



利用模擬結合強化學習於排程之應用

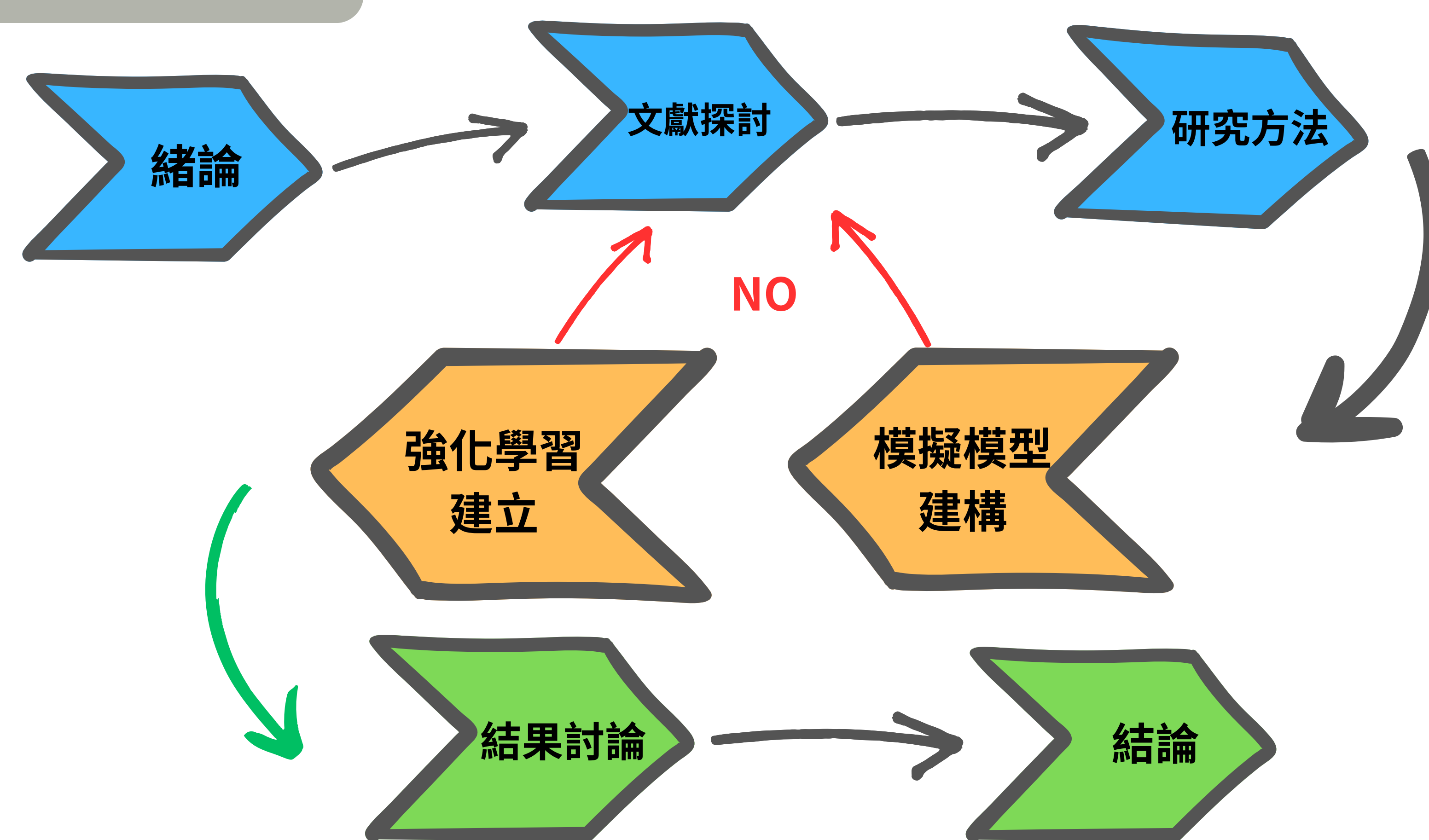
SHOP SCHEDULING USING COMPUTER SIMULATION AND REINFORCED LEARNING

指導教授:任恒毅 研究生:陳冠廷、林保岑、黃柏凱

研究背景目的

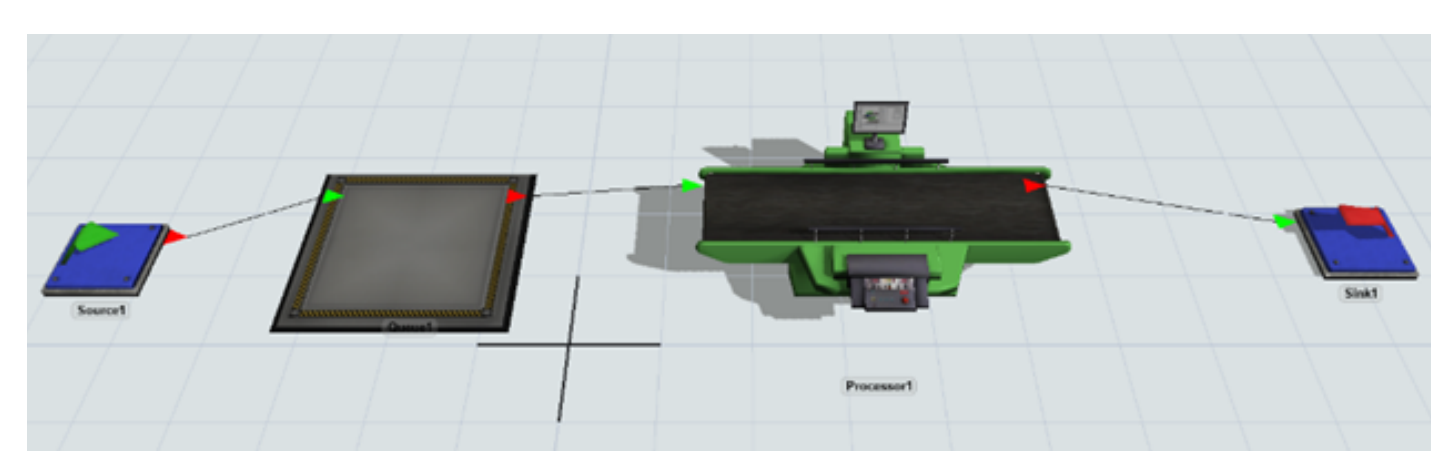
現代科技進步，傳統生產排程方法已無法滿足當前市場靈活性和效率，因此探索最佳排程解決方案變得尤為重要。模擬軟體技術使企業能夠在虛擬環境中進行生產流程的優化更新等，但單純依賴模擬技術仍存在一定局限性。強化學習作為一種基於經驗的自我學習技術，能在模擬環境中學習最佳排程策略，並隨著環境變化及時調整其行為。本研究希望能透過模擬與強化學習的結合，為未來的生產排程提供創新且有效的解決方案。

