

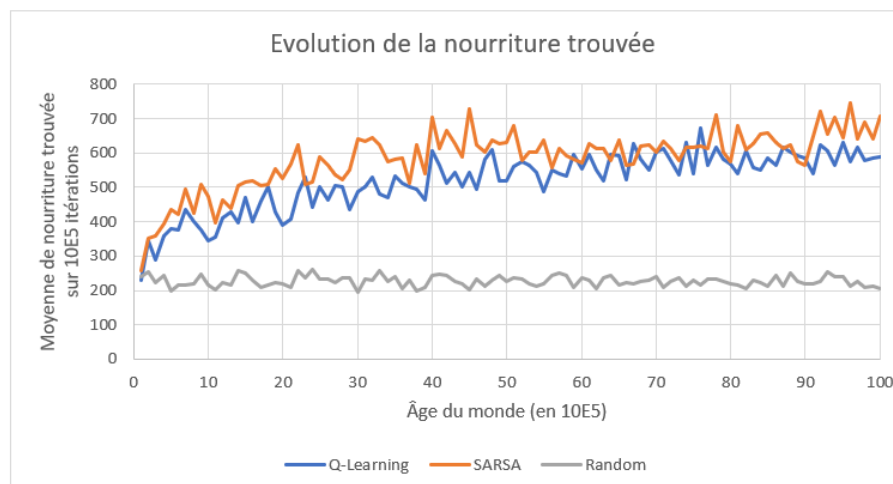
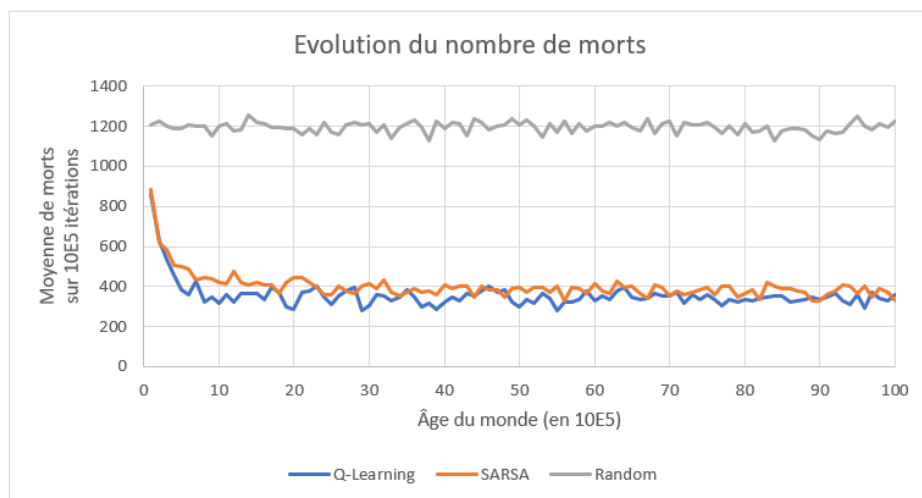
Projet 2 : Apprentissage renforcé

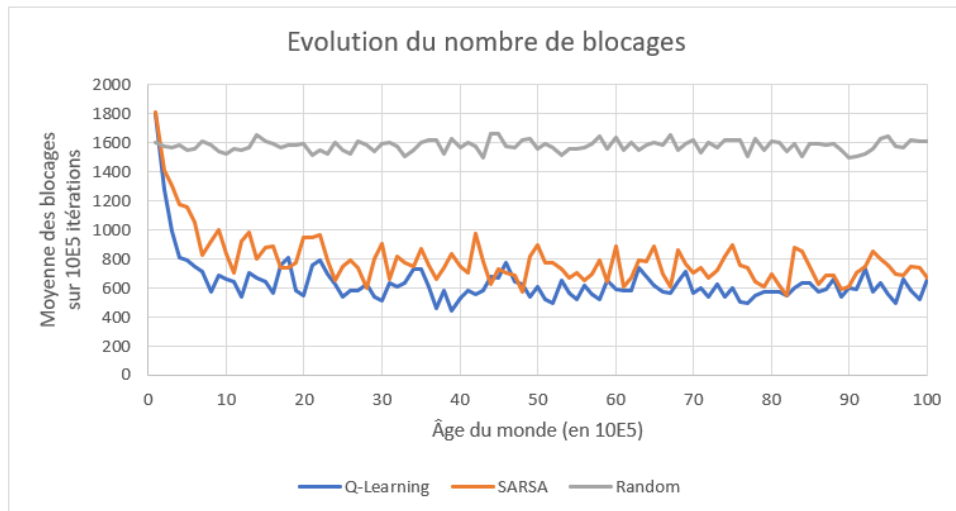
Le projet que nous avons réalisé est constitué de différents algorithmes d'apprentissage renforcé. En première partie, nous avons appliqué un apprentissage basé sur la résolution de l'équation de Bellman. En seconde partie, nous avons effectué un apprentissage de type Q-Learning.

