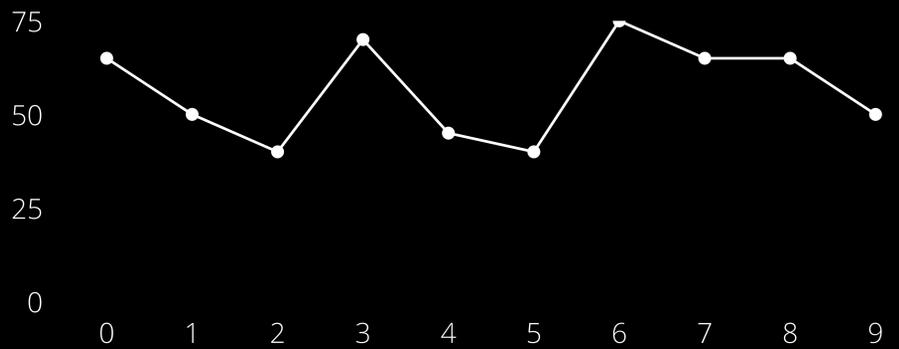


研究 AlphaZero 強化式學習流程並且應用在 Ultimate Tic Tac Toe

摘要 此專題研究如何以人工智能研究公司 DeepMind 開發的程式 AlphaZero 掌握棋類遊戲。為此，我們選擇了 Ultimate Tic Tac Toe 作為中高難度遊戲作為本研究的主題。該程式基於深度學習、強化式學習和賽局樹演算法。要獲得表現良好的 AI，必須進行訓練，並且需要相對較長的時間。在有限的硬體設備下，我們訓練了一個模型來與人類玩 Ultimate Tic Tac Toe 遊戲。在整個過程中，我們記錄了 AI 的表現結果，並見證了它隨著時間的推移而改進。