研究架構



研究方法

1.設立模型

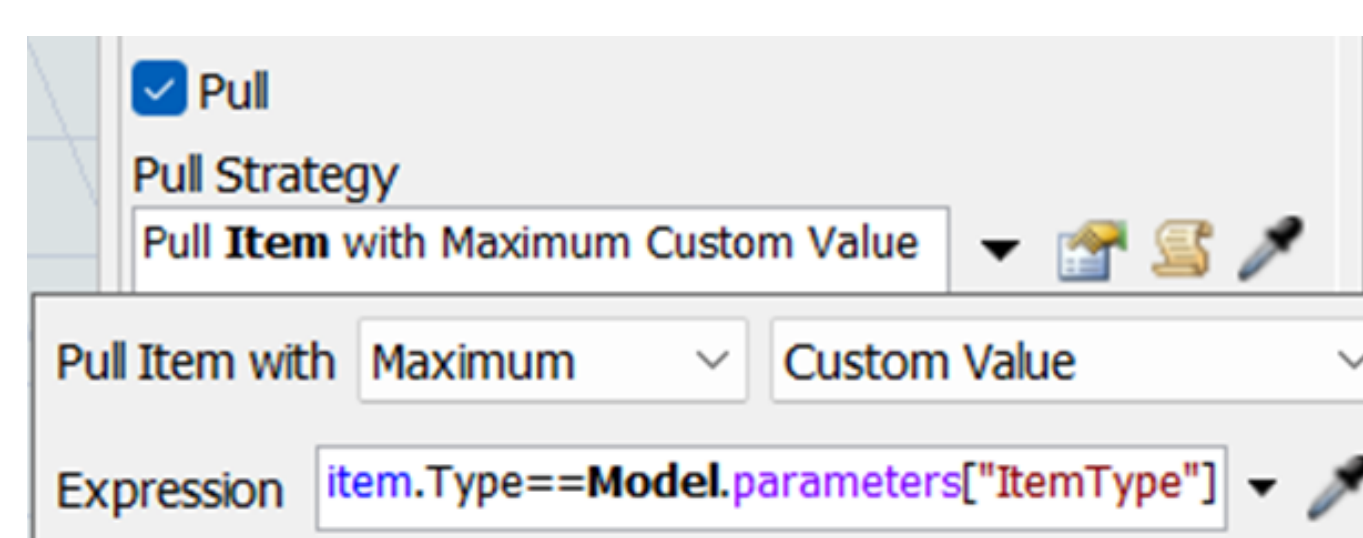


(包含Source、Queue、Processor、Sink各一)

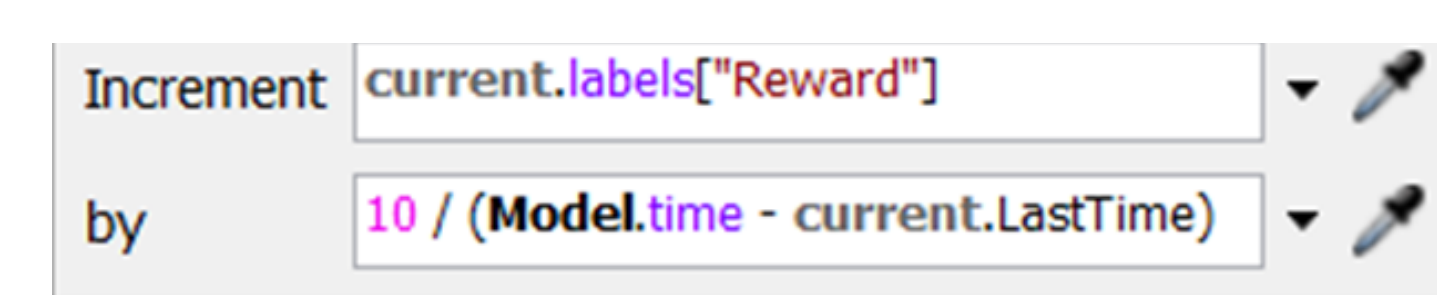
2.Flexsim設定

Model	Changeover Times	Type	Observations	Type	Actions
Col 1	Col 2	Col 3	Col 4	Col 5	
Row 1	0	10	20	30	40
Row 2	10	0	10	20	30
Row 3	20	10	0	10	20
Row 4	30	20	10	0	10
Row 5	40	30	20	10	0

