

Rapport projet DRL

- BENDAVID Natane
- TARDY Louis
- WADE Cheikh Abdourahmane

Dans ce rapport, vous trouvez l'intégralité des algos que nous avons implementé ainsi que les resultat obtenues pour chaq'un des environnement testé sur les algos

Dynamic programming

LINE WORD

GRID WORD

SECRET ENV 1

Monte carlo methods

TIC TAC TOE

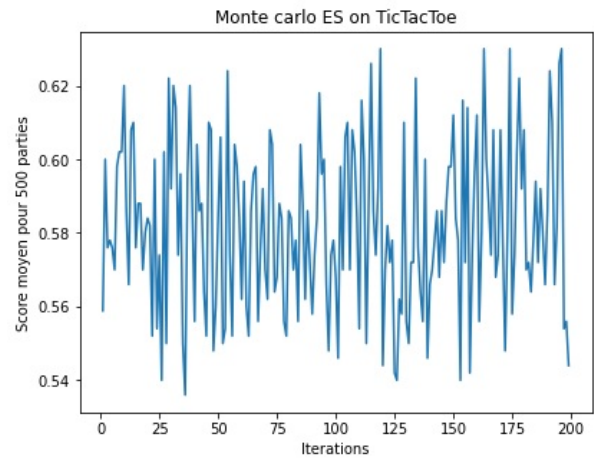
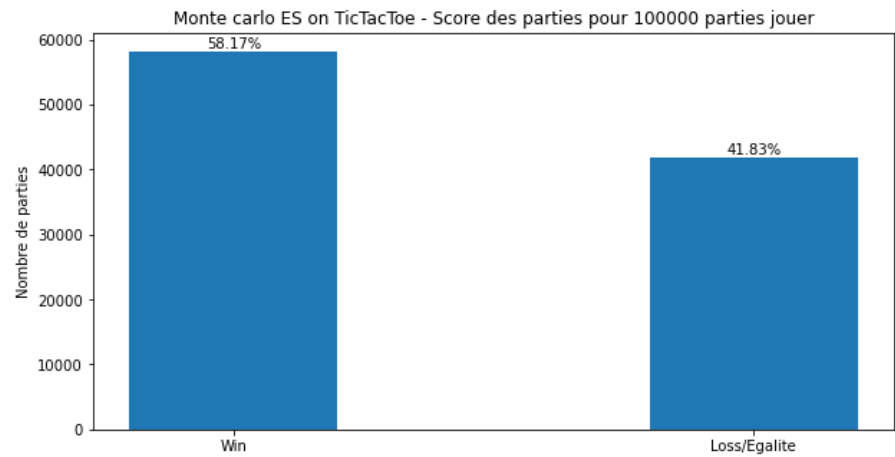
MONTE CARLO ES

Nous avons testé L'algo Monte Carlo ES sur l'env TIC TAC TOE avec un gama de 0.99

Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 58.17% de partie gagner
- 41.83% de partie perdue ou égalité

Voici les graphiques :



