AI期末專題計畫書—

強化學習在對戰類遊戲上的應用

組員：

4107051008 化學三 林仁凱

4107056001 資工三 賴中原

4107056006 資工三 游庭瑋

4107056008 資工三 林泓凱

目錄

1. 研究背景與動機
2. 研究計畫的難度及分析
3. 研究方法
4. 參考資料
5. **研究背景與動機**

隨著AI的發展，AI已經可以完成很多人類難以做到或是必須花很多時間才能做到的事情了；對於遊戲方面，隨著AlphaGo的出現，也越來越多討論。但我們發現，絕大部分的AI都只琢磨於回合制的遊戲上，即AlphaGo在圍棋上的應用，卻鮮少有對即時遊戲的研究。我們想在貪食蛇遊戲上，加入Deep Q Learning的技術，修改部份遊戲規則並提高遊戲難度，驗證AI操控的貪食蛇即使在高難度的情況下仍可以長時間存活。

1. **研究計畫的難度及分析**

在即時對戰類的遊戲上，要做出一個AI agent難度本身就會比較高一些。但現在已經有很多機器學習的框架了，讓我們不必真的自己寫複雜的微分和導數即可做出一個完整的model。比較困難的點在於，AI的訓練其實相當花時間，在大部分的情形下，Deep learning module的處理時間遠遠超過顯示出圖像的時間，這部分就只能仰賴更高效能的硬體設備；且訓練一次需要的場數皆動輒500場以上，對於參數調整來說相當的有挑戰性。

**三、研究方法**

1. 狀態表示

如第一部份所描述，我們想藉由Deep Q learning來完成這個AI agent。首先，我們定義遊戲的狀態如下：

共11個boolean來描述，即：

(以下一行為一個boolean)

前方、

左方、

右方有沒有危險

目前方向向北、

向東、

向南、

向西

食物方向在北、

東、

南、

西

動作則為3個

前方、

左方、

右方

我們會設計方向成這個方式是因為在貪食蛇遊戲中基本上不可能往反方向走，一定是往前方或左右方，雖然這個設定會需要額外維護蛇目前的方向狀態，但可以大幅減少AI的無謂嘗試。

因此簡單來說，我們設計出的AI即是一個輸入有11個值，輸出有三個值的函數

2.類神經網路搭建

在

四、參考資料

[尝试用强化学习算法来玩下FlappyBird？ - 知乎](https://zhuanlan.zhihu.com/p/108068817)[yenchenlin/DeepLearningFlappyBird: Flappy Bird hack using Deep Reinforcement Learning (Deep Q-learning).](https://github.com/yenchenlin/DeepLearningFlappyBird)

[Deep Q-Learning Tutorial: minDQN. A Practical Guide to Deep Q-Networks | by Mike Wang | Towards Data Science](https://towardsdatascience.com/deep-q-learning-tutorial-mindqn-2a4c855abffc)

[Teach AI To Play Snake! Reinforcement Learning With PyTorch and Pygame](https://github.com/python-engineer/snake-ai-pytorch)