
DRLProject2

Étude de la non-localité de
l'observation et de la considération
d'incertitudes

Emilien VIMONT

Rappel des objectifs du projet

Cas de l'étude la non-localité : Pacman-Series

- Amélioration des performances dans les problèmes à états/actions discrets non locaux (e.g Pac-Man)
- Étude de variantes des algorithmes classiques
- Prise en main du framework ACME

Cas de l'étude de la considération des incertitudes : Mountain Car

- Étudier l'influence de différentes sources de bruit sur les performances d'apprentissage
- Analyse graphique des résultats

Deep Reinforcement Learning

→ RL : Apprentissage par l'expérience

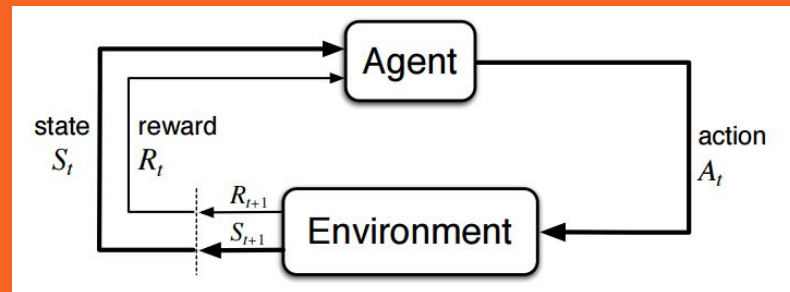
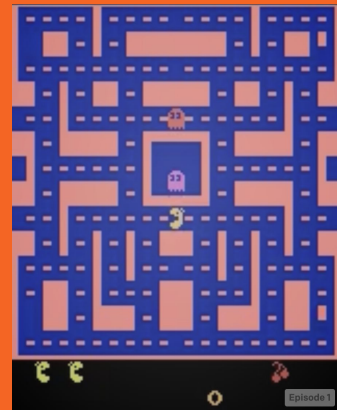
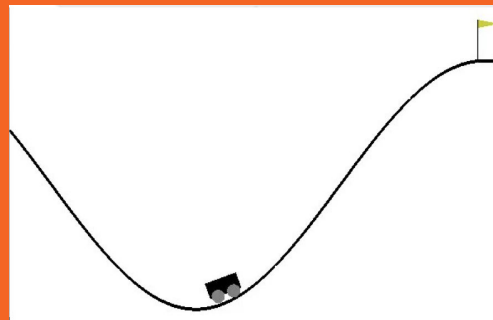
→ Markov Decision Process (MDP)

- ◆ State : "Where I am"
- ◆ Transition : "Where I end up"
- ◆ Action : "What movement I do"
- ◆ Reward : "Did I reach the goal"

→ Outils du DRL

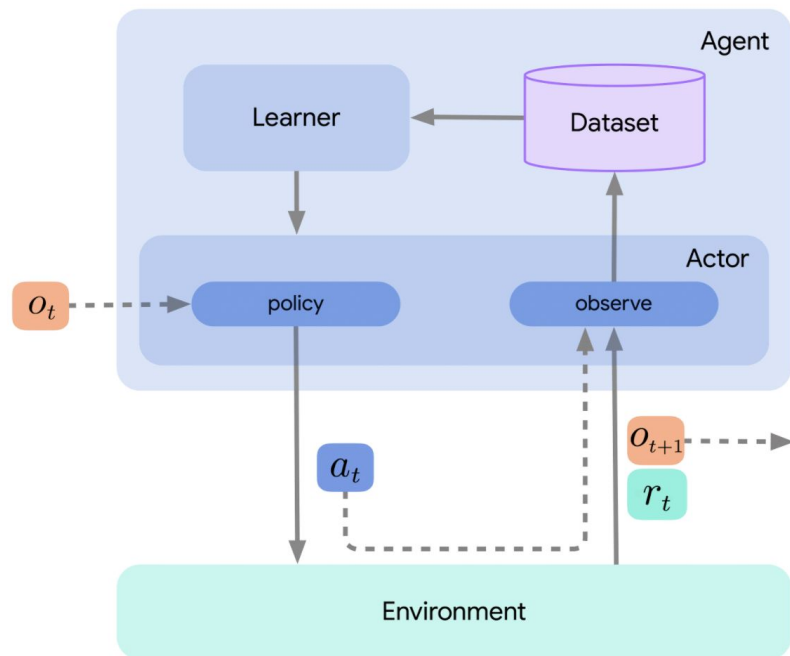
- ◆ Interaction Agent-Environment
- ◆ Policy π