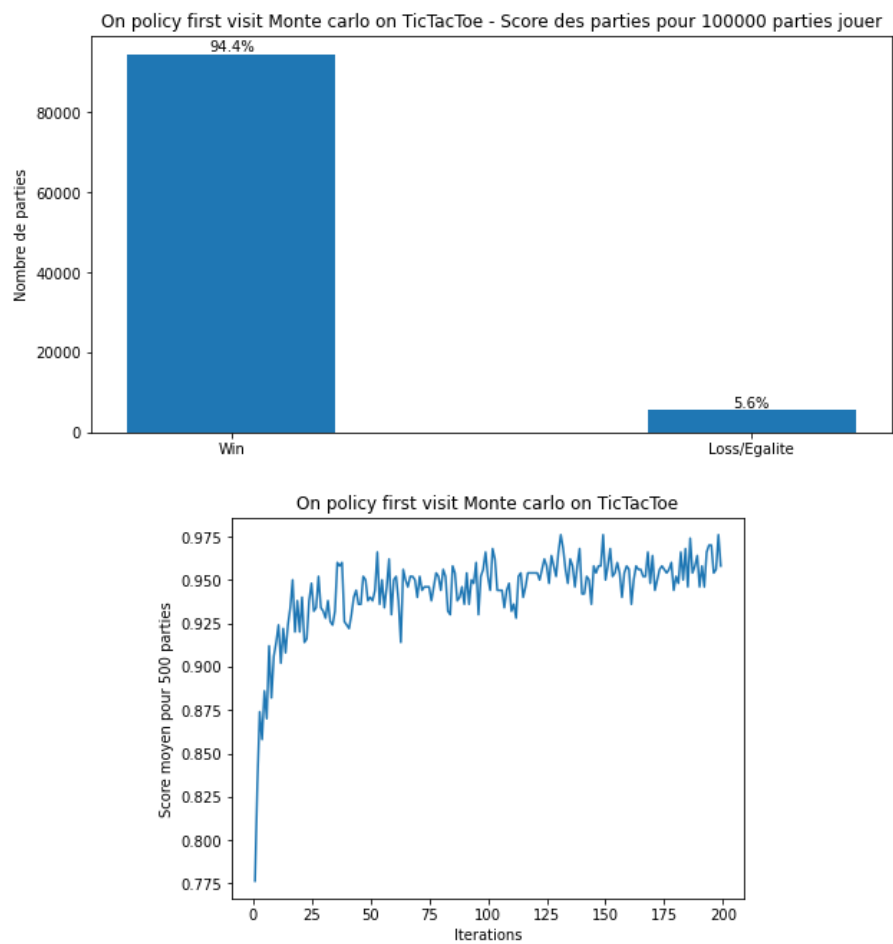
On policy first visit monte carlo control

Nous avons testé L'algo On policy first visit monte carlo control sur l'env TIC TAC TOE avec un gama de 0.99 et un epsilon de 0.1

Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 94.4% de partie gagner
- 5.6% de partie perdue ou égalité

Voici les graphiques :



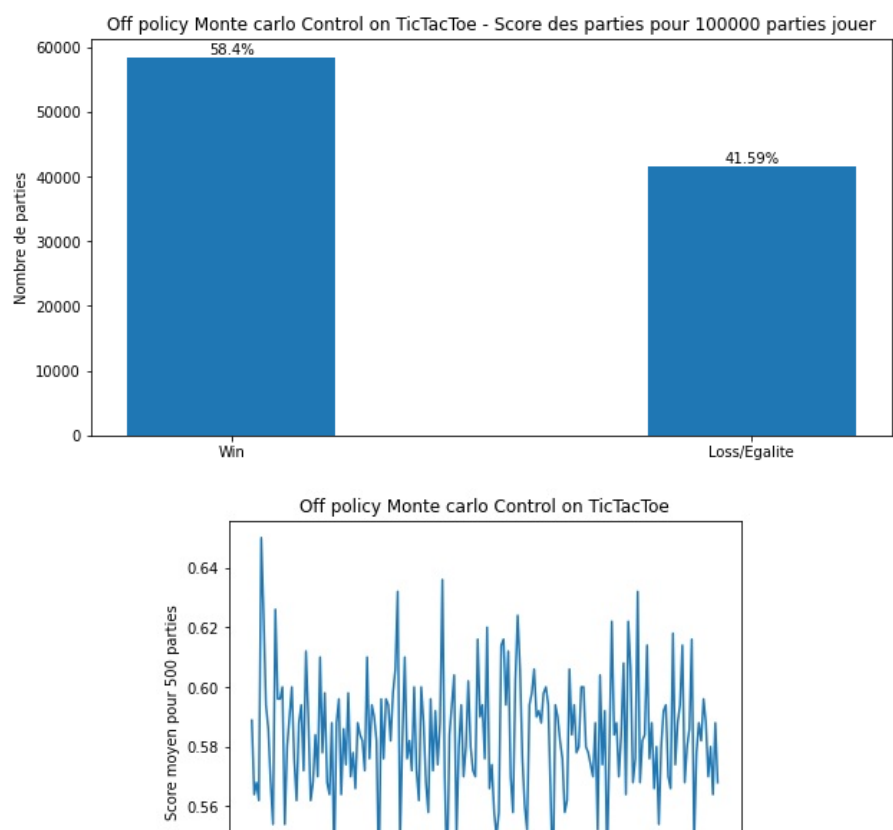
Off policy monte carlo control

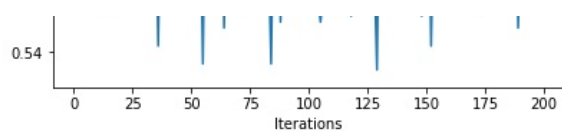
Nous avons testé l'algo Off policy monte carlo control sur l'env TIC TAC TOE avec un gama de 0.99

Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 58.4% de partie gagner
- 41.59% de partie perdue ou égalité

Voici les graphiques :





SECRET ENV 2

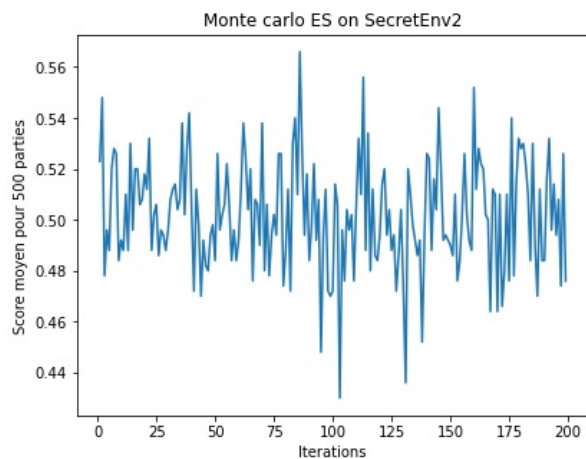
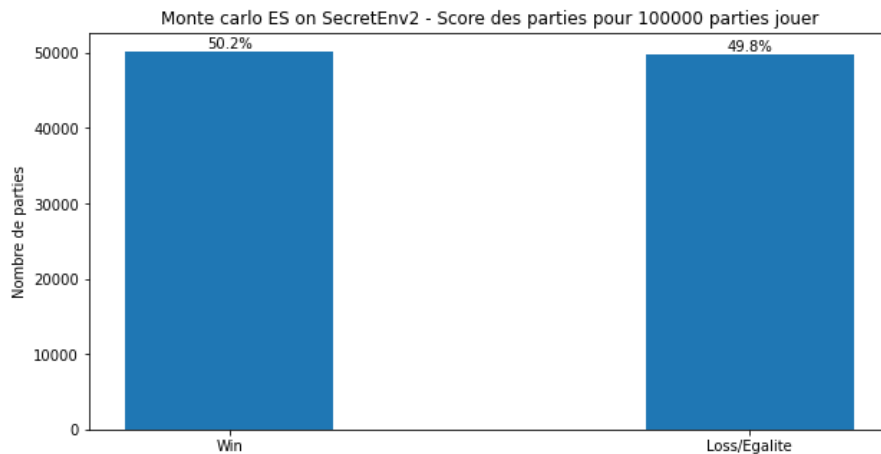
MONTÉ CARLO ES

Nous avons testé L'algo Monte Carlo ES sur l'env Secret ENV 2 avec un gama de 0.99

Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 50.2% de partie gagner
- 49.8% de partie perdue ou égalité

Voici les graphiques :



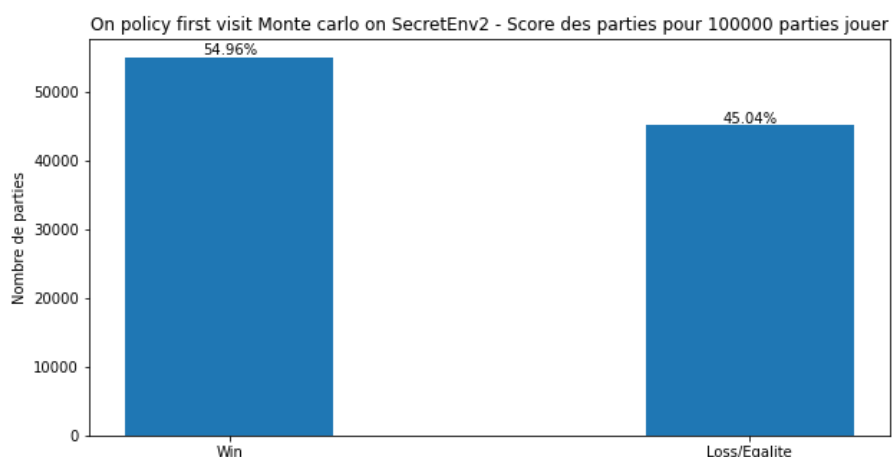
On policy first visit monte carlo control

