

AI Final Project

『Text World』-- A Text Adventure Game Environment

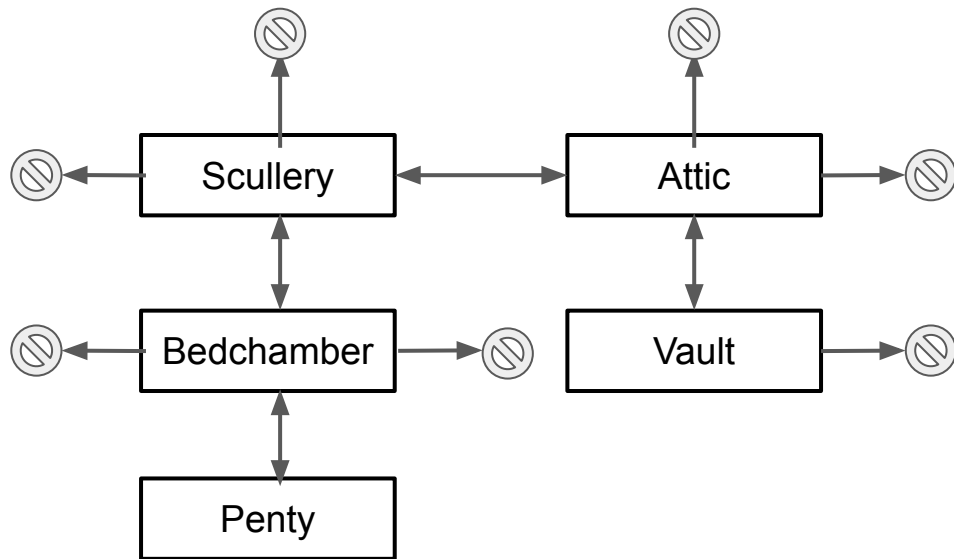
Alexis Lin

環境說明

- Text World 是一個以Python開發的文字冒險類遊戲沙盒(開發環境), 包括遊戲創建和遊玩兩大功能。
- 以 Text World 創建一個遊戲
 - 參數: 5個房間 / 10個物件 / 5個完成遊戲所需最少指令 (一個指令即為一個步數)
 - 指令: `tw-make custom --world-size 5 --nb-objects 10 --quest-length 5 --seed 1234 --output gen_games/`
- 撰寫一個Agent與Text World遊戲環境互動
 - `game_state, reward, done = env.step(command)`
- 遊戲環境提供函式讓agent取得環境資訊
 - `env.compute_intermediate_reward()` 取得立即回饋分數
 - `game_state.description` 取得目前所在房間的狀態的描述文字
 - `game_state.admissible_commands` 取得目前狀態下可下的指令
 - `game_state.state` 取得目前所在房間的位置資訊 (JSON格式)

互動觀察

- 互動方式
 - 以人工方式與遊戲互動
- 資訊取得
 - 遊戲環境僅會顯示部分資訊，沒有明確的遊戲目標指示。
- 房間配置
 - (如右圖)
- 完成遊戲
 - insert the keycard into the toolbox



互動觀察

- 互動方式
 - 以程式與遊戲互動
- 資訊取得
 - 可取得明確的遊戲目標
 - 可取得明確的遊戲環境的描述
 - 可查詢已取得的物品
- 立即回饋
 - 進入正確的房間(go east / go west / go south / go north)會得 +1 分
 - 進入不正確的房間(go east / go west / go south / go north)會得 -1 分
 - 其他都是 0分

進行實驗

實驗一：隨機 (get command randomly)

- 實驗方式

- 以所在房間作為 state
 - room = ['scullery', 'attic', 'bedchamber', 'pantry', 'vault']
- 以 game_state.admissible_commands 取得可執行的 command list
- 以隨機方式取出一個 command 代入 env.step(command) 函式執行
- 10 episodes / 5000 steps in a episode