Exemples :



ACME

- Simplifie l'implémentation d'agent (DQN, D4PG, etc.) pour des problèmes de DRL
- Suivi d'un tutoriel DeepMind avec l'environnement MountainCar



Deep Q-Learning/Networks pour un environnement discret

- Intuitivement : Q : évalue la qualité de chaque action
 - ◆ Q donnée pour une paire (s,a) état action
 - ◆ Représente un potentiel gain (Reward) pour des actions futures
 - ◆ Apprendre cette fonction ne nécessite pas de connaître un environnement initial

Pacman Series

Pacman-Series

→ Environnement Gym : MsPacman-v0

→ Aspects du projet :

- ◆ Implications de la non-localité sur le réseau :

- ◆ Modélisation d'un problème non local ou abandonner Markov ?

→ Justification choix de l'Agent : DQN pour des États Discrets

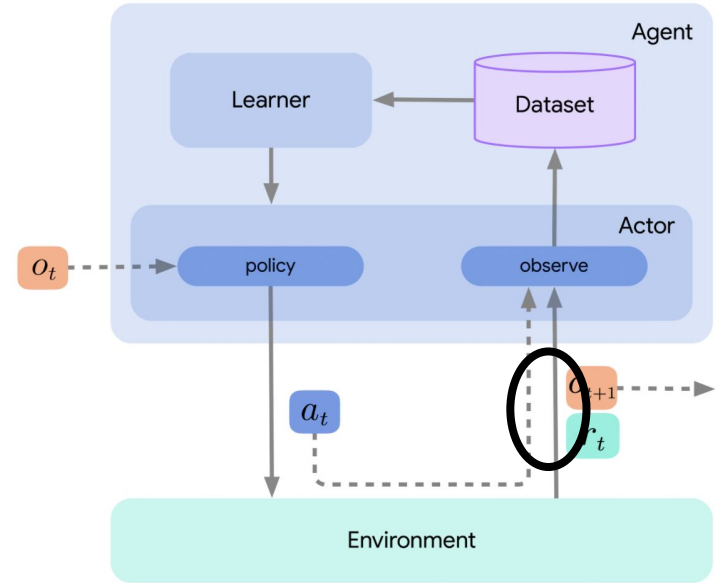
→ Adaptation du code DQN pour Atari : wrappers, environnement, définition de l'agent, etc.