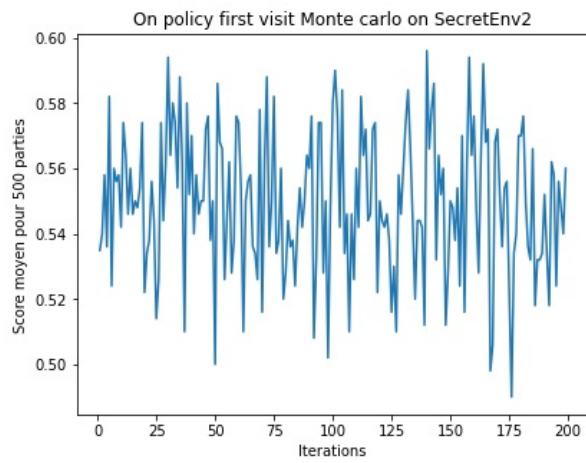
Nous avons testé L'algo On policy first visit monte carlo control sur l'env SECRET ENV 2 avec un gama de 0.99 et un epsilon de 0.1

Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 54.96% de partie gagner
- 45.04% de partie perdue ou égalité

Voici les graphiques :





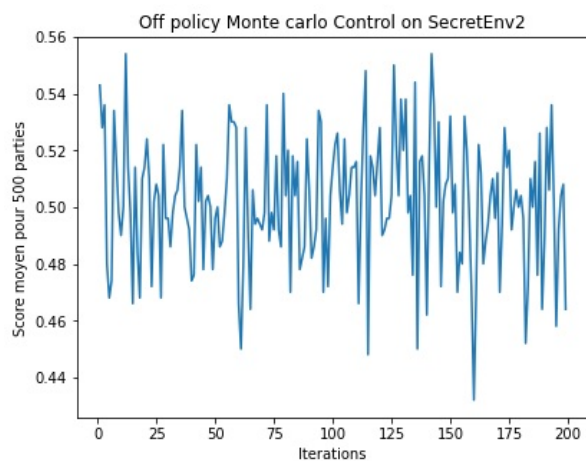
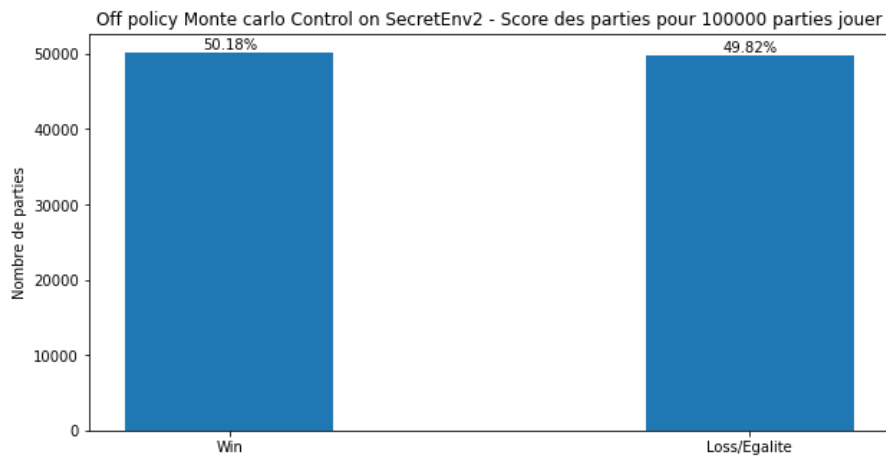
Off policy monte carlo control

Nous avons testé l'algo Off policy monte carlo control sur l'env SECRET ENV 2 avec un gama de 0.99

Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 50.18% de partie gagner
- 49.82% de partie perdue ou égalité

Voici les graphiques :