- 實驗結果

- 可執行的動作 (action space) 共有 110 個
- 平均用 2,195 steps 完成遊戲

```
=====COMPLETED=====
steps: 838
=====COMPLETED=====
steps: 165
=====COMPLETED=====
steps: 1020
=====COMPLETED=====
steps: 352
=====COMPLETED=====
steps: 4144
=====COMPLETED=====
steps: 4734
=====COMPLETED=====
steps: 351
=====COMPLETED=====
steps: 339

avg. steps: 2195.1;
avg. score: 0.8 / 1.
```

進行實驗

實驗二：從記錄中隨機取動作

● 實驗方式

- 以所在房間作為 state
- 以 `game_state.admissible_commands` 取得可執行的 command list
- 以 `env.compute_intermediate_reward()` 取得立即回饋分數，並將 (state, command, intermediate_reward) 資訊記錄起來，並持續累加 intermediate_reward
- 在某比例下，從 command list 中隨機取出一個 command 代入 `env.step(command)` 函式執行
- 在某比例下，從回饋記錄中取出最高分的動作
- 10 episodes / 5000 steps in a episode

● 實驗結果

- 可執行的動作 (action space) 共有 106 個
- 平均用 427 steps 完成遊戲

```
=====COMPLETED=====
steps: 32
=====COMPLETED=====
steps: 434
=====COMPLETED=====
steps: 252
=====COMPLETED=====
steps: 488
=====COMPLETED=====
steps: 866
=====COMPLETED=====
steps: 376
=====COMPLETED=====
steps: 121
=====COMPLETED=====
steps: 703
=====COMPLETED=====
steps: 217
=====COMPLETED=====
steps: 768
```

```
avg. steps: 426.7;
avg. score: 1.0 / 1.
```

進行實驗

實驗三：限制動作選項，隨機取動作

- 實驗方式

- 以所在房間作為 state
- 以 `game_state.admissible_commands` 取得可執行的 command list
- 篩選出以 go, insert, take 開頭的動作存放在 `good_command_list`
- 從 `good_command_list` 中隨機取出一個 command 代入 `env.step(command)` 函式執行
- 10 episodes / 5000 steps in a episode

- 實驗結果

- 可執行的動作 (action space) 共有 28 個
- 平均用 46 steps 完成遊戲

```
=====COMPLETED=====
steps: 33
=====COMPLETED=====
steps: 17
=====COMPLETED=====
steps: 80
=====COMPLETED=====
steps: 72
=====COMPLETED=====
steps: 67
=====COMPLETED=====
steps: 21
=====COMPLETED=====
steps: 55
=====COMPLETED=====
steps: 47
=====COMPLETED=====
steps: 35
=====COMPLETED=====
steps: 22

avg. steps: 45.9;
avg. score: 1.0 / 1.
```

進行實驗

實驗四：限制動作選項，從記錄中隨機取動作

● 實驗方式

- 以所在房間作為 state
- 以 `game_state.admissible_commands` 取得可執行的 command list
- 篩選出以 `go`, `insert`, `take` 開頭的動作存放在 `good_command_list`
- 以 `env.compute_intermediate_reward()` 取得立即回饋分數，並將 `(state, command, intermediate_reward)` 資訊記錄起來，並持續累加 `intermediate_reward`
- 在某比例下，從 command list 中隨機取出一個 command 代入 `env.step(command)` 函式執行
- 在某比例下，從回饋記錄中取出最高分的動作
- 10 episodes / 5000 steps in a episode

● 實驗結果

- 可執行的動作 (action space) 共有 28 個

```
=====COMPLETED=====
steps: 48
=====COMPLETED=====
steps: 48
=====COMPLETED=====
steps: 96
=====COMPLETED=====
steps: 91
=====COMPLETED=====
steps: 31
=====COMPLETED=====
steps: 40
=====COMPLETED=====
steps: 69
=====COMPLETED=====
steps: 29
=====COMPLETED=====
steps: 33
=====COMPLETED=====
steps: 15

avg. steps: 51.0;
avg. score: 1.0 / 1.
```