Problèmes rencontrés entre l'agent et l'environnement

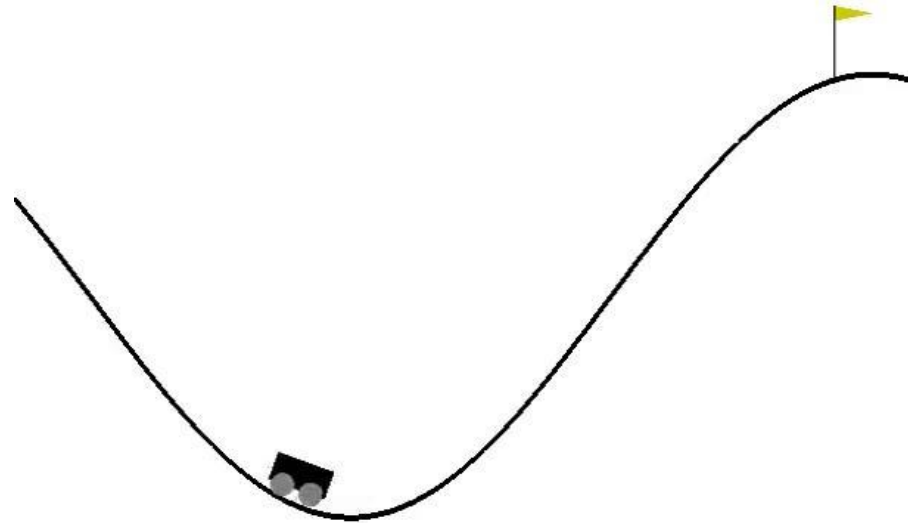
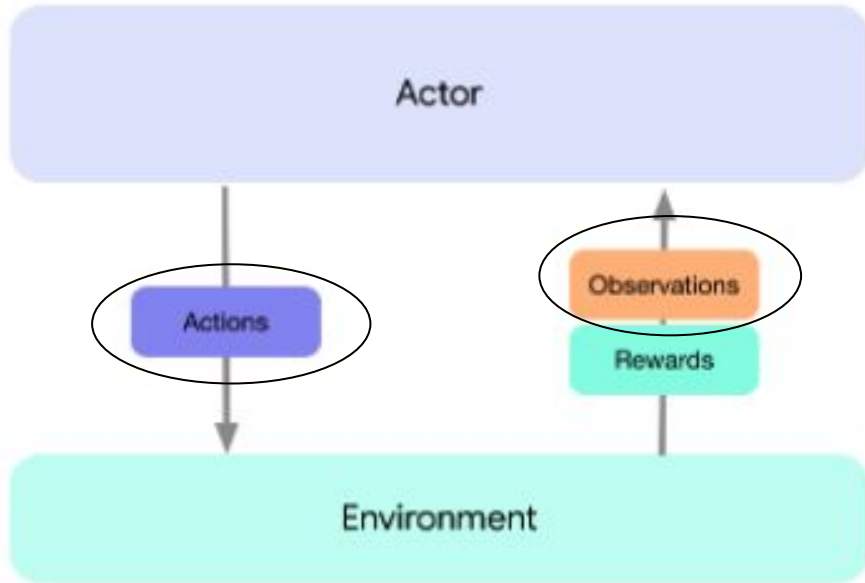
Pacman

- Fastidieux d'utiliser l'agent DQN avec le Pacman
- **Nouvel objectif** : faire correspondre la nature des objets ACME avec celle des objets de l'environnement Atari
 - ◆ Différences trop importantes pour faire tourner Pacman avec un DQN