Temporal difference learning

TIC TAC TOE

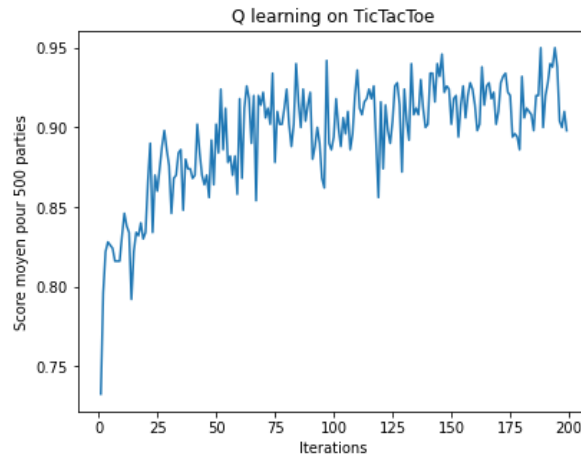
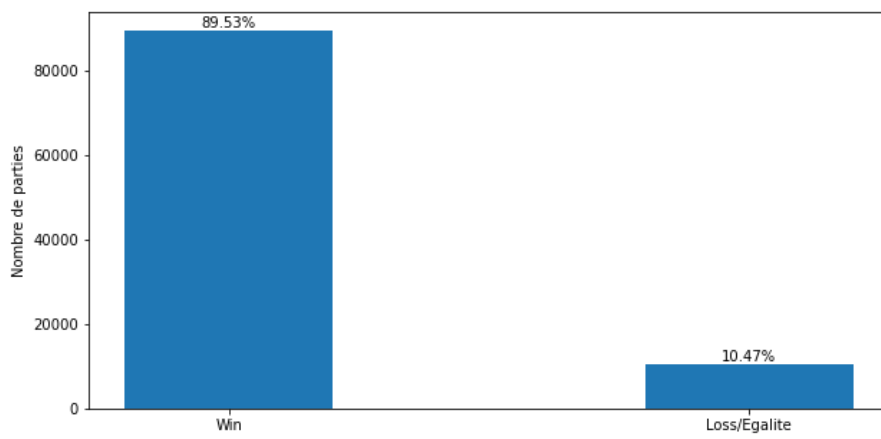
Q_learning

Nous avons testé L'algo Q_learning sur l'env TIC TAC TOE avec un alpha de 0.7, un epsilon de 0.1 et un gamma de 0.9

Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 89.53% de partie gagner
- 10.47% de partie perdue ou égalité

Voici les graphiques :



Expected Sarsa

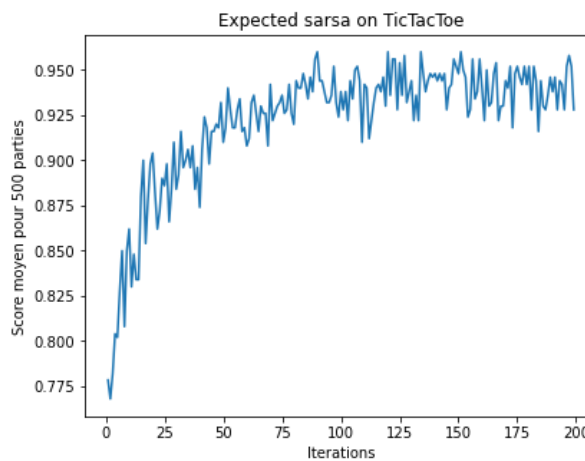
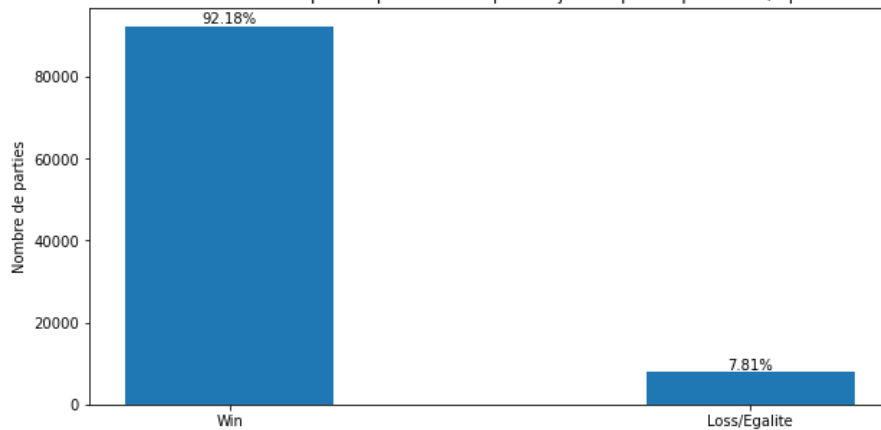
Nous avons testé L'algo Expected Sarsa sur l'env TIC TAC TOE avec un alpha de 0.7, un epsilon de 0.1 et un gamma de 0.9

Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 92.18% de partie gagner
- 7.82% de partie perdue ou égalité

Voici les graphiques :

