

# Cours DQN – Section 1.2

## 1.2 – Rappel des limites du Q-Learning classique

Le Q-Learning est une méthode d'apprentissage par renforcement puissante lorsqu'on peut représenter tous les états possibles dans une table Q. Toutefois, cette approche est confrontée à plusieurs limites majeures lorsqu'on l'applique à des environnements réels ou complexes.

### Limite 1 : Explosion combinatoire des états

Le Q-Learning repose sur une table Q qui associe chaque état et chaque action à une valeur. Dans des environnements simples (par exemple, un jeu avec une grille 5x5 et 4 actions), cela fonctionne parfaitement.

Mais si les états sont des images, comme dans les jeux Atari :

- Chaque pixel peut prendre 256 valeurs (niveau de gris).
- Une image de 84x84 contient 7 056 pixels.
- Le nombre d'états possibles devient de l'ordre de  $256^{7056}$ .

**Conséquence :** La table Q est inutilisable car elle serait trop grande pour être stockée ou mise à jour.

### Limite 2 : Absence de généralisation

La table Q est une structure rigide. Elle ne sait rien faire d'autre que :

- Mémoriser exactement les valeurs pour chaque couple (état, action).
- Ne rien prédire pour des états jamais rencontrés.

**Exemple :** Si l'agent a appris à agir dans un état où la pomme est à gauche, il ne saura pas quoi faire dans un état presque identique où la pomme est légèrement plus à droite.

**Conséquence :** Le Q-Learning classique ne permet pas d'extrapoler ou de généraliser à partir d'expériences passées.

### Limite 3 : Inefficacité dans des environnements riches

- L'apprentissage par table devient lent : l'agent doit explorer chaque état possible un par un.
- Cela nécessite un très grand nombre d'épisodes pour que la table Q se remplisse suffisamment.

Dans des environnements stochastiques ou partiellement observables, la table Q devient non seulement lente mais aussi instable.

## Limite 4 : Impossible de traiter des entrées continues ou perceptuelles

Le Q-Learning classique ne sait gérer que :

- Des espaces d'états finis et discrets.
- Des représentations simples comme des positions sur une grille.

Il est totalement inadapté aux entrées comme :

- Images (vision par ordinateur).
- Signaux audio ou temporels.
- Mesures physiques complexes.

## Conclusion

Le Q-Learning, bien qu'efficace dans des contextes restreints, montre de graves limites dès que :

- L'espace d'états devient grand ou continu.
- L'agent doit généraliser à partir d'expériences passées.
- Les entrées sont des données perceptuelles (images, sons, etc.).

Ces limites techniques justifient l'introduction de DQN, qui vient combiner les forces du Q-Learning avec les capacités de représentation des réseaux de neurones.