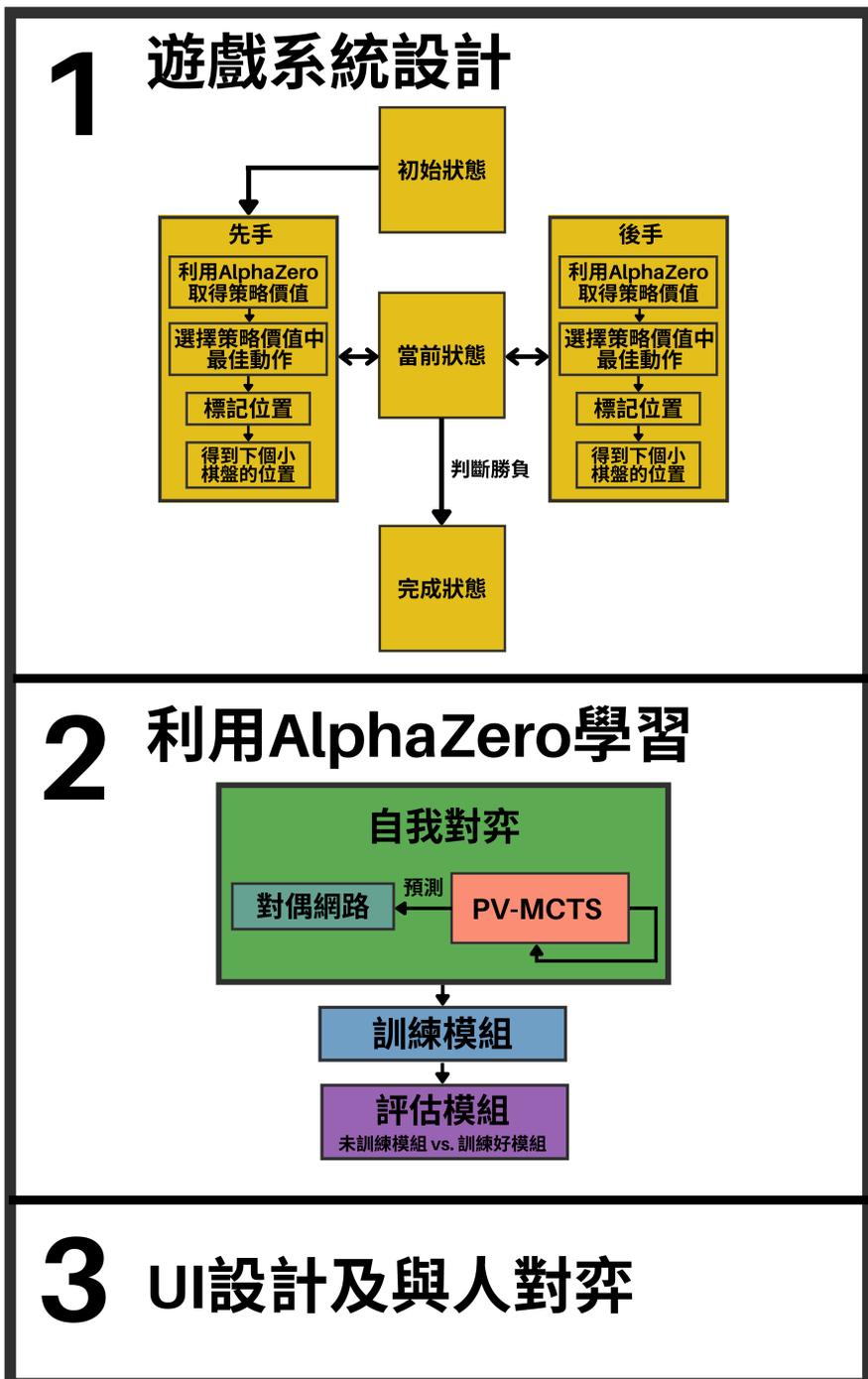
成果展示

未訓練模組 vs. 訓練好模組在每個訓練循環獲得的平均分數

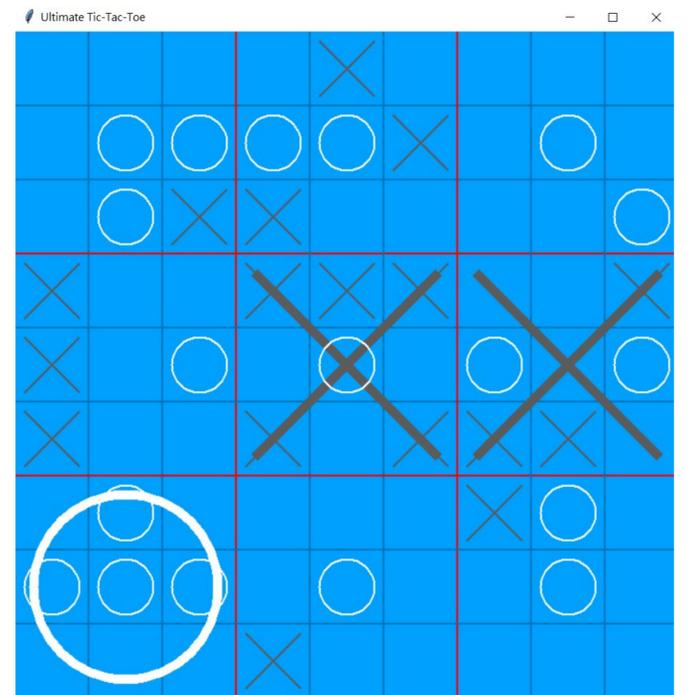


*1個訓練循環 = 1000次自我對弈

系統流程



UI設計



結論與展望

這次的專題我們成功的應用 AlphaZero 演算法。與人類的對戰過程中，可以發現它有在進步，從一開始未經訓練的模型，到最後，可以發現它與我們的思考越來越相近。在專題的最後，我們見證了一個 AI 的成長，如何學到一個遊戲的策略。AlphaZero 是一個偉大的人工智慧，但是它的缺點是需要很強的設備去訓練。未來若是有機會，我們可能會開發更高效的強化式學習演算法，讓它訓練效果提升。