Expected sarsa on TicTacToe - Score des parties pour 100000 parties jouées pour alpha = 0.7, epsilon = 0.1, gamma = 0.9



SECRET ENV 3

Q_learning

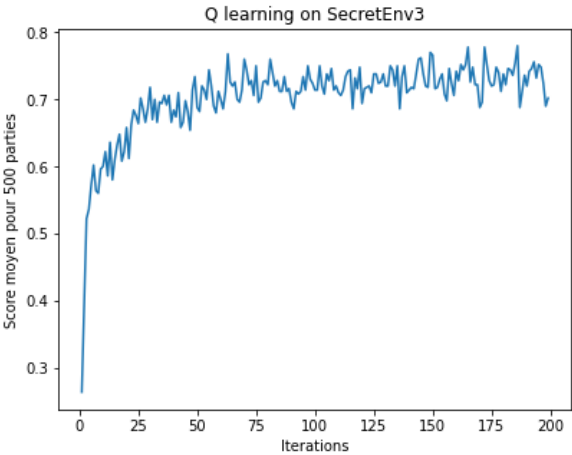
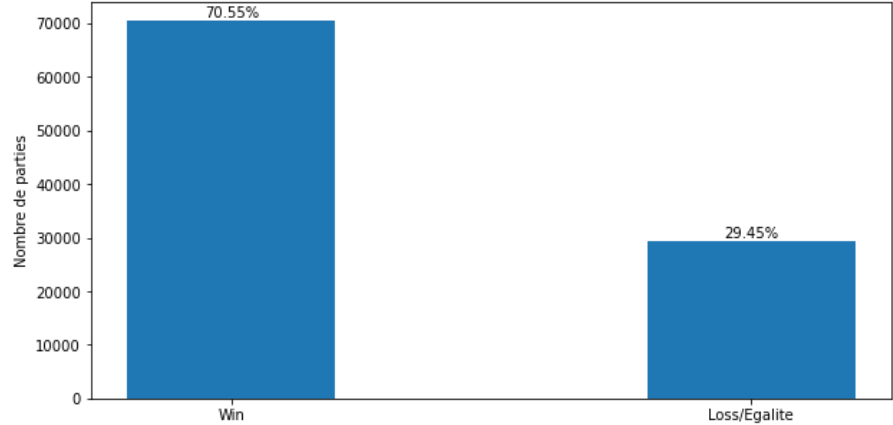
Nous avons testé L'algo Q_learning sur l'env Secret env 3 avec un alpha de 0.3, un epsilon de 0.1 et un gamma de 0.9

Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 70.55% de partie gagner
- 29.45% de partie perdue ou égalité

Voici les graphiques :

Q learning on SecretEnv3 - Score des parties pour 100000 parties jouées pour alpha = 0.3, epsilon = 0.1, gamma = 0.9



Expected Sarsa

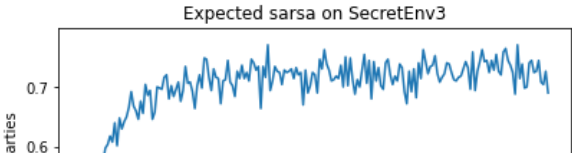
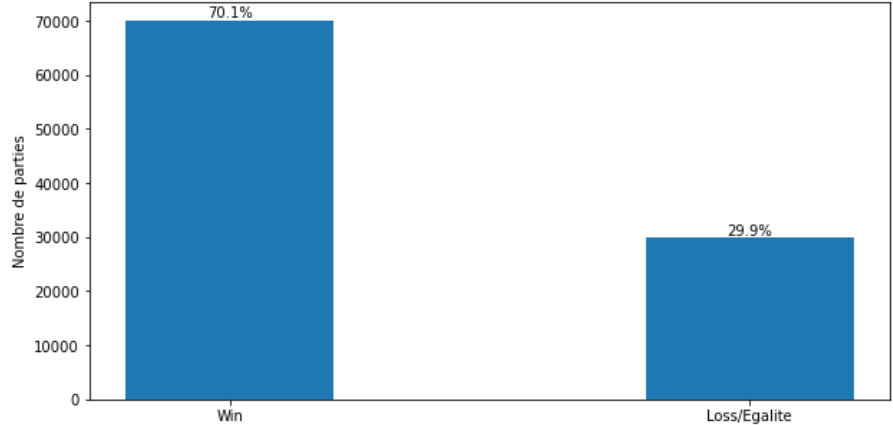
Nous avons testé L'algo Expected Sarsa sur l'env SecretEnv3 avec un alpha de 0.3, un epsilon de 0.1 et un gamma de 0.7