Mountain Car

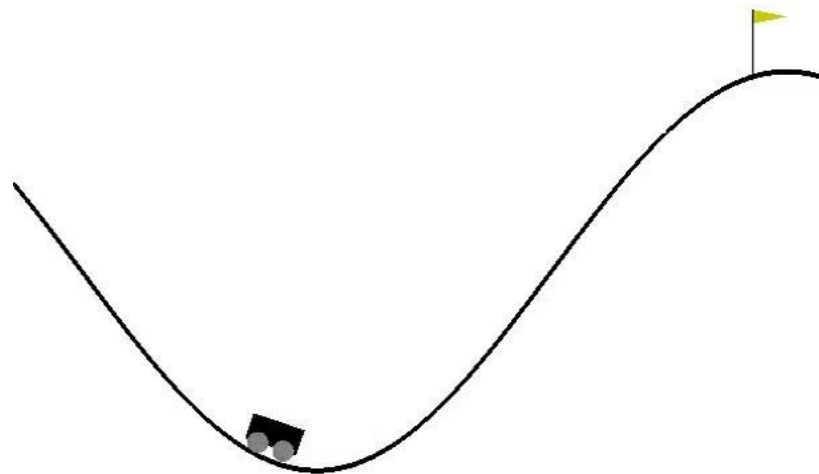
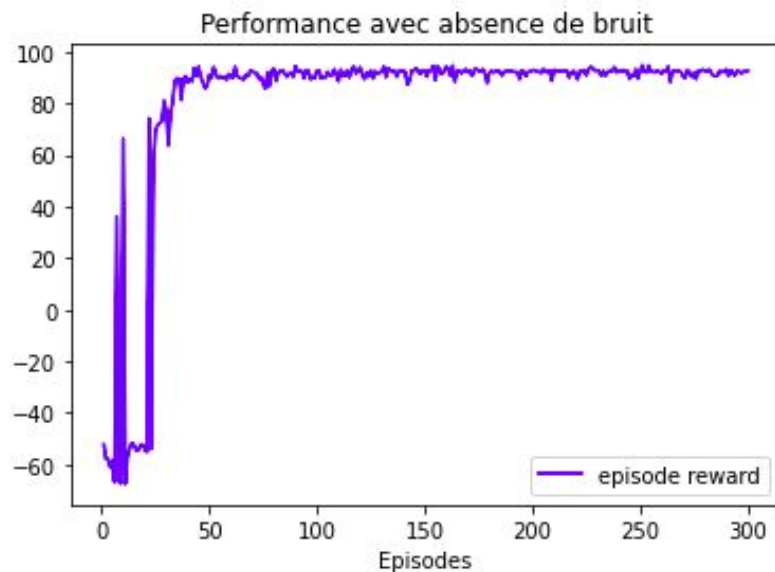
Problème continu et local : Open AI Gym - ACME



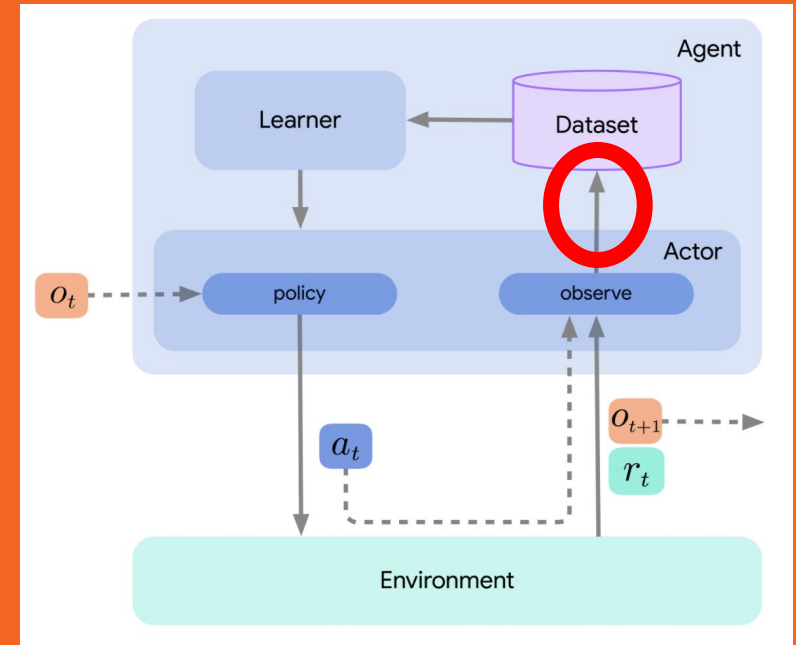
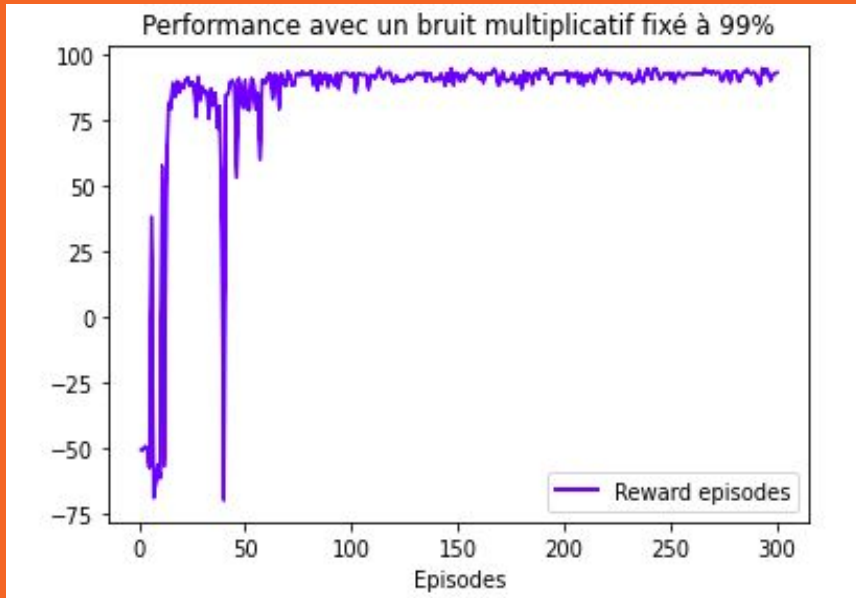
Processus d'apprentissage mis en place :

