



Université Mohammed V
Faculté des Sciences
Rabat

L'apprentissage par renforcement et le Few-Shot Learning

Mehdi Bouskri

mehdi_bouskri@um5.ac.ma

Introduction

L'apprentissage par renforcement et le few-shot learning sont deux types particuliers d'apprentissage profond, où l'ensemble du système d'apprentissage est créé pour imiter l'intelligence humaine.

L'apprentissage par renforcement est considéré comme une catégorie du deep learning, mais le few-shot learning est une partie particulière de l'apprentissage supervisé.

L'apprentissage par renforcement

l'apprentissage par renforcement vise à maximiser un signal de récompense, à travers l'interaction avec un environnement.

Le renforcement fait référence à la manière dont certains comportements sont encouragés et d'autres découragés, selon le résultat qu'on souhaite avoir.

Le processus de ce type d'apprentissage repose généralement sur cinq éléments : l'agent, l'environnement, l'état, l'action, et la récompense.

Les éléments d'apprentissage par renforcement

L'agent: un réseau de neurones dont l'objectif est de maximiser la récompense totale qu'il reçoit sur le long terme.

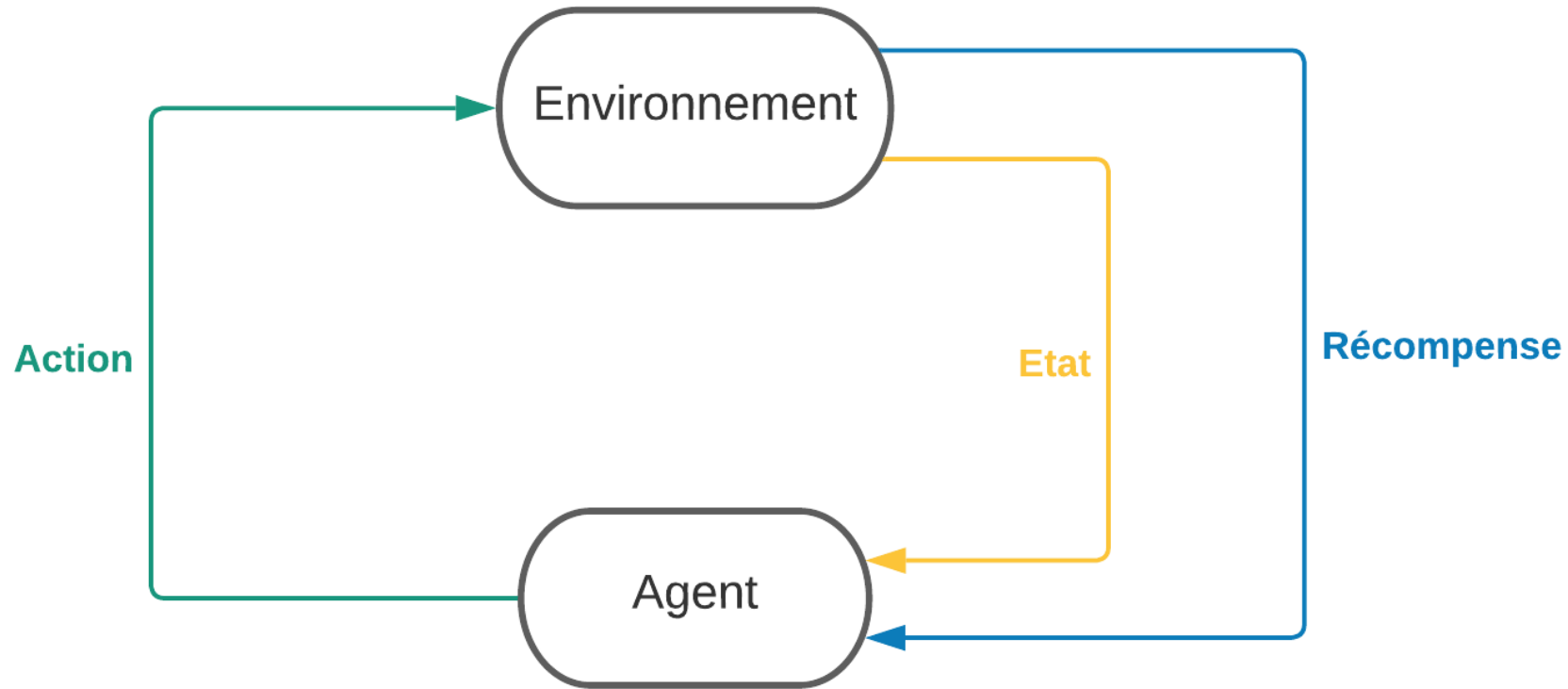
L'environnement: l'espace avec lequel l'agent interagit.

L'état: l'ensemble des informations qui représentent l'état de l'environnement.

