Manœuvre automatique d'un drone :

Mise en œuvre dans un réseau de gestion des déchets ménagers

Omar Ettalbi

Code CNC : ST046T

Juillet 2022

Sommaire

1 Introduction

2 Problématique

3 Le réseau de gestion des déchets ménagers

4 L'intelligence artificielle
- L'apprentissage par renforcement
- L'apprentissage profond
- L'apprentissage par renforcement profond
- Le transfert Sim-To-Real

5 Simulation

6 Conclusion

Introduction

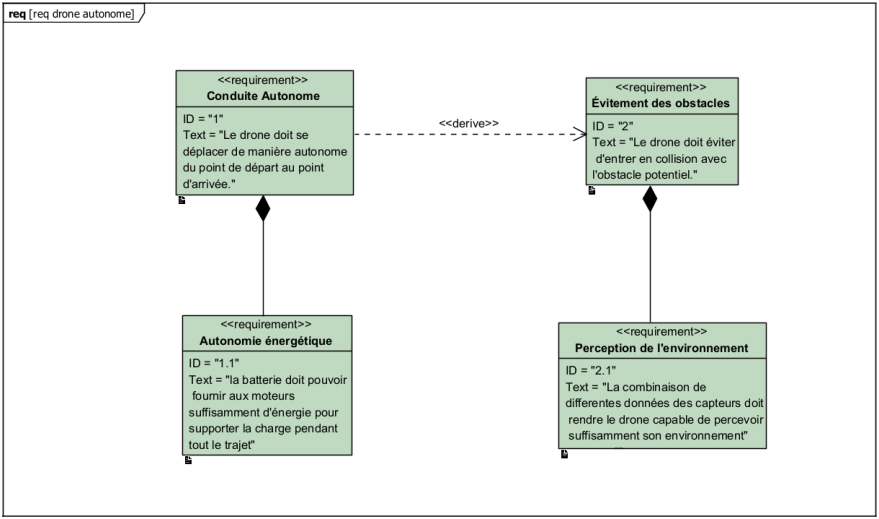
Figure – Diverses utilisations de drones

Figure – Problème des poubelles publiques

Problématique

- Comment rendre un drone capable de faire sortir les déchets ménagers et rentrer chez lui d'une façon autonome ?