Pour la méthode Bellman, notre environnement est une grille de récompense 2D où un agent peut se déplacer dans 4 directions différentes ou rester immobile. Notre implémentation contient les 2 méthodes de résolution de l'équation de Bellman : la résolution exacte par matrices et la résolution itérative. On trouvera un tableau décrivant l'évolution de la politique de la première grille proposée par le squelette en annexe.

Nous avons aussi implémenté une action jumpF, qui à 2 issues possibles. On trouvera en annexe l'évolution de la politique de la grille 2 avec cette action de disponible.

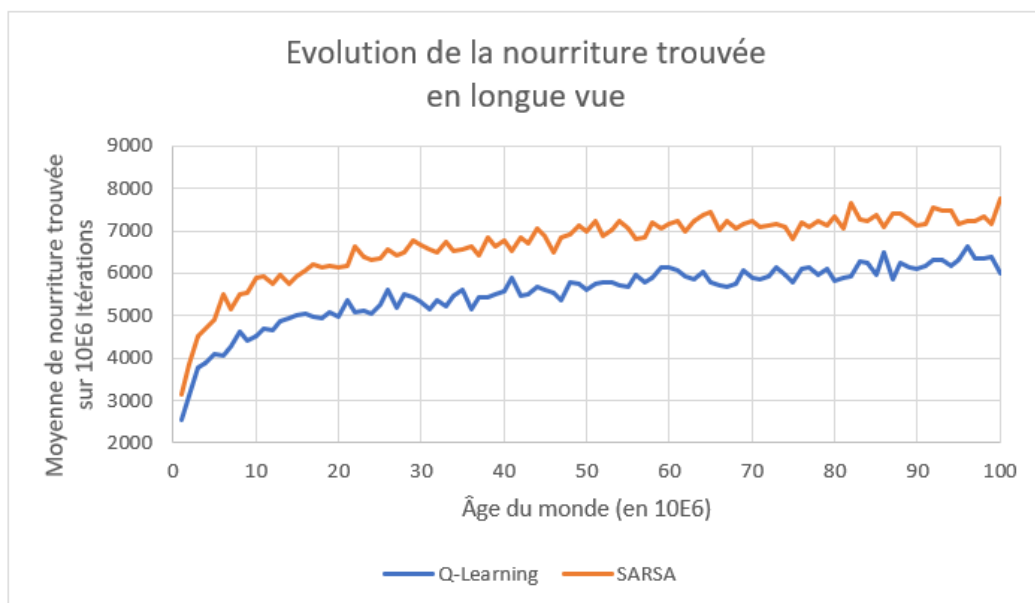
Pour le Q-Learning, notre support est un pac-man. On a un tableau de jeu dynamique, où l'agent peut se déplacer dans les 8 cases qui lui sont adjacentes. Notre agent ne possède d'information que sur les 11 cases les plus proches de lui. Le plateau contient dans une case soit de la nourriture, soit un fantôme qui peut manger notre agent, soit rien du tout.





Les graphes ci-dessus ont été tracés avec les valeurs de gamma, epsilon, alpha, ainsi que celles de toutes les récompenses par défaut. Pour ce modèle, nous avons implémenté les méthodes Q-Learning et la méthode SARSA.