L'action: la manière par laquelle l'agent peut modifier l'environnement.

La récompense: la valeur reçue par l'agent après l'exécution d'une action sur l'environnement.

Le processus d'apprentissage par renforcement

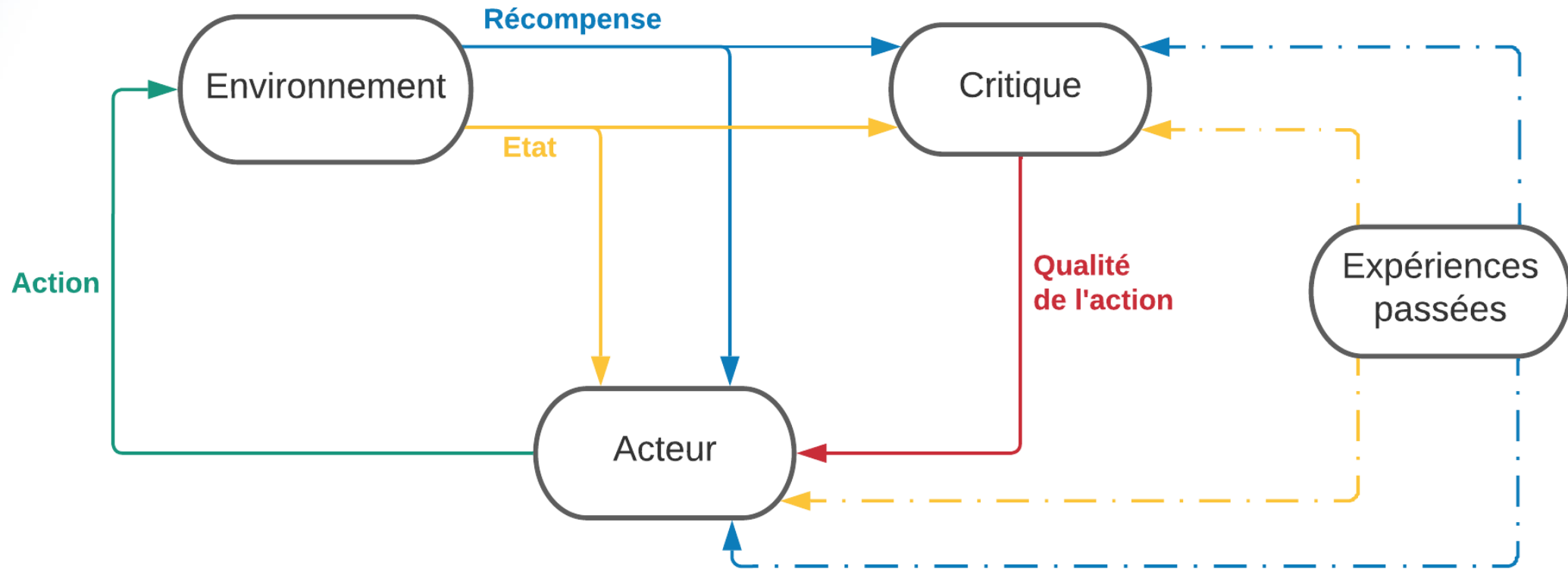


Le Q-learning

Le Q-learning est la manière la plus simple de résoudre un problème d'apprentissage par renforcement.

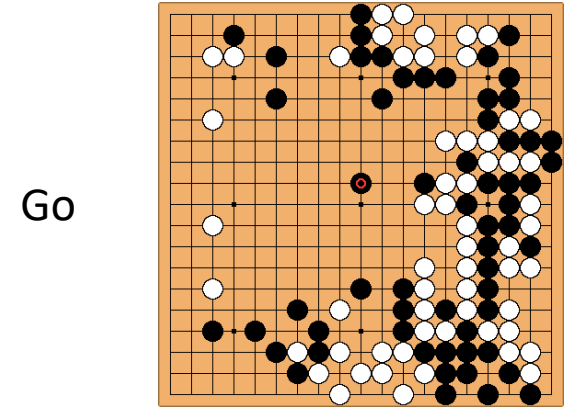
Il consiste à prédire la prochaine récompense pour chaque action, et à choisir la prochaine action avec la plus grande récompense ou Q-value.

Actor-Critic



Quelques applications notables d'apprentissage par renforcement

AlphaGo: le premier programme informatique à défier un joueur humain professionnel de Go.



Go

AlphaZero: un algorithme générique d'apprentissage par renforcement .

OpenAI Five: la première IA à dépasser les champions du monde dans Dota 2.



Dota 2

Quelques applications notables d'apprentissage par renforcement

Contrôle des systèmes de refroidissement commerciaux:
l'utilisation de l'apprentissage par renforcement a conduit à des économies d'énergie d'environ 9 % et 13 % pour deux systèmes.