Cahier de charge



Le réseau de gestion des déchets ménagers

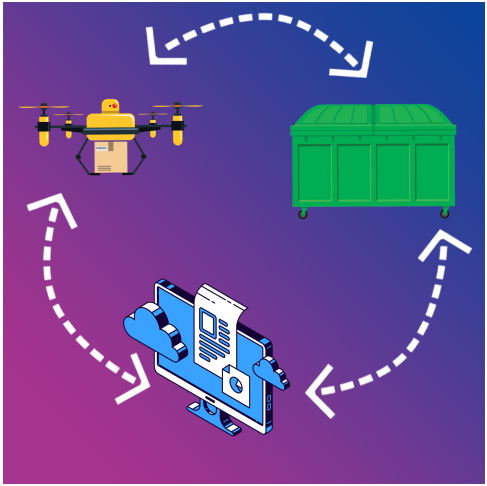


Figure – les éléments du réseau considéré

Particularité de drone

Figure – Drone prêt à être piloté

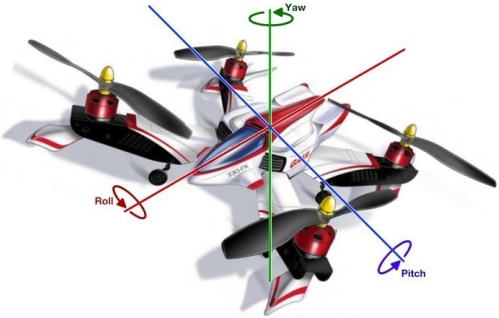


Figure – Les mouvements d'un drone

Type de moteur

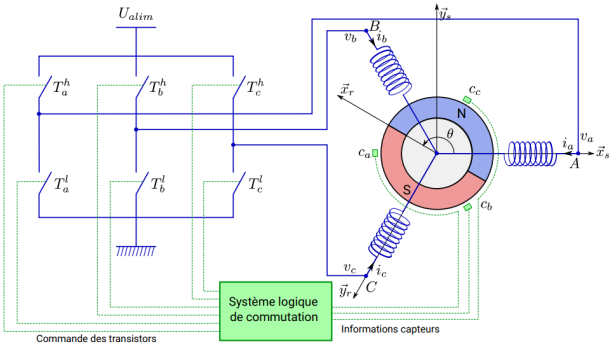


Figure – Variateur de vitesse - Moteur Brushless

Le choix du moteur

- Information sur le moteur ECOII-2807

Motor type	The voltage (V)	Propeller	current (A)	thrust (G)	power (W)	efficiency (G/W)	speed (RPM)
ECOII-2807-1300KV	25.2	Gemfan7040 3-blade propeller	2.3	300	57.96	5.18	8920
	25.2		4.1	500	103.32	4.84	11000
	25.2		6.4	700	161.28	4.34	12800
	25.2		9.3	900	234.36	3.84	14400
	25.2		12.5	1100	315.00	3.49	15700
	25.2		45.2	2490	1139.04	2.19	Max

Motor type	The voltage (V)	Propeller	current (A)	thrust (G)	power (W)	efficiency (G/W)	speed (RPM)
ECOII-2807-1300KV	25.2	HQ 7*3.5*3 3-blade propeller	2.1	300	52.92	5.67	7990
	25.2		4	500	100.80	4.96	10100
	25.2		6.4	700	161.28	4.34	11700
	25.2		9.2	900	231.84	3.88	13200
	25.2		12.4	1100	312.48	3.52	14600
	25.2		47.1	2580	1186.92	2.17	Max

Charge de 5Kg ! Poussée 200% -> 2.5kg chaque moteur

Le choix de la batterie

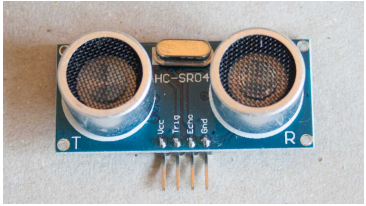
- Batterie Lithium Polymère

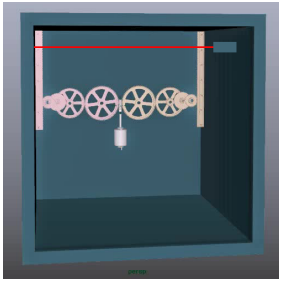
Capacité : 11Ah
-> 285Wh
Puissance absorbée : 1186W
Durée : 14,4 minutes

Particularité de la poubelle

Figure – Ouverture et Fermeture automatique de la couverture de la poubelle

Particularité de la poubelle





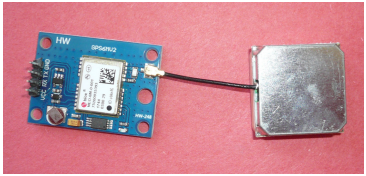


Figure – Les capteurs nécessaires au niveau de la poubelle

Interconnexion des éléments

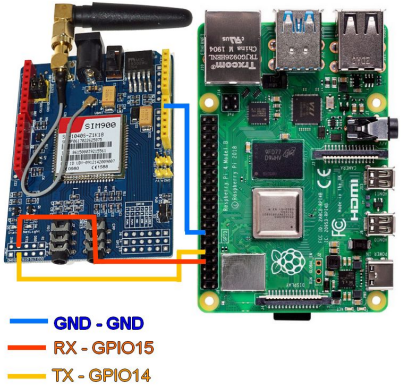


Figure – Les outils nécessaires pour assurer la communication entre les éléments du réseau