Nous avons aussi implémenté le pac-man avec un champ de vision augmenté à 3 cases de distance, soit une vue sur 23 cases. Les graphes sont assez similaires à ceux déjà tracés, du moins lorsqu'on utilise la même échelle. En augmentant le nombre d'itération, on voit que la méthode SARSA profite plus de l'information supplémentaire. Ceci est logique car cet algorithme est plus basé sur l'exploration.



Ce projet met donc en lumière les différents usages des méthodes d'apprentissage par renforcement vues en cours. Il est important de choisir un outil adapté au problème que l'on souhaite résoudre et à nos différents paramètres. Par exemple, on choisira plutôt Bellman pour un problème statique comme la grille, et du Q-Learning pour un problème dynamique comme le pac-man. Ceci reste vrai pour les méthodes à l'intérieur des méthodes, par exemple discriminer SARSA et la Q-Learning.

Annexe 1 : Politique Grille num_g = 0

Itérations	1	2	3	4	5
Etats					
0	stay	stay	stay	stay	stay
1	left	left	left	left	left
2	stay	left	left	left	left
3	stay	stay	left	left	left
4	stay	stay	stay	left	left
5	stay	stay	right	right	right
6	stay	right	right	right	right
7	right	right	right	right	right
8	up	up	up	up	up
9	stay	left	left	left	left
10	stay	stay	left	left	left
11	stay	stay	stay	left	left
12	stay	stay	stay	stay	left
13	stay	stay	stay	right	right
14	stay	stay	right	right	right
15	stay	up	up	up	up
16	stay	up	up	up	up
17	stay	stay	left	left	left
18	stay	stay	stay	stay	stay
19	stay	stay	stay	stay	stay
20	stay	stay	stay	stay	stay
21	stay	stay	stay	stay	stay
22	stay	stay	stay	right	right
23	stay	stay	up	up	up
24	stay	down	down	down	down
25	stay	stay	down	down	down
26	stay	stay	stay	down	down
27	stay	stay	stay	stay	down
28	stay	stay	stay	stay	stay
29	stay	stay	stay	stay	right
30	stay	stay	stay	right	right
31	stay	stay	down	down	down
32	down	down	down	down	down
33	stay	left	left	left	left
34	stay	stay	left	left	left
35	stay	stay	stay	left	left
36	stay	stay	stay	stay	left
37	stay	stay	stay	right	right
38	stay	stay	right	right	right
39	stay	right	right	right	right

Annexe 2 : Politique Grille num_g = 1 avec jumpF

Itérations	1	2	3	4	5
Etats					
0	down	down	down	down	down
1	stay	down	down	down	down
2	stay	stay	stay	stay	stay
3	stay	stay	down	jumpF	jumpF
4	stay	stay	right	right	right
5	stay	down	down	down	down
6	stay	stay	stay	stay	stay
7	left	left	left	left	left
8	stay	stay	stay	stay	stay
9	stay	jumpF	jumpF	jumpF	jumpF
10	stay	stay	stay	stay	stay
11	right	right	right	right	right
12	stay	stay	stay	stay	stay
13	stay	up	up	up	up
14	stay	stay	left	left	left
15	stay	stay	up	left	up
16	stay	stay	stay	stay	stay
17	stay	up	up	up	up
18	stay	stay	stay	stay	stay
19	stay	stay	up	up	up
20	stay	stay	stay	stay	stay
21	stay	stay	stay	up	up
22	stay	stay	stay	stay	stay
23	stay	stay	up	up	up
24	stay	stay	stay	stay	stay
25	stay	stay	down	down	down
26	stay	stay	stay	stay	stay
27	stay	stay	stay	stay	stay
28	stay	stay	stay	right	right
29	stay	stay	jumpF	jumpF	jumpF
30	jumpF	jumpF	jumpF	jumpF	jumpF
31	stay	left	left	left	left
32	stay	stay	left	left	left
33	stay	stay	stay	stay	stay
34	stay	stay	stay	stay	stay
35	stay	stay	stay	stay	stay

D'autres tableaux de ce type peuvent être construit à partir de notre programme en important les fichiers texte construits sur un tableur.