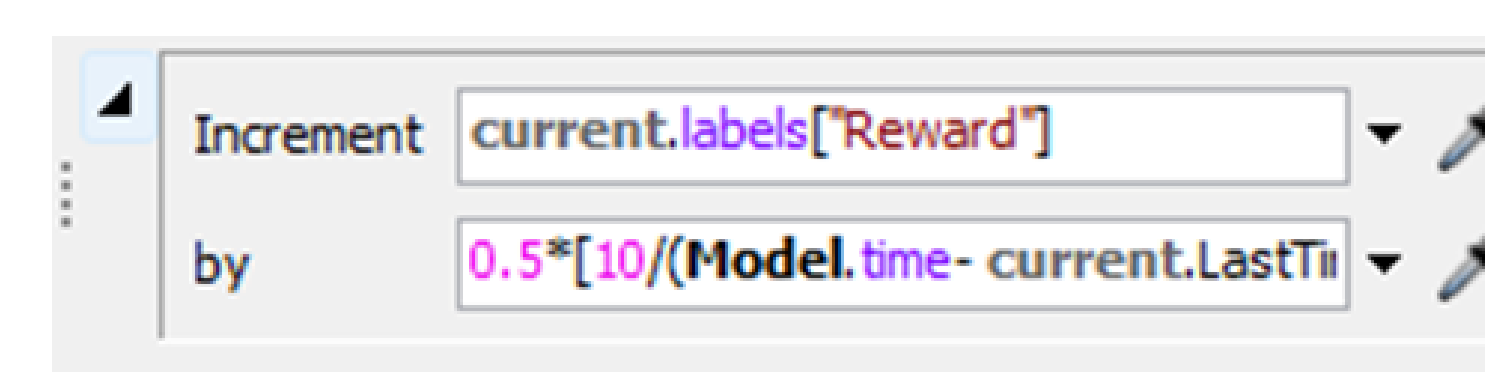
產品設置時間的轉換時間表



設定強化學習所需之拉式策略

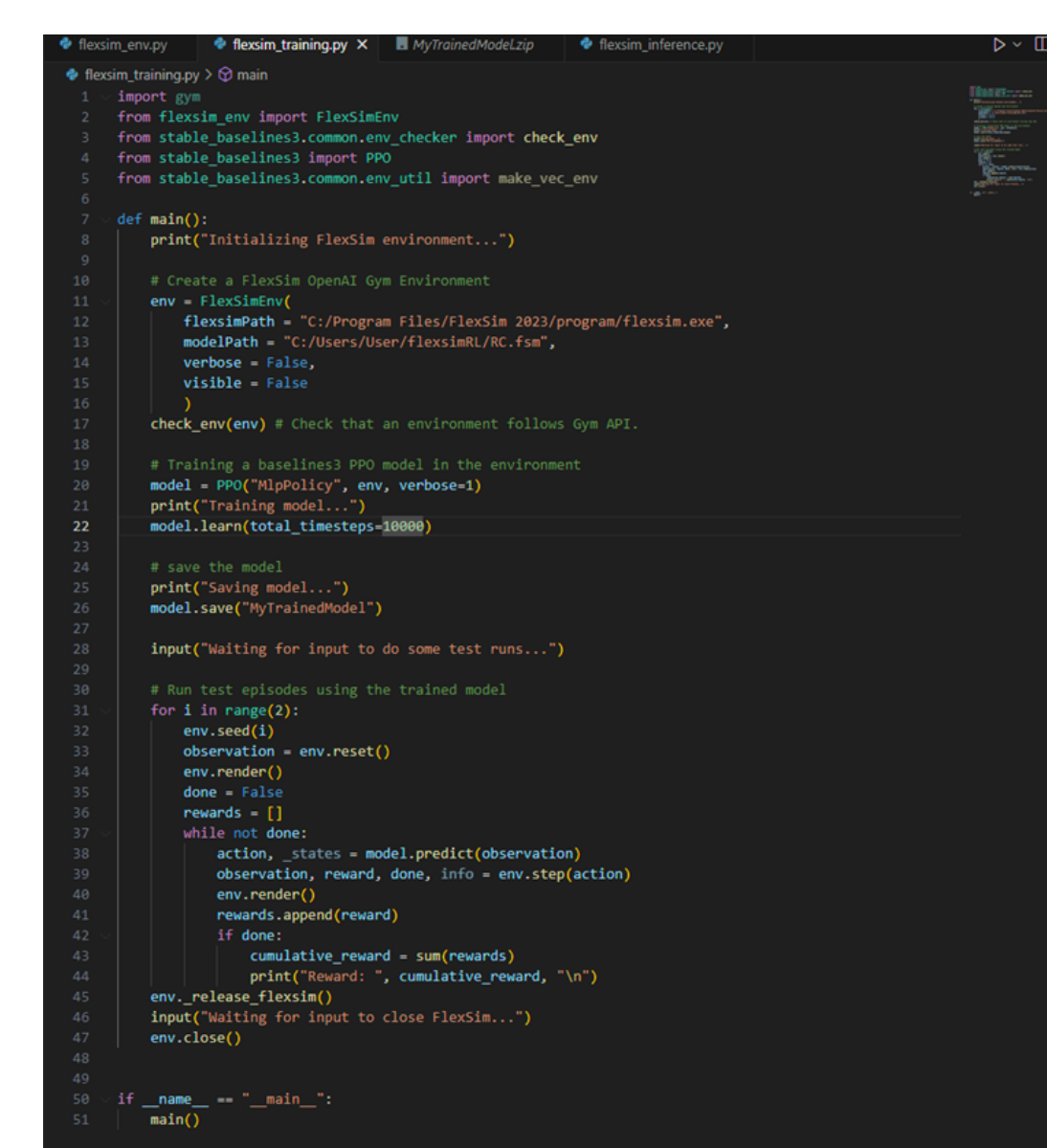


初始獎勵函數，只考慮最短生產間隔



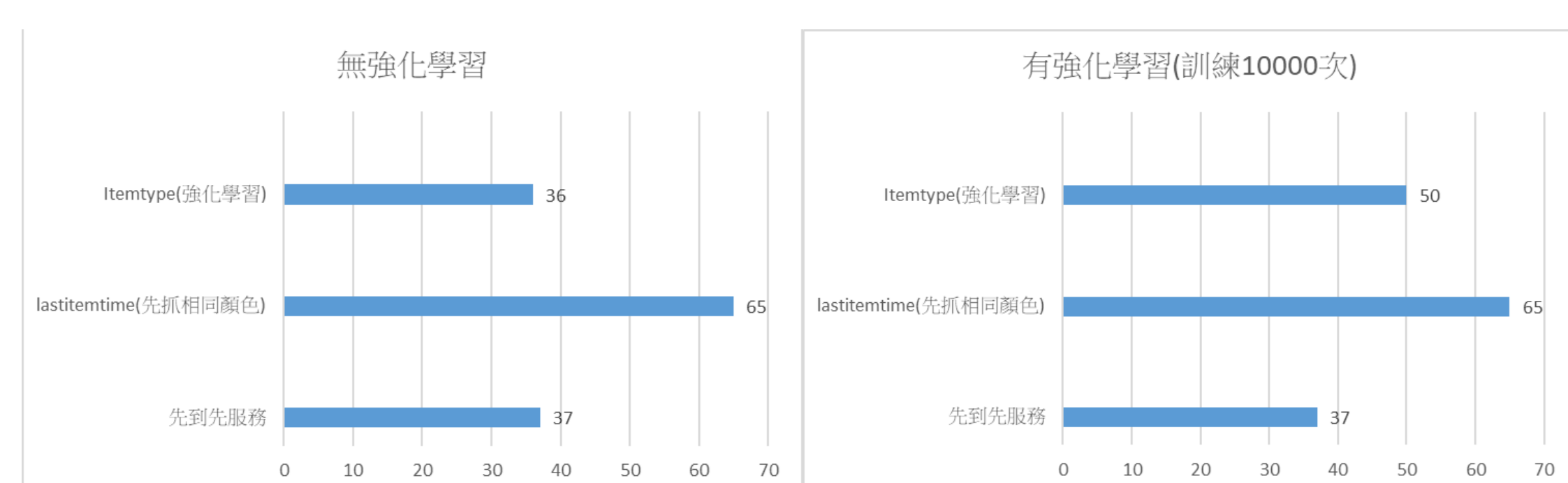
修改後的獎勵函數，可透過修改權重來設定要以最短設置時間或是優先生產價值高之產品

3.強化學習設定

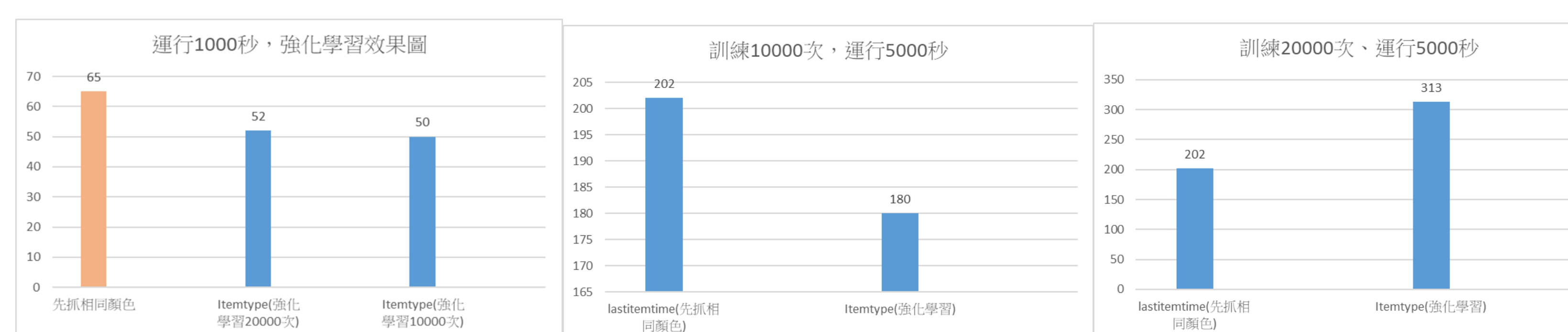


python 程式碼主要使用在訓練模型和推論訓練完成的模型

研究結果分析



- 無強化學習的生產總量可看出**優先生產的派工法則最高**
- 有強化學習的生產總量可看出**強化學習產量明顯提升**



- 拉長運行時間跟增加訓練次數**都能有效地提升強化學習效果

結論

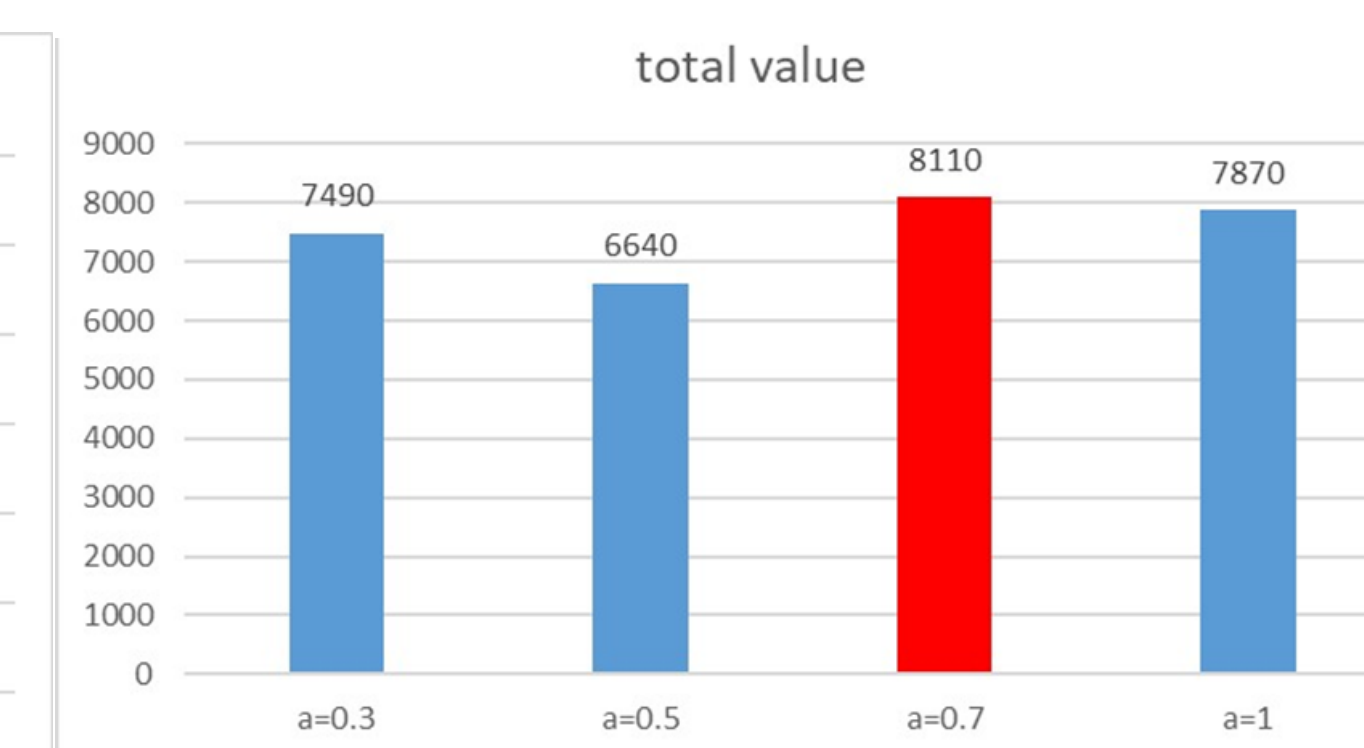