Gestion de la base de données sur PhpMyAdmin

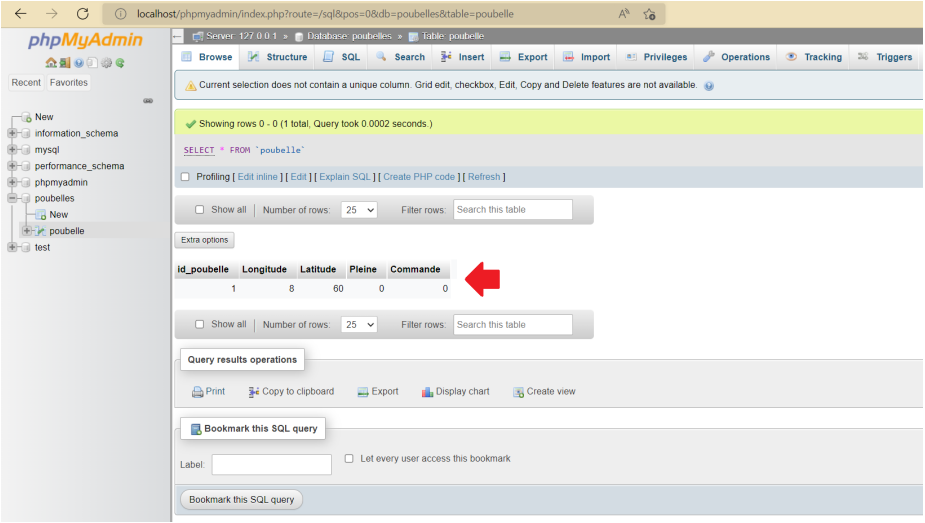
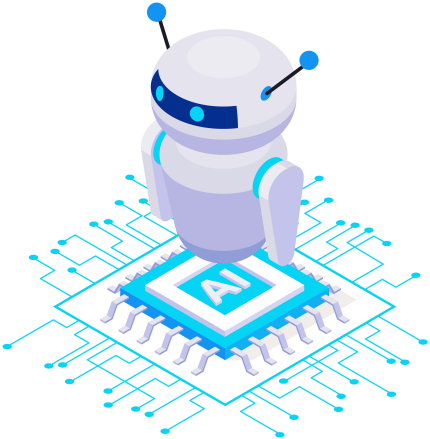
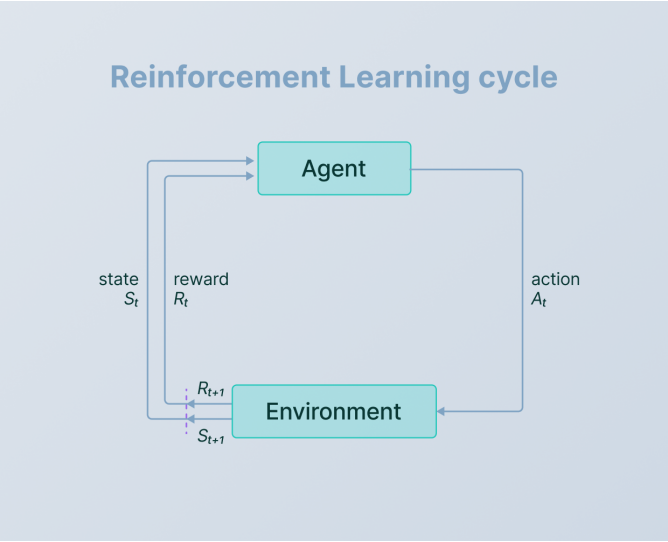