L'état est un ensemble numérique de mesures de capteurs telles que les températures, le flux d'eau et l'état de l'équipement.

L'action est un vecteur à 12 dimensions représentant la séquence d'opérations du système.

La récompense est le négatif de la consommation totale d'énergie pendant une durée déterminée.

Few-shot learning

S'agit de la classification des données lorsque le nombre d'éléments d'entraînement disponibles est limité.

En général, il existe trois cas particuliers de few-shot learning:

- Few-shot: lorsque l'ensemble d'apprentissage contient plus qu'un élément par classe.
- One-shot: lorsque l'ensemble d'apprentissage contient un seul élément par classe
- Zero-shot: lorsque la classe prédite n'existe pas dans l'ensemble d'apprentissage.

La classification N-way-K-Shot

Le few-shot learning est caractérisé par les éléments suivant:

- L'ensemble d'apprentissage nommé Support, qui consiste de N classes et K éléments pour chaque classe.
- L'ensemble des nouveaux éléments qu'on veut classifier, nommé Query.

Exemple de classification 4-Way-1-Shot

Support

Ili Pika



Koala



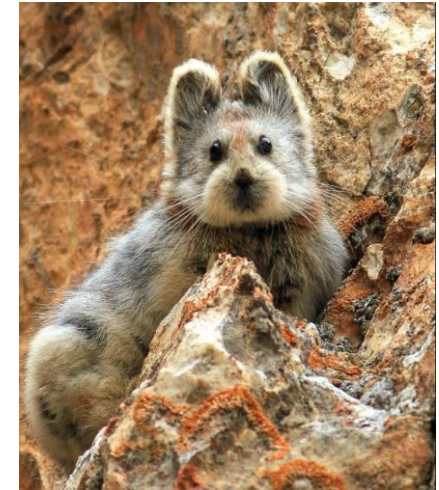
Lemur



Écureuil



Query



Meta-Learning

Le Meta-learning généralement est le processus d'apprendre à apprendre.

Le modèle apprend comment classer étant donné un ensemble de données d'entraînement. Ensuite, le modèle utilise ce qu'il a appris de chaque ensemble de problèmes de classification pour classer d'autres nouveaux ensembles.

Matching Networks

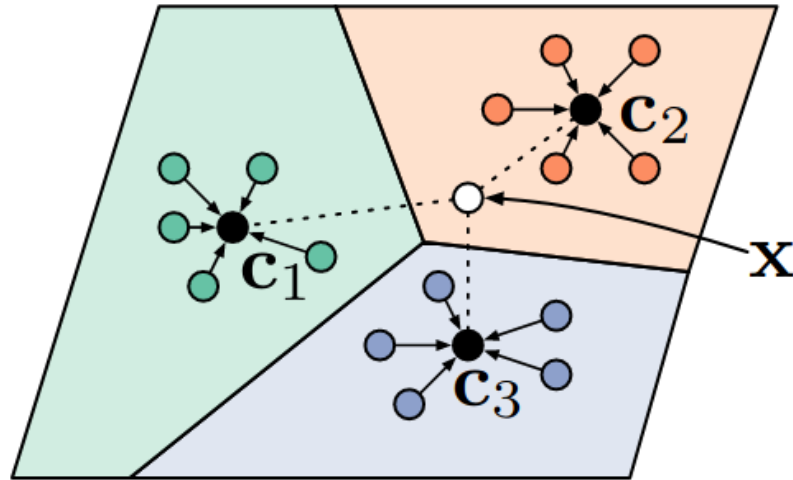
Matching networks apprennent à calculer les embeddings des images.

Chaque image de l'ensemble de Support et de l'ensemble de Query est envoyée à un réseau qui produit des embeddings pour ces images.

Ensuite chaque image de Query est classée en utilisant la distance entre ses embeddings et les embeddings de l'ensemble de Support.

Prototypical Networks

Les Prototypical networks forment des prototypes de classe, qui sont des embeddings de classe obtenues par la moyenne des embeddings d'images de cette classe. Les embeddings des images de Query sont ensuite comparées uniquement à ces prototypes de classe.



Model-Agnostic Meta-Learning

Le MAML apprend des paramètres d'initialisation qui permettent au modèle de s'adapter efficacement à une nouvelle tâche avec de nouvelles classes.

