

設定QLearning 迷宮的起點與終點座標



設定QLearning訓練參數，三個參數的範圍都是0~1。

學習率α：

折扣率γ：控制下個狀態值對當前狀態值的影響力。

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



設定各狀態給予的獎勵值。



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。