Figure – Serveur web de test local - XAMPP

L'intelligence Artificielle



L'apprentissage par renforcement



Processus de Markov

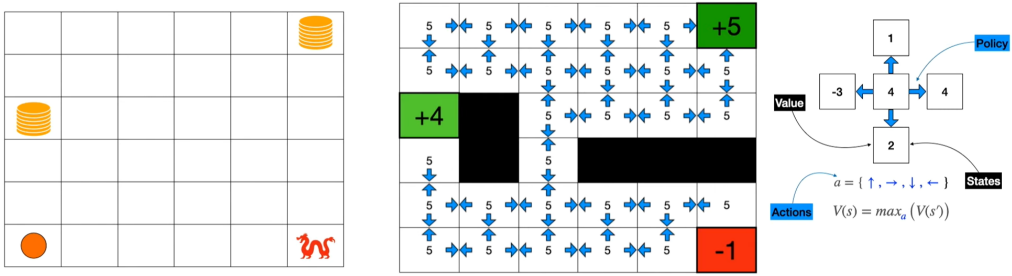


Figure – Exemple classique de Markov Decision Processes

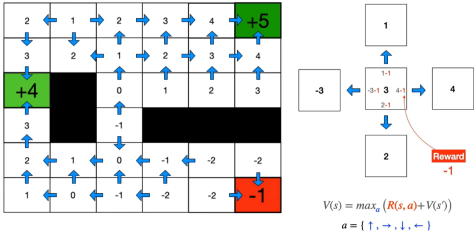


Figure – Récompense pour chaque pas

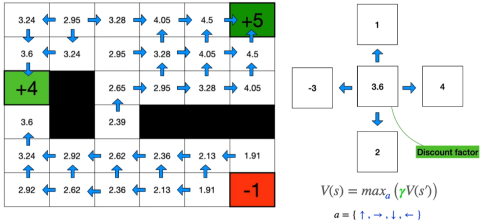


Figure – Constante de convergence

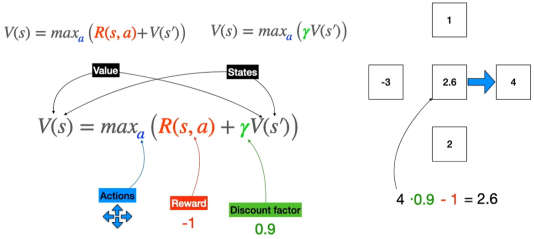


Figure – L'équation de Bellman

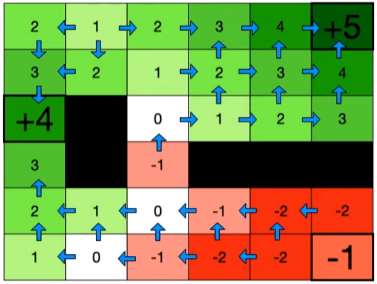
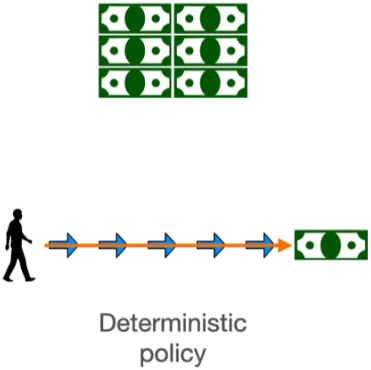
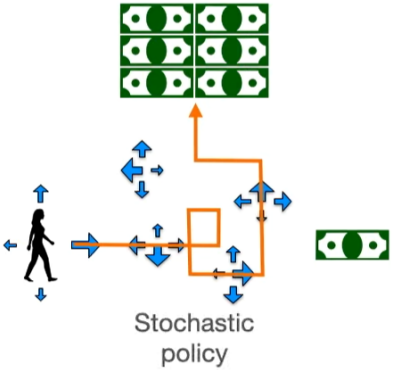


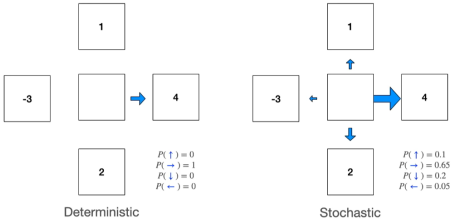
Figure – Sa solution

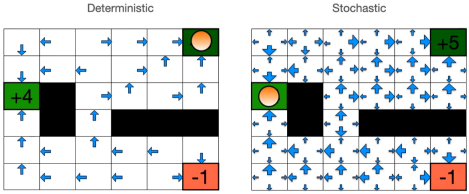
Procédure Deterministe vs Procédure Stochastique





Procédure Deterministe vs Procédure Stochastique





Les récompenses

$$G_t = R_{t+1} + R_{t+2} + R_{t+3} + \cdots + R_T,$$

Expected Return

$$\begin{aligned} G_t &= R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \gamma^3 R_{t+4} + \cdots \\ &= R_{t+1} + \gamma \left(R_{t+2} + \gamma R_{t+3} + \gamma^2 R_{t+4} + \cdots \right) \\ &= R_{t+1} + \gamma G_{t+1} \end{aligned}$$

Related successive returns

$$\begin{aligned} G_t &= R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots \\ &= \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}. \end{aligned}$$

Discounted Return

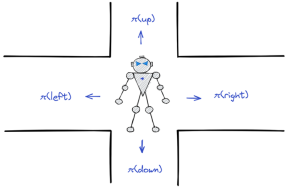
$$G_t = \sum_{k=0}^{\infty} \gamma^k = \frac{1}{1-\gamma}.$$

$$\gamma < 1,$$

Convergence

Policy / Value Functions

$$\pi(a|s) \qquad A_t = a \qquad S_t = s$$



Politique

$$\begin{aligned} v_\pi(s) &= E_\pi\left[G_t \mid S_t = s\right] \\ &= E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid S_t = s\right]. \end{aligned}$$

State-Value function

$$\begin{aligned} q_\pi(s,a) &= E_\pi\left[G_t \mid S_t = s, A_t = a\right] \\ &= E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid S_t = s, A_t = a\right]. \end{aligned}$$

Action-Value function

Qu'apprennent les algorithmes d'apprentissage par renforcement ?

