

Reinforcement Learning (RL)

Tendances IT

Abdelhakim Zetati – Lahcen Ezzara

zetati.abdelhakim@ensam-casa.ma – ezzara.lahcen@ensam-casa.ma
Université Hassan II de Casablanca
ENSAM Casablanca

October 14, 2024



- ① Introduction au RL
- ② Interface Agent-Environnement
- ③ Fondations Théoriques
- ④ Terminologie du RL
- ⑤ Deep Q-Learning

- ① Introduction au RL
- ② Interface Agent-Environnement
- ③ Fondations Théoriques
- ④ Terminologie du RL
- ⑤ Deep Q-Learning

Comprendre l'Apprentissage par Renforcement

Reinforcement Learning (RL):

- **L'apprentissage par renforcement (RL)** se distingue en apprenant à travers des interactions avec un environnement pour maximiser une fonction de récompense, sans étiquettes correctes prédéfinies, ce qui le rend utile pour la prise de décision dans des environnements complexes.

- 1 Introduction au RL
- 2 Interface Agent-Environnement**
- 3 Fondations Théoriques
- 4 Terminologie du RL
- 5 Deep Q-Learning

L'Interface Agent-Environnement d'un Système d'Apprentissage par Renforcement

L'état de l'agent est composé de ses variables, et il interagit avec l'environnement à travers des actions, recevant des récompenses qui guident ses transitions d'état. Le processus d'apprentissage implique de trouver un équilibre entre l'exploration (essayer de nouvelles actions) et l'exploitation (choisir des actions avec des récompenses connues) pour maximiser les récompenses cumulées au fil du temps.

L'Interface Agent-Environnement d'un Système d'Apprentissage par Renforcement

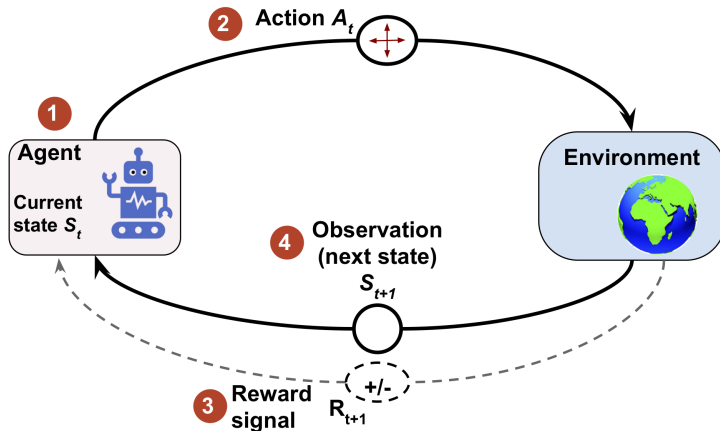


Figure 1: L'Interface Agent-Environnement d'un Système d'Apprentissage par Renforcement

- 1 Introduction au RL
- 2 Interface Agent-Environnement
- 3 Fondations Théoriques**
- 4 Terminologie du RL
- 5 Deep Q-Learning

Les Processus de Décision de Markov

En général, le type de problèmes que l'apprentissage par renforcement (RL) traite est généralement formulé comme des processus de décision de Markov (MDP). L'approche standard pour résoudre les problèmes MDP est d'utiliser la programmation dynamique, mais l'apprentissage par renforcement offre certains avantages clés par rapport à la programmation dynamique.

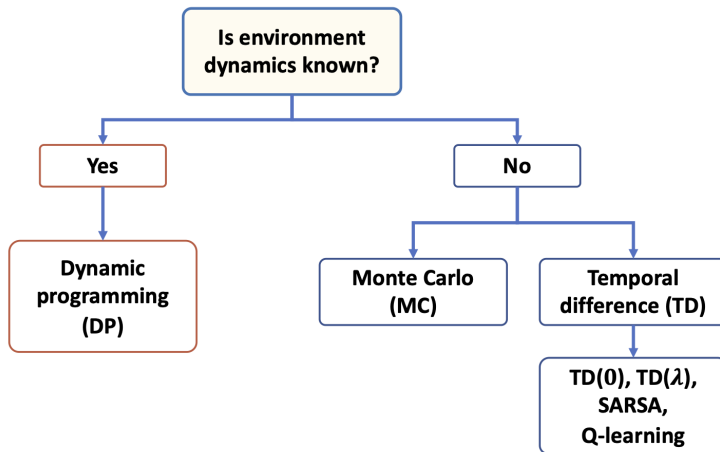
La Formulation Mathématique des Processus de Décision de Markov

La distribution de probabilité pour $S_{t+1} = s'$ et $R_{t+1} = r$ peut être écrite comme une probabilité conditionnelle sur l'état précédent S_t et l'action prise A_t comme suit :

$$p(s', r | s, a) \equiv P(S_{t+1} = s', R_{t+1} = r | S_t = s, A_t = a)$$

Deux approches principales pour traiter ce problème sont les méthodes Monte Carlo (MC) sans modèle et les méthodes de différence temporelle (TD). Le tableau suivant présente les deux principales catégories et les branches de chaque méthode :

La Formulation Mathématique des Processus de Décision de Markov



La Formulation Mathématique des Processus de Décision de Markov

La dynamique de l'environnement peut être considérée comme déterministe si des actions particulières pour des états donnés sont toujours ou jamais prises, c'est-à-dire $p(s', r | s, a) \in \{0, 1\}$. Sinon, dans le cas plus général, l'environnement aurait un comportement stochastique.

- ① Introduction au RL
- ② Interface Agent-Environnement
- ③ Fondations Théoriques
- ④ Terminologie du RL**
- ⑤ Deep Q-Learning

Lorem Ipsum Dolor Sit Amet

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus dignissim mi et faucibus interdum. Etiam nisl sapien, auctor in hendrerit non, laoreet auctor dolor. Pellentesque vel ex id dui blandit porttitor ut vitae risus.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus lacinia odio vitae vestibulum vestibulum. Cras venenatis euismod malesuada.

- ① Introduction au RL
- ② Interface Agent-Environnement
- ③ Fondations Théoriques
- ④ Terminologie du RL
- ⑤ Deep Q-Learning**

Lorem Ipsum Dolor Sit Amet

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus dignissim mi et faucibus interdum. Etiam nisl sapien, auctor in hendrerit non, laoreet auctor dolor. Pellentesque vel ex id dui blandit porttitor ut vitae risus.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus lacinia odio vitae vestibulum vestibulum. Cras venenatis euismod malesuada.

Merci pour Votre Attention !

Abdelhakim Zetati – Lahcen Ezzara