

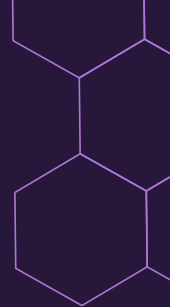
Application sur Super Mario Bros

Projet IFT - 7201

BOUCHERY Loïc
BUIS Anh My



Structure de l'environnement Super Mario Bros (NES)



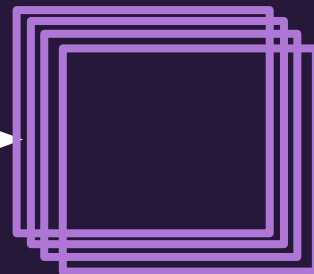
3 x 240 x 256
RVB + fenêtre



1 x 240 x 256
Niveaux de gris



84 x 84
Redimensionnement



4x84x84
Saut de trame
Empilement de trames

Environnement : Premier niveau de Super Mario Bros
Librairie utilisée : gym-super-mario-bros (de OpenAI gym)



ESPACE D'ETATS

Tous les états possibles du jeu

ESPACE D'ACTIONS

Mouvements possibles de Mario

Position de Mario,
obstacles, powerups,
ennemis

Score courant

Direction des
éléments mouvants

Attendre

Courir/Sauter à droite

*Courir/sauter à
gauche*

Fonction de récompense

Fonction de récompense implémentée dans la librairie

VELOCITE	$V = \text{Position1} - \text{Position0}$
TEMPS	$T = \text{Temps0} - \text{Temps1}$
DEATH PENALTY	Si en vie : $D = 0$, sinon $D = -15$
RECOMPENSE	$R = V + T + D$

Indice 0 : avant le step
Indice 1 : après le step

D'après <https://pypi.org/>

Mouvements possibles

Pas bouger $V = 0$

Aller à droite $V > 0$

Aller à gauche $V < 0$

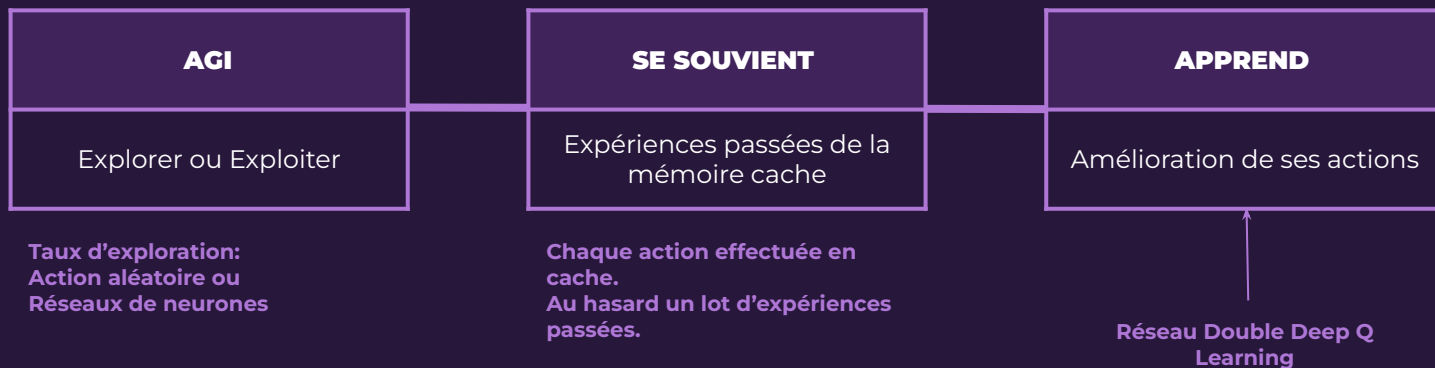
Morts possibles

Tomber dans un trou

Toucher un ennemi

Manque de temps

Construction de l'agent : Mario



DOUBLE Q-LEARNING

$$Q_T(s_t, a_t) = r_{t+1} + \gamma \max_a Q_T(s_{t+1}, a_t)$$

$$Q_l(s_t, a_t) \leftarrow Q_l(s_t, a_t) + \alpha \mathcal{L}[Q_T(s_t, a_t), Q_L(s_t, a_t)]$$

2 réseaux : cible (t) et local (l)

Prédiction des Q - r. local

Apprentissage - r. cible

REFERENCE

- ◆ https://pytorch.org/tutorials/intermediate/mario_rl_tutorial.html
- ◆ <https://arxiv.org/pdf/1509.06461.pdf>
- ◆ <https://pypi.org/project/gym-super-mario-bros/>