1. *Génération de l'environnement par Gym : ici le Mountain Car*
2. *Construire un policy network adapté (MLP)*
3. *Créer l'acteur (FeedForwardActor)*
4. *Critiquer l'acteur : D4PG or Distributed Distributionnal DDPG*
5. *Training Loop*
6. *Evaluation de l'agent*

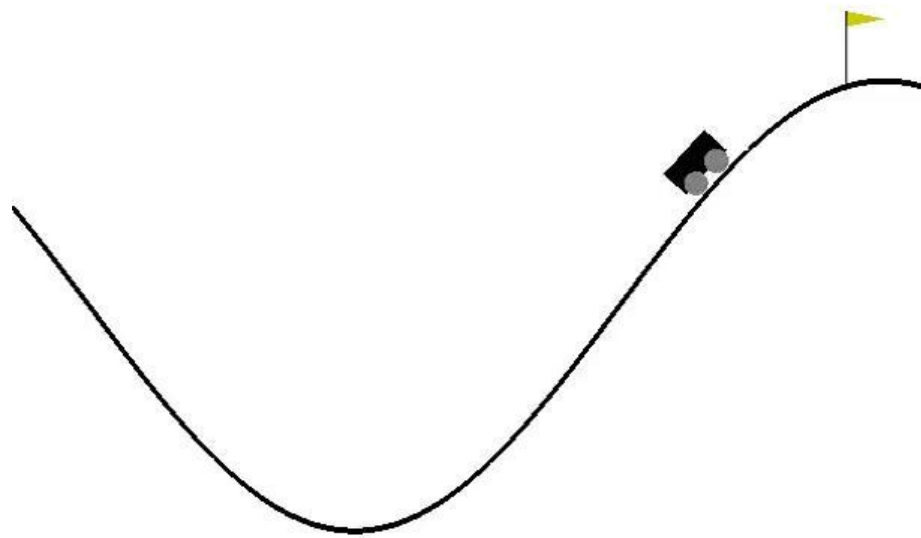
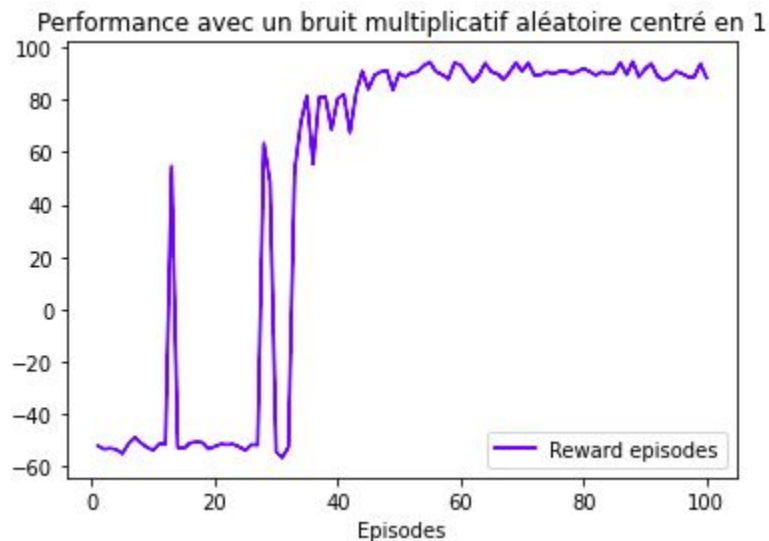
Evaluation de D4PG pour le Mountain Car classique



