

控制 MMORPG 內通貨膨脹：基於強化學習的 NPC 商店多商品動態定價

Controlling In-Game Inflation in MMORPGs: Reinforcement Learning-Based Multi-Product Dynamic Pricing for NPC Stores

指導教授：陳朝鈞

專題成員：張凱盛

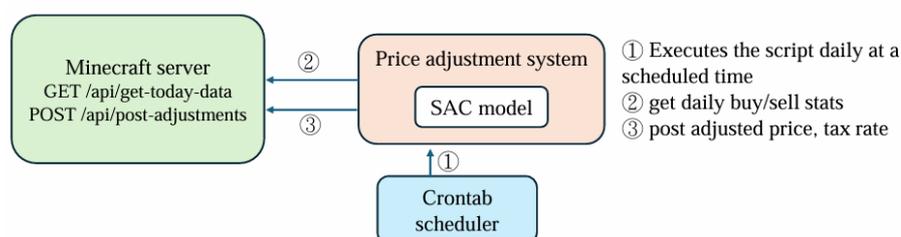
開發工具：Python、Pytorch、OpenAI Gym

測試環境：Windows11、Ubuntu

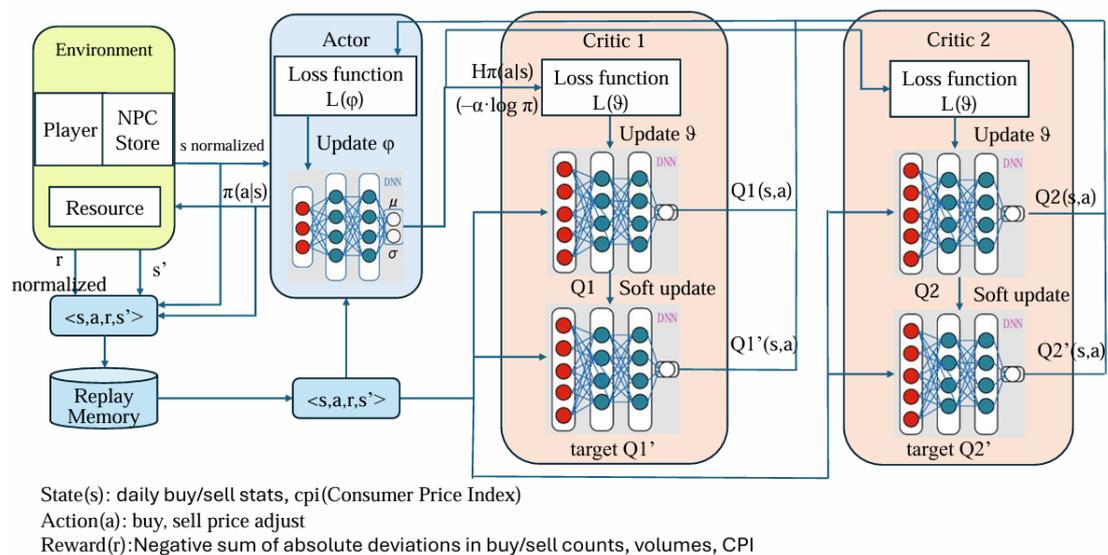
一、簡介：

遊戲通貨膨脹是指貨幣供應量過多導致貨幣價值下降，例如某些資源過於容易獲取，玩家可透過高效率農場或重複性任務迅速累積，並大量販售給遊戲商店換取金幣。如果沒有適時調整商品價格，金幣供應量將急遽增加，導致金幣購買力下降最終失去價值。然而依靠人工每日調整商品價格既耗時又難以及時因應市場變化。為實現自動化與即時的經濟調控，我們引入一個以 Soft Actor-Critic (SAC) 為核心的強化學習模型。該模型根據每日交易狀況，動態調整價格，透過與環境的互動學習最優策略。強化學習即時調整與長期回饋的特性，讓模型能有效掌握市場變化，實現供需平衡。

系統每日會將商店所有商品的買入與賣出狀況，整合成一個「狀態」輸入到 SAC 的 Actor-Critic 網路。Actor 輸出一個針對所有商品組合的價格與稅率微調向量，(既能抑制單品逐一調整帶來的套利機會，也能兼顧不同商品間的互補與替代關係)；Critic 則評估全局調整後的結果。獎勵函數以買賣人次還有商品成交數量偏離理想值的負絕對差為主軸，並適度懲罰價格跳動過大以鼓勵平滑穩定的調價策略。以下為系統架構圖：



圖一：系統架構圖



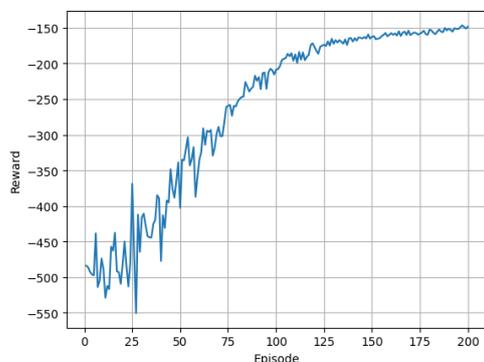
圖二：SAC 架構圖

二、測試結果：

由於強化學習模型的訓練需仰賴大量的訓練回合才能學習到穩定有效的策略，而實際遊戲中每日玩家的交易次數有限，無法提供足夠的資料以支撐完整的訓練流程。為此，本研究基於玩家的歷史交易紀錄，利用 **OpenAI Gym** 框架自訂並建構了一套模擬環境，使模型能在脫離真實遊戲的情境下進行大量測試與強化訓練。儘管模型已實際部署於遊戲平台，本研究的測試數據仍主要來自此模擬環境，以確保在不干擾真實玩家體驗的前提下，驗證模型的學習成效。

以下為測試結果：

以調整十個商品價格，玩家人數為200人為例，由圖三可以看到在經過25個 episode 後 reward 明顯上升，並在125個 episode 左右收斂，(reward function 在圖二)。



圖三：SAC 價格調整模型 Reward 圖