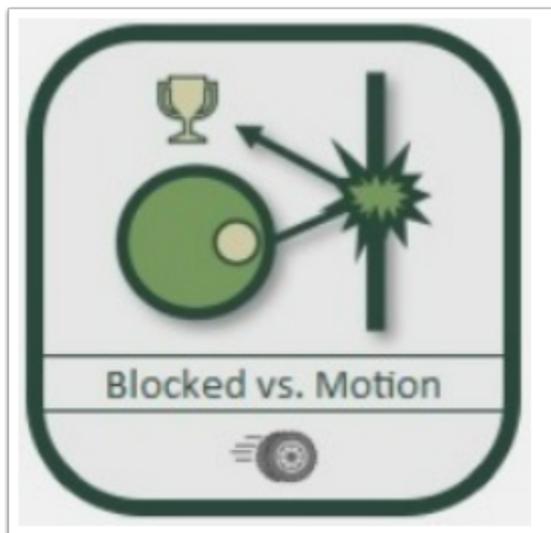


單元十：強化式學習法 - 動 vs. 不動（初階）

學習目標

在這一個單元，會為大家介紹一種高階的AI 學習法 - 強化式學習法。透過正確設定獎勵（和懲罰）機制，機械人可以自學得到智力。

實驗的目標是要訓練機械人可以在四方形的場地內不停地走動。



學習內容

強化式學習法（機械人自學）的原理是透過獎勵和懲罰去鼓勵機械人去做對某些動作。機械人會去改變它的行為從而獲取最高的累積的平均獎勵分數。

- 向前行走得越快越多，會得到更多獎勵。相反，停頓或向後行，就會被懲罰（被扣分）。
- 要將等級提升，機械人要盡力向前行，同時間要減少停頓或向後行。

活動

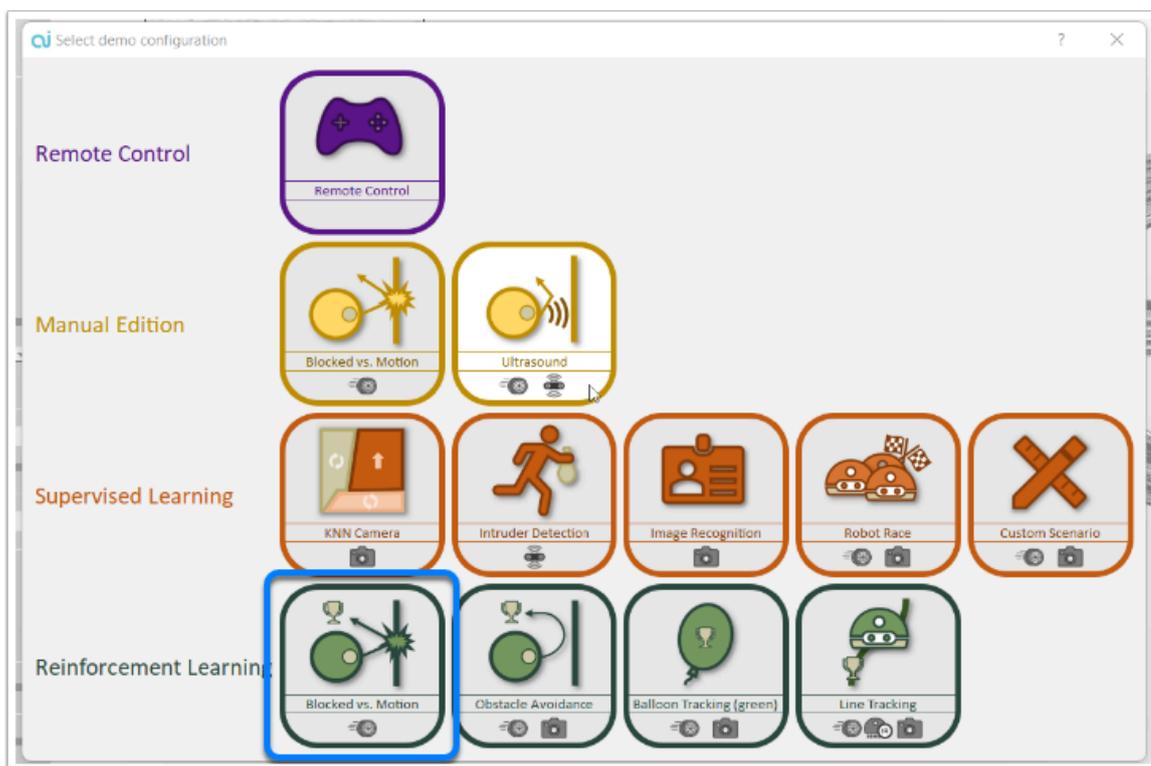
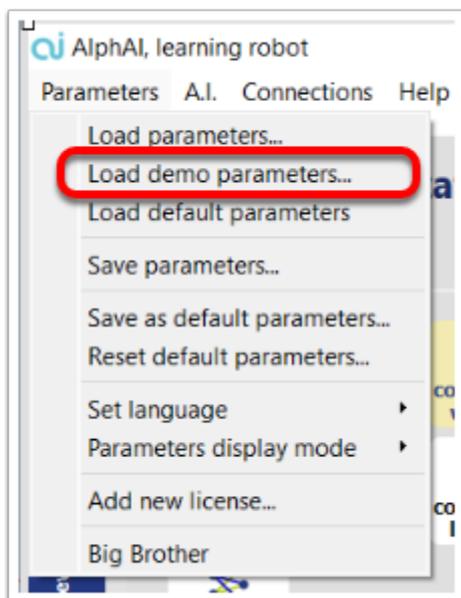
材料：

- 四方形有四面紅色牆的賽場。

AI 設置的參數

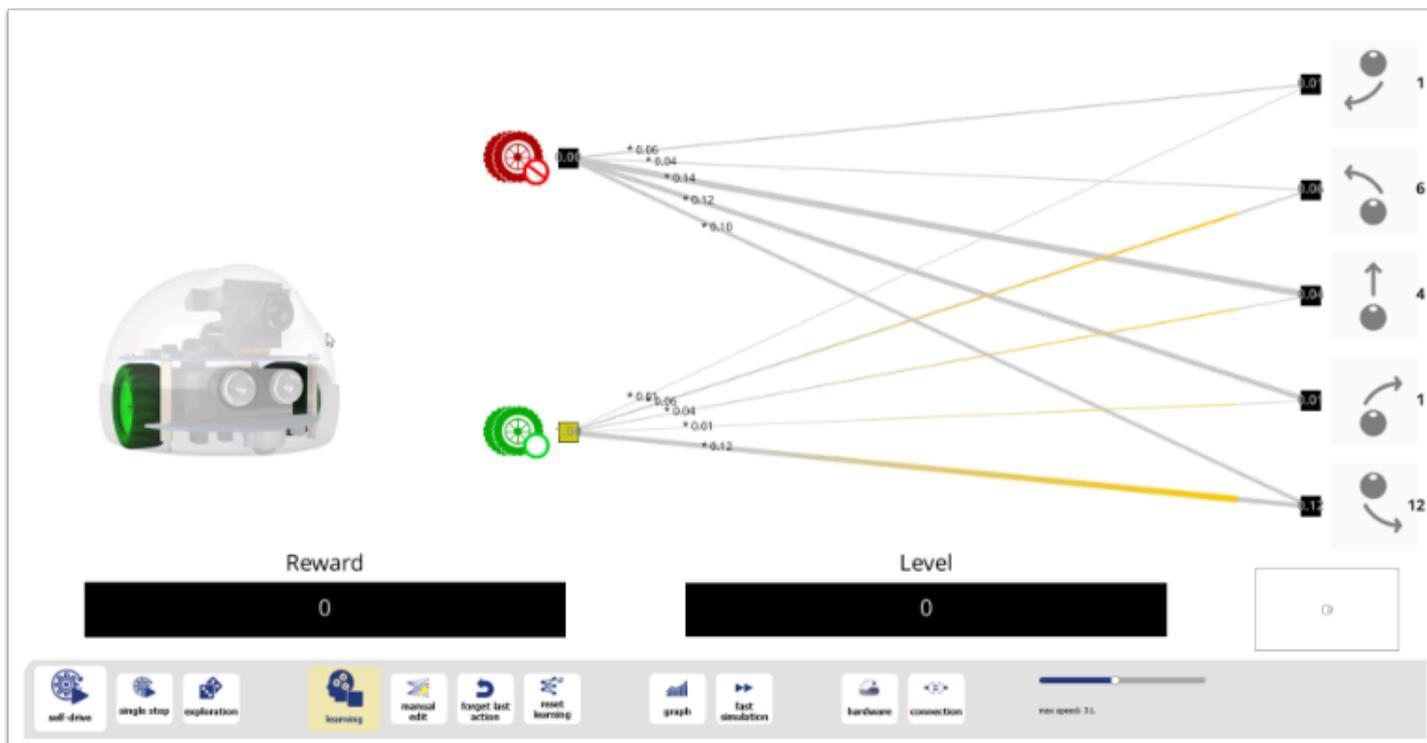
- 用 參數 <Parameters> 下拉餐單的“載入示範參數” <Load demo parameters ..>

- 像下面一樣，揀選“動 vs. 不動” <Blocked vs. Motion>。



開始訓練

- 啓動AlphaAI 機械人及聯線到你的電腦。需要時按螢光幕上的“聯線” <Connection> 按鈕。
- 如果設定參數正確，螢光幕將會顯示以上的神經網絡。



