ECE
ÉCOLE D'INGÉNIEURS
ENGINEERING SCHOOL

LYRIDS
Laboratory of Research in Interdisciplinary
Digital Engineering & Sciences
ECE RESEARCH CENTER

STAGE RECHERCHE

RÉUNION DU 18/06/2024

Etudiant: Guy MBOSSO MBOSSO

Encadrant: Jae Yun JUN KIM

Ce qui a été fait

Début de l’implémentation

Lecture des articles

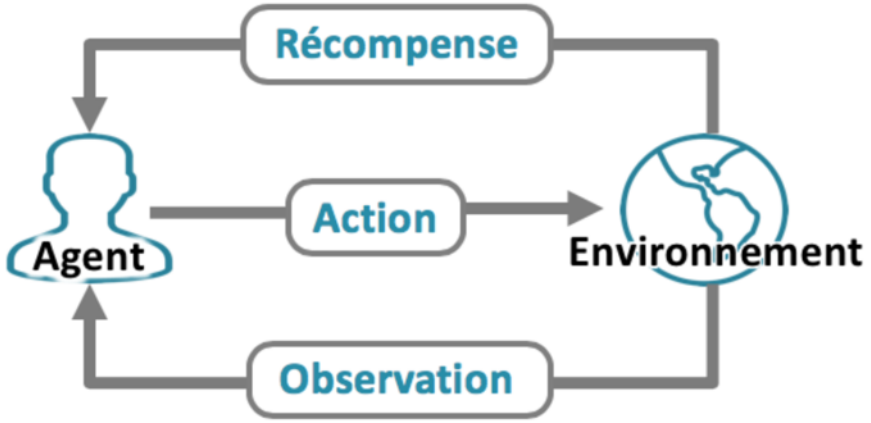
Implémentation

Etat = { coordonnées GPS , vitesse }

Action = { augmenter la Vitesse,

diminuer la Vitesse,

gauche, droite}



$$\max_{\pi} \sum_{t} R_t$$

Résumé des articles

Application du DL pour le RL

- Une variante de l'algorithme du Q-learning pour entrainer le réseau.

$$R_t = \sum_{t'=t}^{T} \gamma^{t'-t} r_{t'}$$

$$Q^*(s,a) = \max_\pi \mathbb{E}\left[R_t | s_t = s, a_t = a, \pi\right]$$

- Calculer une estimation des valeurs Q optimales pour les paires (s, a) grâce au Deep Q-network (DQN).
- Dans l'état s, choisir l'action présentant la plus grande valeur état-action.

Résumé des articles

$$\nabla_{\theta_i} L_i(\theta_i) = \mathbb{E}_{s,a\sim\rho(\cdot);s'\sim\mathcal{E}}\left[\left(r + \gamma \max_{a'} Q(s', a'; \theta_{i-1}) - Q(s, a; \theta_i)\right) \nabla_{\theta_i} Q(s, a; \theta_i)\right]. \quad (3)$$

Algorithm 1 Deep Q-learning with Experience Replay

Initialize replay memory $\mathcal{D}$ to capacity N
Initialize action-value function Q with random weights
for episode $= 1, M$ **do**
  Initialise sequence $s_1 = \{x_1\}$ and preprocessed sequenced $\phi_1 = \phi(s_1)$
  for $t = 1, T$ **do**
    With probability ϵ select a random action a_t
    otherwise select $a_t = \max_a Q^*(\phi(s_t), a; \theta)$
    Execute action a_t in emulator and observe reward r_t and image x_{t+1}
    Set $s_{t+1} = s_t, a_t, x_{t+1}$ and preprocess $\phi_{t+1} = \phi(s_{t+1})$
    Store transition $(\phi_t, a_t, r_t, \phi_{t+1})$ in $\mathcal{D}$
    Sample random minibatch of transitions $(\phi_j, a_j, r_j, \phi_{j+1})$ from $\mathcal{D}$
    Set $y_j = \begin{cases} r_j & \text{for terminal } \phi_{j+1} \\ r_j + \gamma \max_{a'} Q(\phi_{j+1}, a'; \theta) & \text{for non-terminal } \phi_{j+1} \end{cases}$
    Perform a gradient descent step on $(y_j - Q(\phi_j, a_j; \theta))^2$ according to equation 3
  end for
end for