Introduction d'un bruit multiplicatif constant



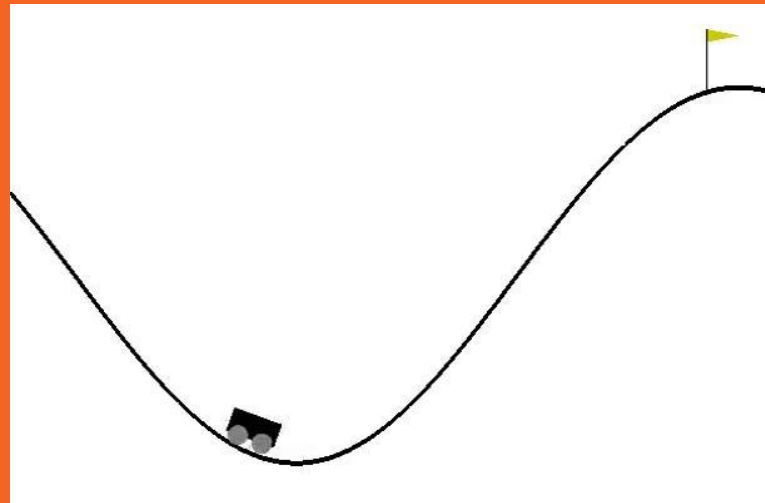
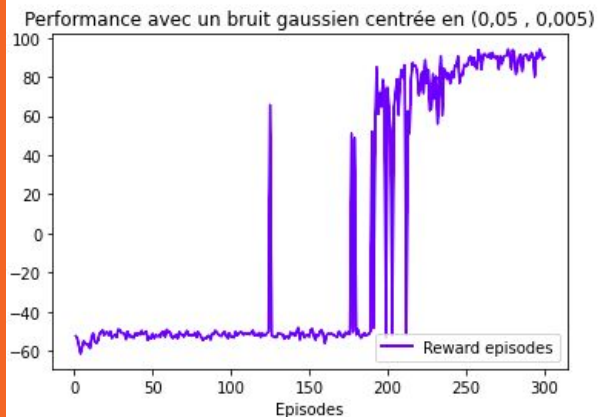
Multiplication de l'observation par un bruit Gaussien centré en 1 et avec un écart-type sigma tq $\sigma^2=0,05$



Introduction d'un bruit gaussien additif sur l'observation centrée en 0 d'écart type σ tq $(\sigma(x)^2=0,05, \sigma(y)^2=0,005)$

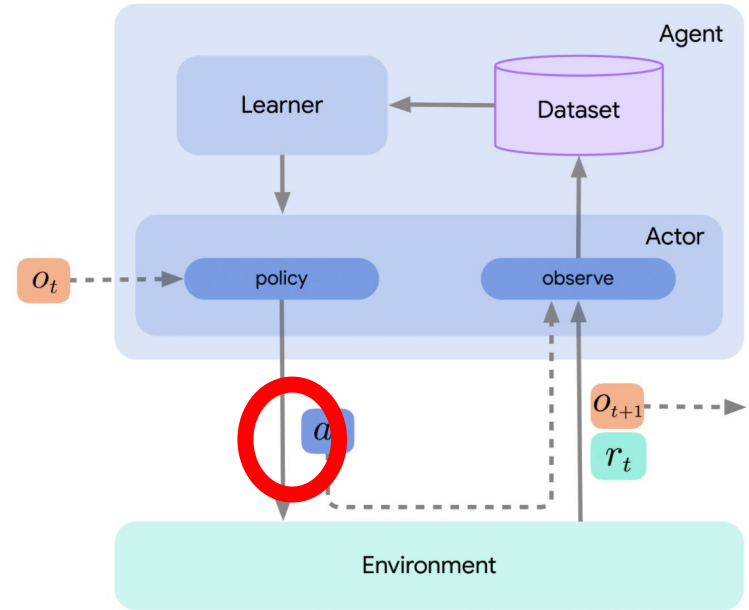
observation $x \in [-1,2,0,6]$

observation $y \in [-0,07,0,07]$



Introduction d'un bruit gaussien additif sur l'action centré en 0 et d'écart-type sigma tq ($\sigma^2=0,05$)

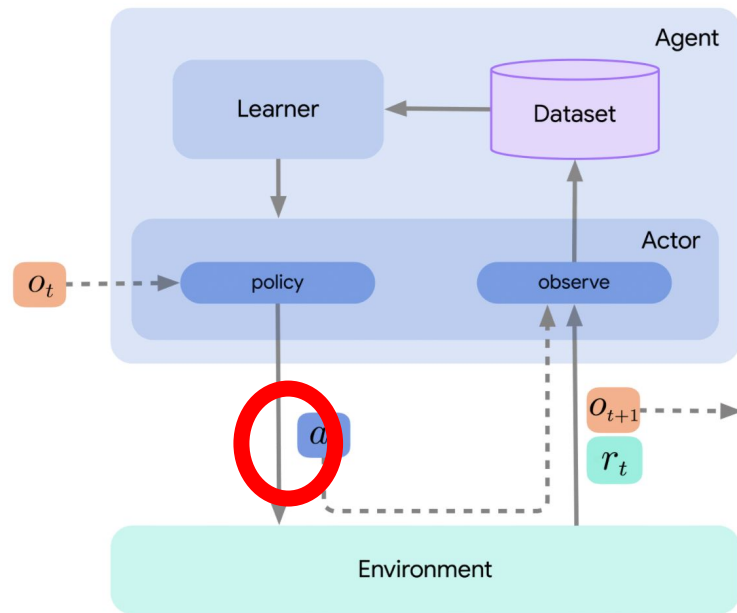
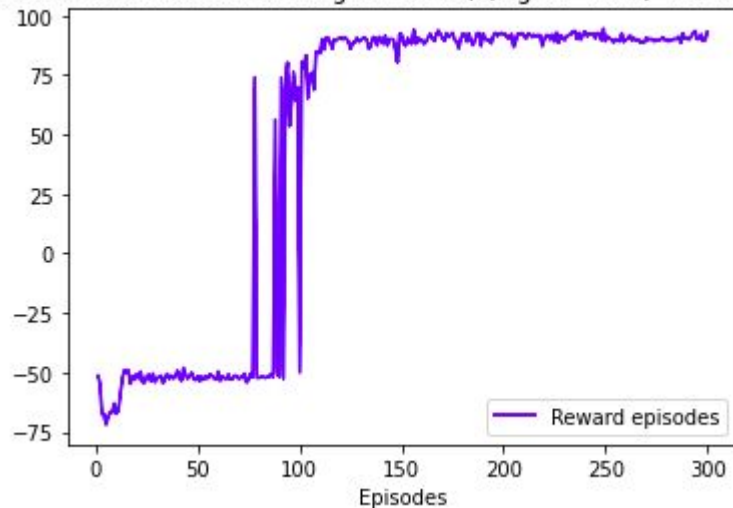
$action \in [-1,1]$



