

HW 5

一、專案簡介

`mini_behavior` 是由 Stanford Vision and Learning Lab 發布的強化學習 (Reinforcement Learning, RL) 研究平台，旨在提供一個輕量、模組化、可擴展的仿真環境，用於研究智能體在「物理互動與任務導向行為」中的學習能力。

這個平台專注於 **合成模擬環境中的操作性任務 (manipulation tasks)**，如：撿拾物體、堆疊積木、按按鈕、打開抽屜等，並支援視覺輸入與多種任務設計，非常適合用於：

- 行為學習 (behavior learning)
- 多任務學習 (multi-task learning)
- 模仿學習 (imitation learning)
- 決策策略 (policy learning) 訓練與分析

二、核心功能與實現方式

1. 模擬環境設計

- 基於 OpenAI Gym 架構進行擴展，維持與標準 RL 接口兼容。
- 所有任務都定義在 `mini_behavior/envs/tasks/` 資料夾下。
- 具備模組化的場景設定（如地板大小、物體擺放、交互機制等）。

2. 任務多樣性與通用性

- 預設支援數個經典任務：
 - `pick_place`：撿物 → 放到目標位置
 - `stack_blocks`：堆疊兩塊積木
 - `button_press`：走到按鈕處並觸發
 - `open_drawer`：打開抽屜
- 任務皆支援目標設定 (goal-oriented)、視覺輸入 (RGB image observation)、獎勵設計等。

3. 訓練流程與工具整合

- 使用 `RLlib` 進行訓練（高度模組化，可切換 DQN、PPO、SAC 等演算法）
- 整合 `WandB` 與 `TensorBoard` 作為視覺化訓練監控平台。
- 提供指令式入口：
 - `train.py`：啟動指定任務的訓練流程
 - `rollout.py`：使用訓練好的模型執行行為並產生 `.mp4` 結果影片

三、可應用領域與研究價值

可應用於：

- 強化學習演算法的快速原型驗證（特別是行為學習任務）
- 調查智能體在視覺→動作決策上的泛化能力
- 多任務或遷移學習環境測試平台
- 行為模仿與計劃性探索（behavioral cloning, planning with RL）

研究價值：

- 提供一致性且高度控制的實驗環境，有利於學術比較與重現。
- 在不需實體機器人的前提下模擬實體操作任務，降低研究門檻。
- 適合延伸整合 vision-language policy、指令式任務描述等先進主題。

Result:



