

QLearning 設起點X: 1 Y: 1 ,終點X: 8 Y: 8

設定 QLearning 迷宮的起點與終點座標

QLearning 設訓練參數: 學習率 $\alpha$  0.1 , 折扣率 $\gamma$  0.9 , 貪婪率 $\epsilon$  0.8

設定 QLearning 訓練參數，三個參數的範圍都是 0~1。

學習率  $\alpha$ ：

折扣率  $\gamma$ ：控制下個狀態值對當前狀態值的影響力。

貪婪率  $\epsilon$ ：挑選下一個動作時，有  $\epsilon$  機會採隨機挑選，有  $(1-\epsilon)$  機會會挑選具最大獎勵的動作。

QLearning 設動作獎勵: 撞牆 -1 , 空格 -0.01 , 終點 2

設定各狀態給予的獎勵值。

QLearning 迭代一次, 模式 訓練模式 , 最大步數 5000

以設定好的參數進行訓練，並傳回在這個階段走完迷宮所需要的步數。若步數大於最大步數則傳回-1。

範例程式：設定參數並進行訓練，將訓練持續到總步數小於 30 步為止。

