = 0.9$  et  $\epsilon = 0.01$ .

**Contraintes :**

Information à afficher :

- la valeur  $V$  pour chaque état ;
- l'action retenu pour chaque état ;
- Le chemin optimal.

Paramètres modifiables :

- la valeur  $\gamma$  ;
- la récompense lorsque l'agent atteint le point B ;
- la récompense lorsque l'agent passe sur un marécage ;
- la position des marécages ;
- les positions de départ et d'arrivée.

### 2.2 Présentation

Les soutenances auront lieu les 11 et 12 mai 2023 en fonction de vos disponibilités. Une présentation d'environ cinq minutes suivie d'une discussion d'une dizaine de minutes vous permettront d'une part d'exposer vos résultats en précisant leurs points forts, leurs points faibles, et les améliorations possibles et d'autre part d'effectuer une démonstration de l'implantation de cette stratégie. Le support de la présentation et les sources doivent être déposés sur le site arche dédié au module Intelligence Artificielle au plus tard le mercredi 10 mai 2023. Merci de nommer votre archive avec le nom du projet et les vôtres (ex. robot-Nom1-Nom2-Nom3.zip).