<< **Python ML - Midterm Project** >>

> ### *Q-Learning with OpenAI Gym Environment*

On-line Report Turn-in Due : **2022/05/19 12:00pm**

[期中報告注意事項] :

1. 請將各組期中專案報告壓縮檔以 **zip** 格式壓縮，上傳至學校的教學平台，檔案名稱如下：
　　　　110-2-midterm_Group_OO.ipynb (例如：110-2-Midterm_Group_1.ipynb)

2. 期中報告 **.ipynb** 電子檔內，須註明報告標題以及各組員之科系、年級、學號和姓名。

3. 期中專案的問題、演算法、內容、檔案格式等相關規定，定義於第 2 頁中！

4. 期中專案報告缺交和遲交者，不能補交，並以零分計算！

　　[建議]： 無論是否能完成所有問題需求，請務必於期限內，上傳期中報告電子檔！

< 前 言 >

本專案旨在利用 **OpenAI Gym Environment 的開發套件 [Ref.3]** 建構一個虛擬環境，以提供實作 Reinforcement Learning (RL，強化學習) 的 "**Q-Learning**" 演算法。其中，經由模擬 TAXI 載客環境，建立 **Q table**，解決 RL 訓練問題。
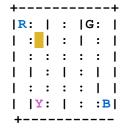

[有關 TAXI 載客 問題 — 簡述]

請用 OpenAI Gym 產生 6x6 的街道環境 (如圖一，**:**代表街道，**|**代表分隔島)。

TAXI 需要載運兩位乘客 (分別在兩處地點，以藍色字母標記)。接到第一位乘客時，TAXI 顏色需轉換成綠色；當接到第二位乘客時，TAXI 顏色需轉換成紅色。之後，將兩位乘客載送至共同目的地 (以紫色字母標記)。

```
+-----------+
|R: | : |G: |
| :█| : | : |
| : : : : : |
| | : | : | |
| : : : | | |
| |Y: | : :B|
+-----------+
```

[ALGORITHM] : Q-Learning Algorithm

Q-learning (off-policy TD control) for estimating $\pi \approx \pi_*$

Algorithm parameters: step size $\alpha \in (0, 1]$, small $\varepsilon > 0$
Initialize $Q(s, a)$, for all $s \in \mathcal{S}^+, a \in \mathcal{A}(s)$, arbitrarily except that $Q(terminal, \cdot) = 0$

Loop for each episode:
    Initialize S
    Loop for each step of episode:
        Choose A from S using policy derived from Q (e.g., ε-greedy)
        Take action A, observe R, S'
        $Q(S, A) \leftarrow Q(S, A) + \alpha \big[R + \gamma \max_a Q(S', a) - Q(S, A) \big]$
        $S \leftarrow S'$
    until S is terminal

[REFERENCE]

1. Richard S. Sutton and Andrew G. Barto, **Reinforcement Learning: An Introduction**, MIT Press, 2018. https://www.andrew.cmu.edu/course/10-703/textbook/BartoSutton.pdf

   [MDP] : Chapter 3 & [Q-Learning] : Chapter 6

2. Joy Zhang, **Tutorial: An Introduction to Reinforcement Learning Using OpenAI Gym's Taxi** , 2021/05/05. https://www.gocoder.one/blog/rl-tutorial-with-openai-gym

3. **OpenAI – Gym Documentation > Environment Creation**

   https://www.gymlibrary.ml/content/environment_creation/

4. **OpenAI – Gym, "Taxi-v3"**, https://gym.openai.com/envs/Taxi-v3/

   [GitHub]: https://github.com/openai/gym/blob/master/gym/envs/toy_text/taxi.py

[**Problem**] : Turn in the "TAXI Passengers Project" report with **the entire OpenAI Gym virtual environment setup files** and **Jupyter-Notebook codes**.

    [Hint] : **MDP – State Space (how many states?), Action Space, Rewards, Q-values (Q-Table) !!**

Good luck !!