$$\pi \geq \pi' \text{ if and only if } v_\pi(s) \geq v_{\pi'}(s) \text{ for all } s \in \boldsymbol{S}.$$

Optimal Policy

$$v_*(s) = \max_\pi v_\pi(s)$$

Optimal State-Value function

$$q_*(s,a) = \max_\pi q_\pi(s,a)$$

Optimal Action-Value function

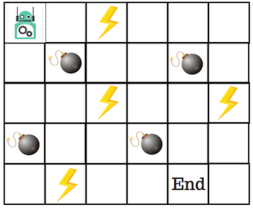
$$q_*(s,a) = E\left[R_{t+1} + \gamma \max_{a'} q_*(s',a')\right]$$

Bellman Optimality Equation

Q-Learning Algorithm

states \ actions	a_0	a_1	a_2	...
S_0	$Q(s_0,a_0)$	$Q(s_0,a_1)$	$Q(s_0,a_2)$	...
S_1	$Q(s_1,a_0)$	$Q(s_1,a_1)$	$Q(s_1,a_2)$	...
S_2	$Q(s_2,a_0)$	$Q(s_2,a_1)$	$Q(s_2,a_2)$	...
⋮	⋮	⋮	⋮	⋮

Initialize Q-table
↓
Choose an action
↓
Perform action
↓
Measure reward
↓
Update Q-table



Actions :	↑	→	↓	←
Start	0	0	0	0
Nothing / Blank	0	0	0	0
Power	0	0	0	0
Mines	0	0	0	0
END	0	0	0	0

Update Current Q-value — Learning Rate — Estimated reward from next action

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\left(R_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1},a_{t+1}) - Q(s_t,a_t)\right)$$

Current Q-value — Rewards — Discount factor

L'apprentissage profond

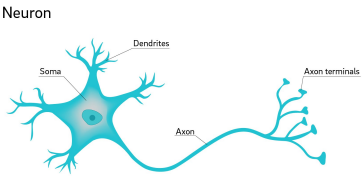


Figure – Neuron Biologique

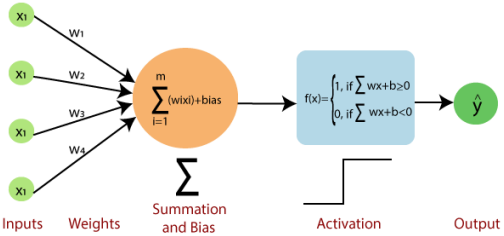
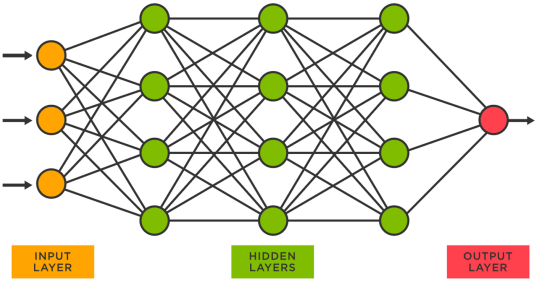
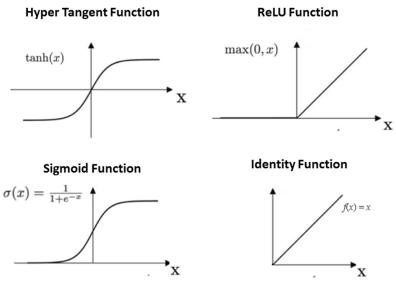


Figure – Perceptron

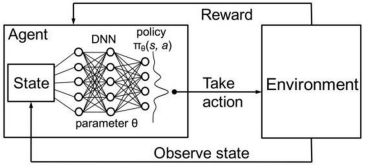
Réseau de neurones artificiels

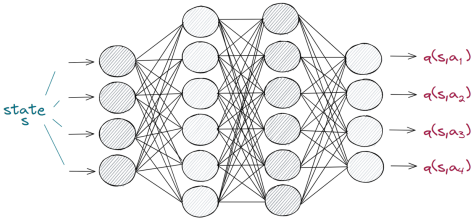




1. Forward Propagation
2. Cost Function
3. Back Propagation
4. Gradient Descent

L'apprentissage par renforcement profond





TD-target is UNKNOWN!