獎勵 (Reward) 和等級 (Level)

向前移動將會得到獎勵

- 向前行 **+100 分**
- 向前左轉 **+30 分**
- 向前右轉 **+30 分**

懲罰

- 停止 (例如撞牆) 或向後行向後轉 **-50 分**

等級的計算：

- 在過去2分鐘得到獎勵的平均數

標籤輸出

- 根據動作的意思解釋

(A) 實驗1 - 沒有學習的駕駛

- 啓動“重設學習” <reset learning> 的按鈕。
- 根據下圖
 - 關閉“學習” <learning> 按鈕。
 - 關閉“探索、作新嘗試” <exploration> 按鈕。
 - 啓動“自動駕駛” <self drive> 的按鈕。
- 留意機械人的行為
-

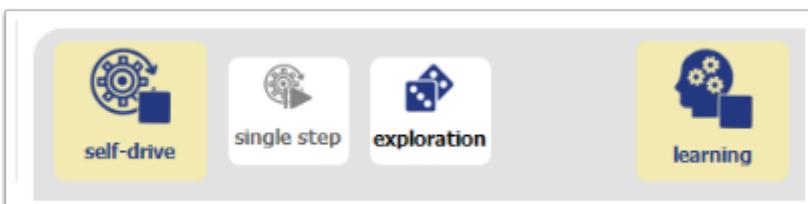


討論

- 形容機械人的行為。它怎樣移動呢？
- 請問“獎勵和等級”有沒有改變呢？如何改變？
- 機械人有沒有學習到什麼？

(B) 實驗2 - 沒有“探索”的學習

- 啓動“重設學習” <reset learning> 的按鈕。
- 根據下圖
 - 啓動“學習” <learning> 按鈕。
 - 關閉“探索、作新嘗試” <exploration> 按鈕。
 - 啓動“自動駕駛” <self drive> 的按鈕。
- 留意機械人的行為

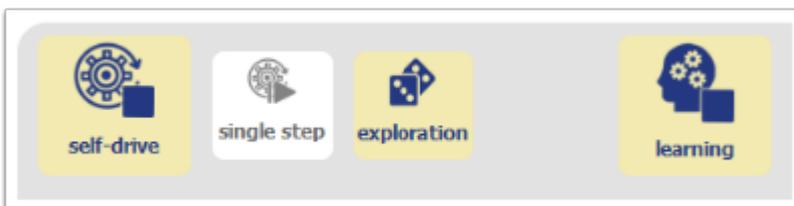


討論

- 形容機械人的行為。它怎樣移動呢？
- 請問“獎勵和等級”有沒有改變呢？如何改變？
- 機械人有沒有學習到什麼？
- 機械人是否已經滿足得到的獎勵和等級呢？

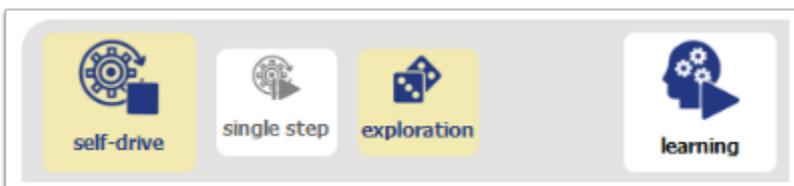
(C) 實驗3 - 會嘗試“探索”的學習

- 啓動“重設學習” <reset learning> 的按鈕。
- 根據下圖
 - 啓動“學習” <learning> 按鈕。
 - 啓動“探索、作新嘗試” <exploration> 按鈕。
 - 啓動“自動駕駛” <self drive> 的按鈕。
- 留意機械人的行為。
- 當等級達到50，可以停止實驗。



討論

- 形容機械人的行為。它怎樣移動呢？
- 請問“獎勵和等級”有沒有改變呢？如何改變？
- 機械人有沒有學習到什麼？
- 機械人可以拿到 +100分的獎勵嗎？
- 當箭阻轉藍色，代表機械人正在探索中。你有沒有留意到呢？
-



測試

- 讓我們測試，透過強化式學習法學習到的智能是否有用？

- 關閉學習” <learning> 按鈕。
- 啓動“探索、作新嘗試” <exploration> 按鈕。
- 啓動“自動駕駛” <self drive> 的按鈕。
-
- 讓機械人自由走動 3 到 5 分鐘。
- 看看它得到的最高的等級是多少？

- 學習到的智能是否運作良好呢？是否能夠保持機械人不斷走動呢？

討論

- 試形容什麼是強化式學習法？

- 「探索、作新嘗試」對自學是否重要？

- 要強化式學習法成功，最重要的因素是什麼？