

AI HW3-Maze 心得

資管二 A_蘇湘婷_109403009

1. 訓練完的 Q-table

Q-table:

```
(<built-in function all>,          left      right      up      down
0   -754.634013 -454.826323 -888.232540 -454.438959
1   -455.077439 -455.003015 -872.569720 -866.460782
2   -455.144362 -455.212921 -907.869054 -932.942279
3   -455.386479 -868.373351 -902.627396 -455.343803
4     0.000000   0.000000   0.000000   0.000000
..
226  0.000000   0.000000   0.000000   0.000000
227 -944.554491 -455.453689 -942.210699 -945.527300
228 -455.442540 -943.706342 -455.504888 -940.081736
229  0.000000   0.000000   0.000000   0.000000
230 -935.253034 -936.013983 -455.118948 -946.758535

[231 rows x 4 columns])
```

2. 步數最少且寶藏數最多的截圖（步數 + score 分數）

=> 步數：935，Score：5

```
P P P P W W P W P W P P P P P P P P P P
P W W P W P P P P W W P P P W P P P W P P
P W O W P W P W W W P W P W P W W W P W P
P P W P P P P P W P P W P P P P P W P P P
P W P P W P W P P P W P P W P P W W W P W
P P P P W W W P W W P P P P W W P P P P P
P W W W P P W P W P W P P P P W W W P W P
P P P P W P W P W P W W P P P P P W P W P
P W P P W P P P W P W P P W P W P P W W P
P W W P W P W W P P P P P P P P W P P P
P W P P W P P P P P P P P P P W P P W P
['Episode 16: total_steps=935 Score=5']
```

3. Reward 設定截圖並說明

用 if-else 語法來設定向右、左、上、下走遇到不同狀況時(抵達終點、超出界線、撞牆、拿到寶藏) 分別的作法。

```

def get_env_feedback(S,A,path,treasure): #主要要寫的地方
    R_treasure = 1000
    R_bound = -500
    R_GOAL = 10
    R_Normal = 3
    global SCORE, hall
    if A=='right':
        if (S == GOAL - 1) and (SCORE==5): #抵達終點
            S_ = "terminal"
            R = R_GOAL
        elif S % N_STATES_x == N_STATES_x - 1: #要超出界線
            S_ = S
            R = R_bound
        elif S+1 in hall:
            S_ = S
            R = R_bound
        elif S+1 in treasure:
            S_ = S + 1
            R = R_treasure
            SCORE = SCORE +1

        else:
            S_ = S + 1
            R = R_Normal

```

```

    if A=='left':
        #if S == GOAL - 1:
        #    S_ = "terminal"
        #    R = 10
        if S % N_STATES_x == 0:
            S_ = S
            R = R_bound
        elif S-1 in hall:
            S_ = S
            R = R_bound
        elif S-1 in treasure:
            S_ = S - 1
            R = R_treasure
            SCORE = SCORE +1
        else:
            S_ = S - 1
            R = R_Normal

```

```

if A=='up':
    #if S == GOAL - 1:
    #    S_ = "terminal"
    #    R = 10
    if S < 21:
        S_ = S
        R = R_bound
    elif S-21 in hall:
        S_ = S
        R = R_bound
    elif S-21 in treasure:
        S_ = S - N_STATES_x
        R = R_treasure
        SCORE = SCORE +1
    else:
        S_ = S - 21
        R = R_Normal

```

```

if A=='down':
    if (S == GOAL - 21) and (SCORE==5):
        S_ = "terminal"
        R = R_GOAL
    elif S > 209:
        S_ = S
        R = R_bound
    elif S+21 in hall:
        S_ = S
        R = R_bound
    elif S+21 in treasure:
        S_ = S + 21
        R = R_treasure
        SCORE = SCORE +1
    else:
        S_ = S + 21
        R = R_Normal

return S_, R

```

4. 心得

一開始花了些時間搞懂 `qlearning` 的概念，以及 `action` 跟 `state` 如何被記錄在 `qtable` 中，主要是依走過、到終點、撞到牆壁、拿到寶藏、可以繼續走這幾個去設 `State` 跟 `Reward`，這是三次作業以來最有趣的一份了，但也是花最多時間理解的一份作業。

5.Colab 連結

<https://colab.research.google.com/drive/1IujNmgT4vFhbGHO2adVckwbvX48cMBhb#scrollTo=T1gLR8nE-65L>