Current Q-value

$$LossFunction = Q_{best}(s_t, a_t) - Q(s_t, a_t)$$

$$q_*(s,a) - q(s,a) = loss$$

$$E\left[R_{t+1} + \gamma \max_{a'} q_*(s', a')\right] - E\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1}\right] = loss$$

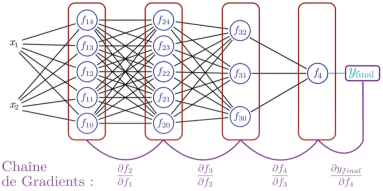
Estimated TD-target

Discount factor

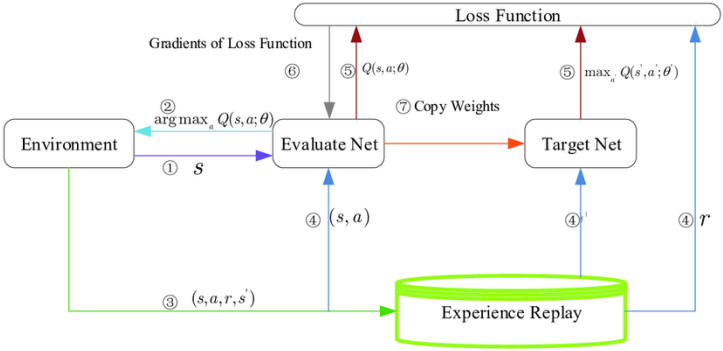
$$R_{t+1} + \gamma \, max Q(s_{t+1}, a)$$

Reward

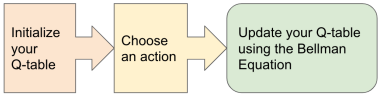
Maximum next-state Q-value



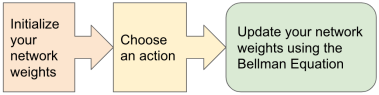
Comment un réseau de neurones de DQN apprend ?



Q-Learning vs Deep Q-Learning



Q-Learning

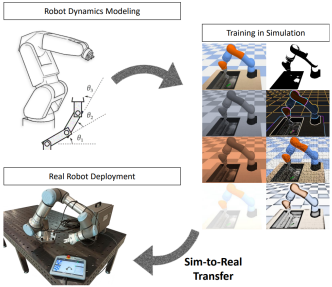


Deep Q-Learning

$$Q(S_t, A_t) = (1 - \alpha)\, Q(S_t, A_t) + \alpha * (R_t + \lambda * max_a\, Q(S_{t+1}, a))$$

Bellman Equation

Le transfert Sim-To-Real



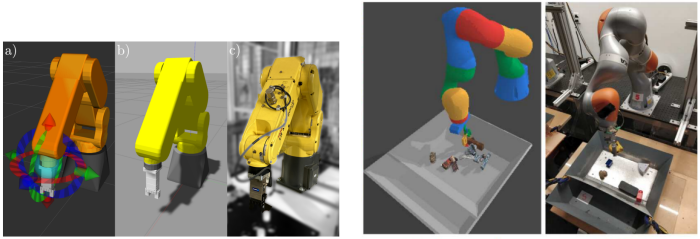
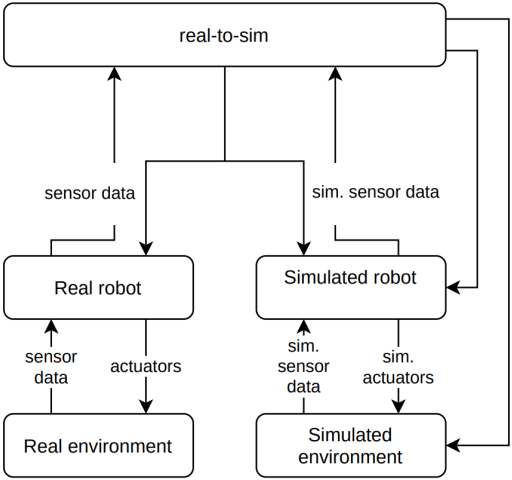


