

Le manuel développeur – Projet MN-IA

KIRLOS Youssef et DIRANI SAFAWI Hani

Execution:

L'exécution commence par l'exécution du programme principal 'main' dans la classe **App**, qui demande de choisir l'exemple à lancer, soit TicTacToe, soit Simple. Pour l'exemple choisi, le programme demande de choisir un algorithme à exécuter :

- **Simple:** Epsilon-Greedy, UCB ou Bandit Gradient pour Simple.
- **TicTacToe:** Value Iteration, Policy Iteration ou Q-Learning pour TicTacToe.

Pour ajuster les paramètres :

Pour 'Simple' :

Les paramètres sont à modifier dans la classe **SimpleRunner** :

- **epsilon** : C'est le taux d'exploration de l'algorithme Epsilon-Greedy. L'agent choisit une action de manière aléatoire avec une probabilité epsilon et exploite la meilleure action avec une probabilité $1 - \epsilon$.
- **alpha** : C'est le taux d'apprentissage, qui détermine dans quelle mesure l'agent met à jour sa fonction de valeur. Un alpha plus grand fait que l'agent apprend plus rapidement.
- **Gamma** : C'est le facteur d'actualisation des récompenses futures. Un gamma proche de 1 valorise fortement les récompenses futures.
- **poidExploration**: Ce paramètre est utilisé pour l'algorithme Upper Confidence Bound (UCB) et contrôle le poids de l'exploration par rapport à l'exploitation.

Pour 'TicTacToe' :

Les paramètres sont directement modifiés dans la classe **App** :

- **epsilon** : Le seuil de convergence pour les algorithmes de Value Iteration, Policy Iteration, ou Q-Learning.
- **gamma** : Le facteur d'actualisation pour ces algorithmes, affectant l'importance des récompenses futures.
- **nbepisodes** : Le nombre d'épisodes à exécuter pour l'algorithme Q-Learning.