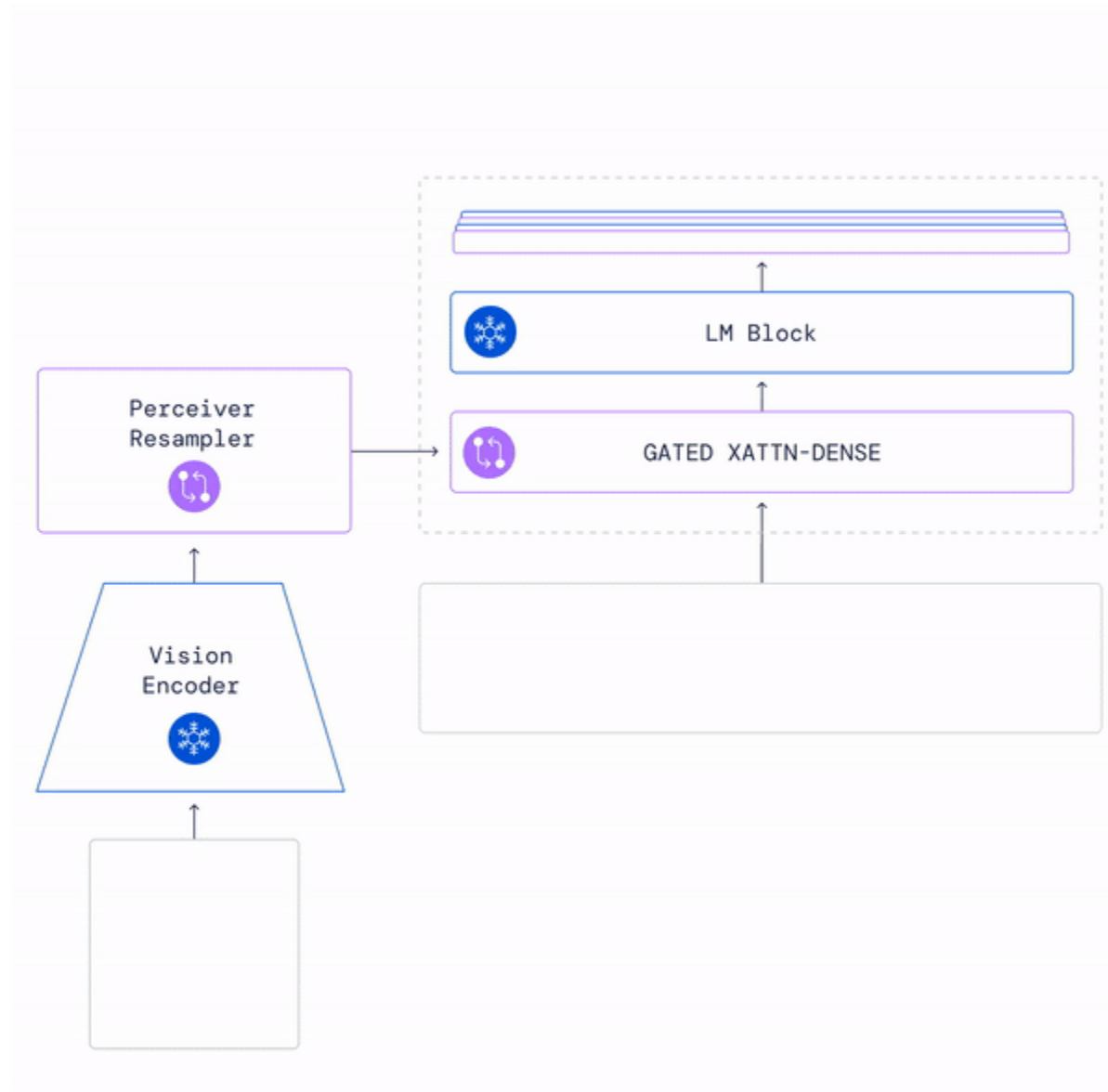
Le MAML crée une copie du modèle et l'initialise avec un ensemble de poids, entraîne le modèle sur l'ensemble de Support, l'applique à l'ensemble de Query, et enfin il rétropropage la perte résultant de la classification du Query à travers tout le processus, et met à jour les poids d'initialisation.

Modèle de langage visuel pour le Few-shot learning

Flamingo: est un modèle polyvalent qui peut être appliqué à des tâches de compréhension d'images et de vidéos avec un nombre limité d'exemples spécifiques à une tâche.

Il fusionne de grands modèles de langage avec des modèles visuels préentraînés en plus de nouveaux composants architecturaux entre les deux.

L'architecture de Flamingo



Exemples d'utilisation de Flamingo



P This is a picture of two teddy bears on the moon.

What are they doing?

P They are having a conversation.

What object are they using?

P It looks like a computer.

Is this surprising?

P Yes, it is surprising.

Why is this picture surprising to you?

P I think it is surprising because teddy bears are not usually found on the moon.



P This is an apple with a sticker on it.

What does the sticker say?

P The sticker says "iPod".

Where is the photo taken?

P It looks like it's taken in a backyard.

Do you think it is printed or handwritten?

P It looks like it's handwritten.

What color is the sticker?

P It's white.