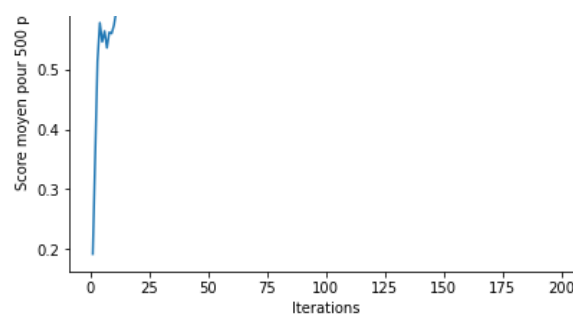
Nous avons obtenue les resultats suivant pour 100 000 partie jouer :

- 70.1% de partie gagner
- 29.9% de partie perdue ou égalité

Voici les graphiques :

Expected sarsa on SecretEnv3 - Score des parties pour 100000 parties jouées pour alpha = 0.3, epsilon = 0.1, gamma = 0.7





Deep reinforcement learning

TIC TAC TOE

Episodic semi gradient sarsa

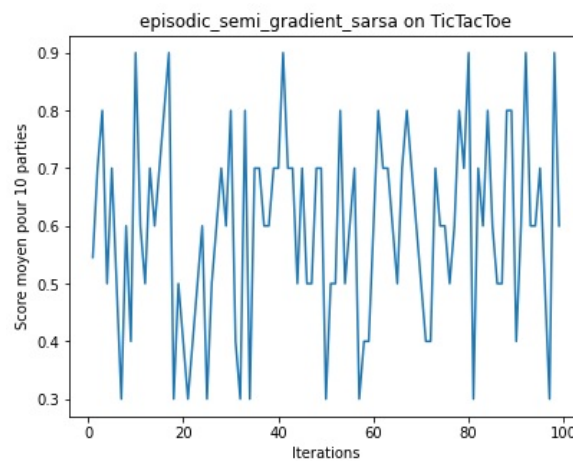
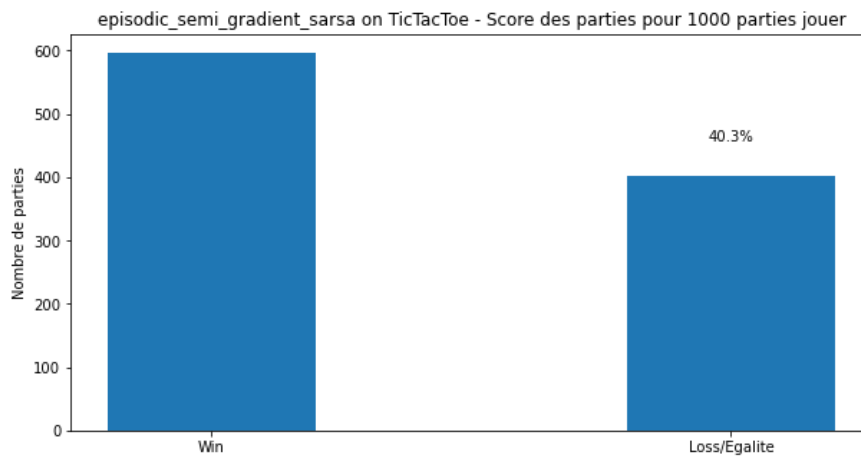
Nous avons testé L'algo Episodic semi gradient sarsa sur l'env TIC TAC TOE avec un epsilon de 0.1 et un gamma de 0.9.

Nous avons créer un model avec 6 couche Dense de : 16, 64, 128, 256, 128, 1 couches du modèle

Nous avons obtenue les resultats suivant pour 1000 partie jouer :

- 59.7% de partie gagner
- 40.3% de partie perdue ou égalité

Voici les graphiques :



SECRET ENV 5

Episodic semi gradient sarsa

Nous avons testé L'algo Episodic semi gradient sarsa sur l'env TIC TAC TOE avec un epsilon de 0.1 et un gamma de 0.9.

Nous avons créer un model avec 6 couche Dense de : 16, 64, 128, 256, 128, 1 couches du modèle