Figure – Exemple de projet déjà réalisé

Le transfert Sim-To-Real



Les observations de notre agent

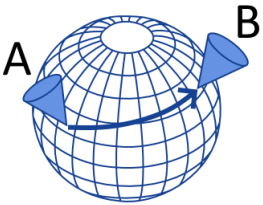


Figure – Distance

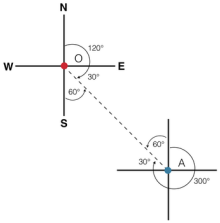


Figure – L'angle de relèvement

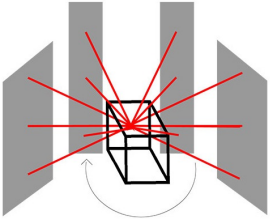


Figure – Lidar

1. Distance : Scalaire,
2. Angle : Scalaire,
3. Vecteur : (Angle, Distance),
4. Altitude : Scalaire.

Les actions de notre agent

Figure – Contrôleur de vol pour le drone

Simulation

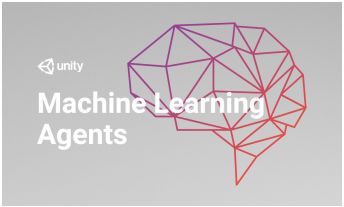
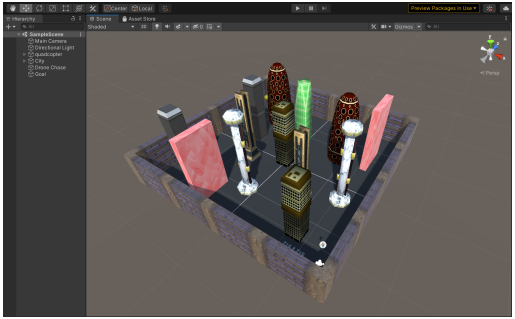
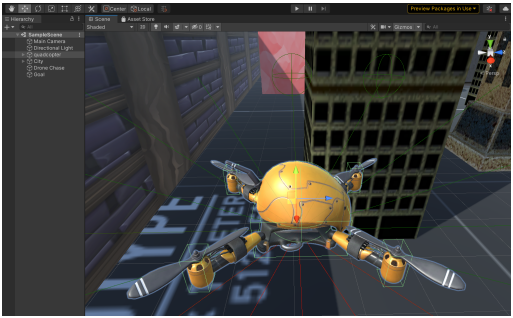


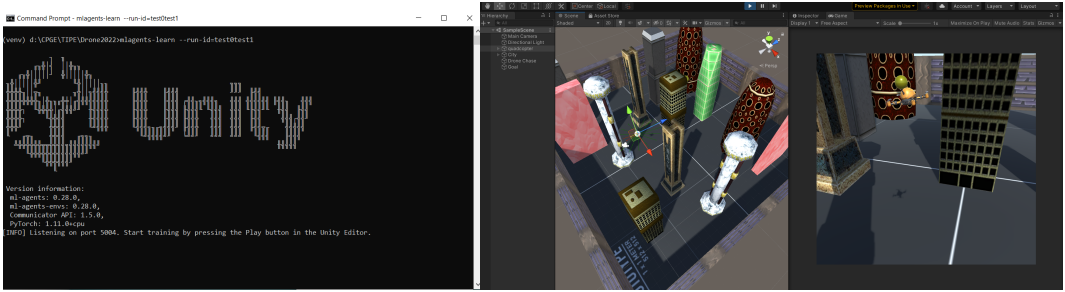
Figure – Les outils utilisés

Univers créer sur Unity 3D

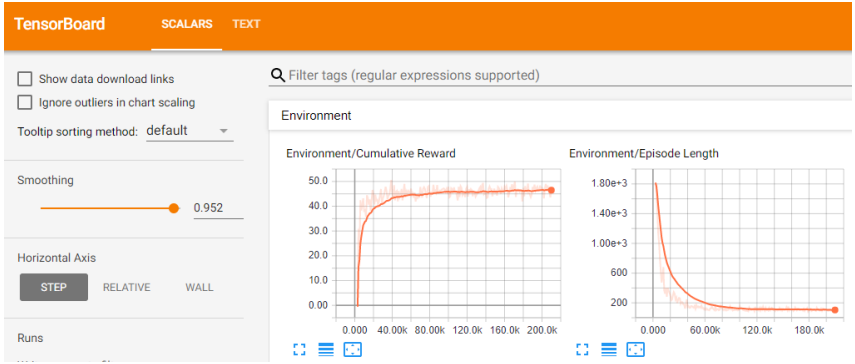




Entraînement de l'agent IA



Le progrès de l'apprentissage



Conclusion

- Alors, quel est l'avenir de l'intelligence artificielle et de son intersection avec la robotique ?

Manœuvre automatique d'un drone :

Mise en œuvre dans un réseau de gestion des déchets ménagers

Omar Ettalbi

Code CNC : ST046T

Juillet 2022