Introduction d'un bruit gaussien additif sur l'action centré en 0 et d'écart-type sigma tq ($\sigma^2=0,05$)

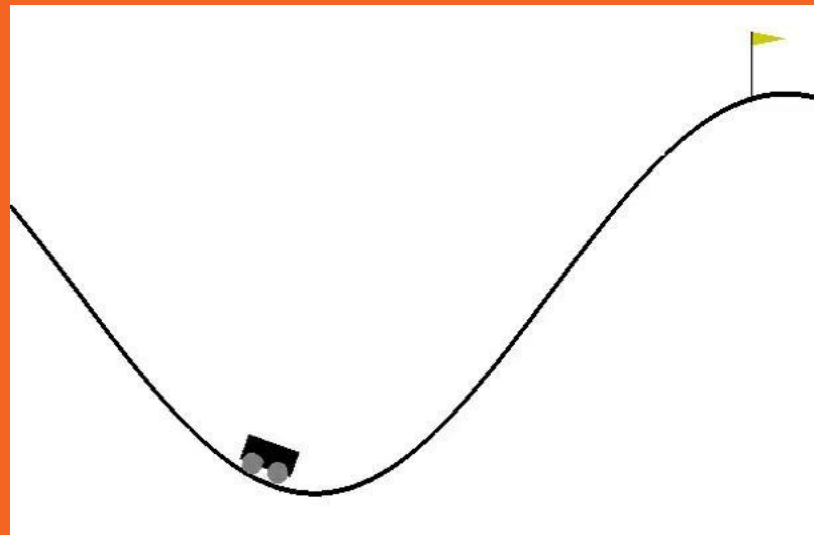
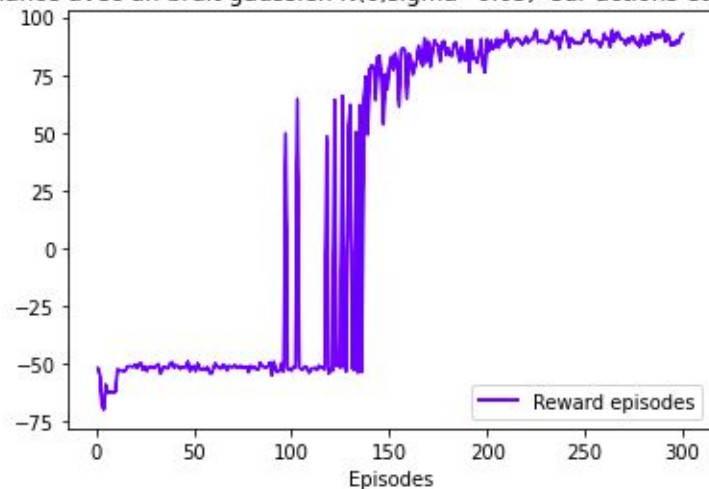
$action \in [-1,1]$

Performance avec un bruit gaussien $N(0, \sigma=0.05)$ sur actions



Introduction d'un bruit gaussien additif sur l'observation centrée en 0 d'écart type sigma tq $(\sigma(\text{observation}))^2=0,05, \sigma^2(\text{action})=0,005$ et introduction d'un bruit gaussien additif centrée en 0 d'écart-type tq $(\sigma^2=0,05)$

Performance avec un bruit gaussien $N(0, \text{sigma}=0.05)$ sur actions et observations



Analyse des résultats expérimentaux et conclusion

- Ajout de bruit blanc sur l'observation
- + Multiplication par un bruit blanc sur l'observation
- + Multiplication par un bruit constant de l'observation
- Ajout de bruit blanc sur l'action
- + Multiplication par un bruit blanc sur l'observation
- + Ajout de bruit blanc sur l'action et l'observation

Ouverture :

Ajout d'un bruit plus proche de la réalité : effectuer une approche mécanique du Mountain Car
Ajout d'un autre type de bruit que le bruit blanc

