

# APPRENTISSAGE PAR RENFORCEMENT

## M2 DAC

### TME 3. Q-Learning

Ce TME a pour objectif d'expérimenter les approches de renforcement value-based vues en cours, notamment l'algorithme Q-Learning (version tabulaire), et les appliquer au problème du GridWorld considéré la semaine dernière. Par contre attention, la commande `getMDP` est maintenant interdite: on se met dans le cas de figure où les dynamiques de l'environnement sont inconnues. Il faut explorer pour apprendre les trajectoires efficaces.

- Implémenter la version basique du Q-learning (tester avec divers paramètres d'exploration)
- Implémenter la version SARSA
- Implémenter l'extension Dyna-Q
- Comparer les résultats sur différentes cartes du jeu (courbes d'apprentissage)
- Bonus: Tester l'implémentation des traces d'éligibilité