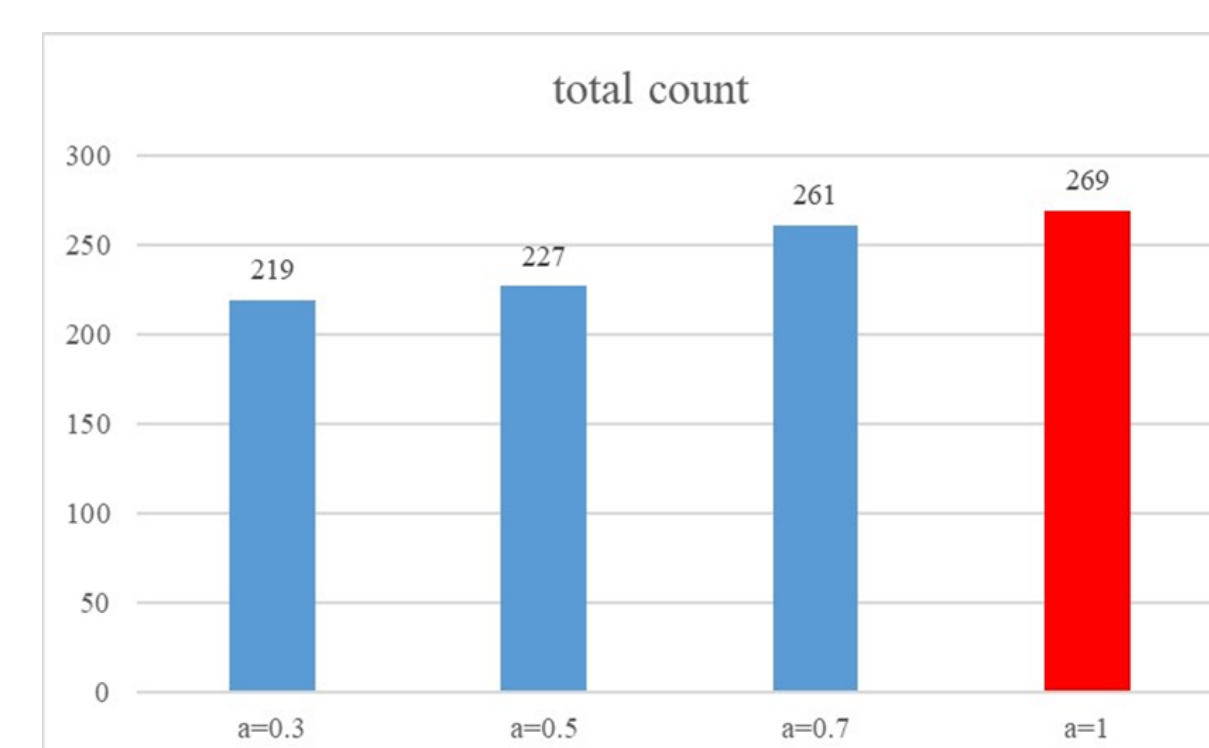
強化學習效果

- 在沒有強化學習下，優先生產相同產品的拉式策略是最好的
- 拉長運行時間跟增加訓練次數**都能有效地提升強化學習效果，且最終生產總量是最高的

考量生產價值的模型

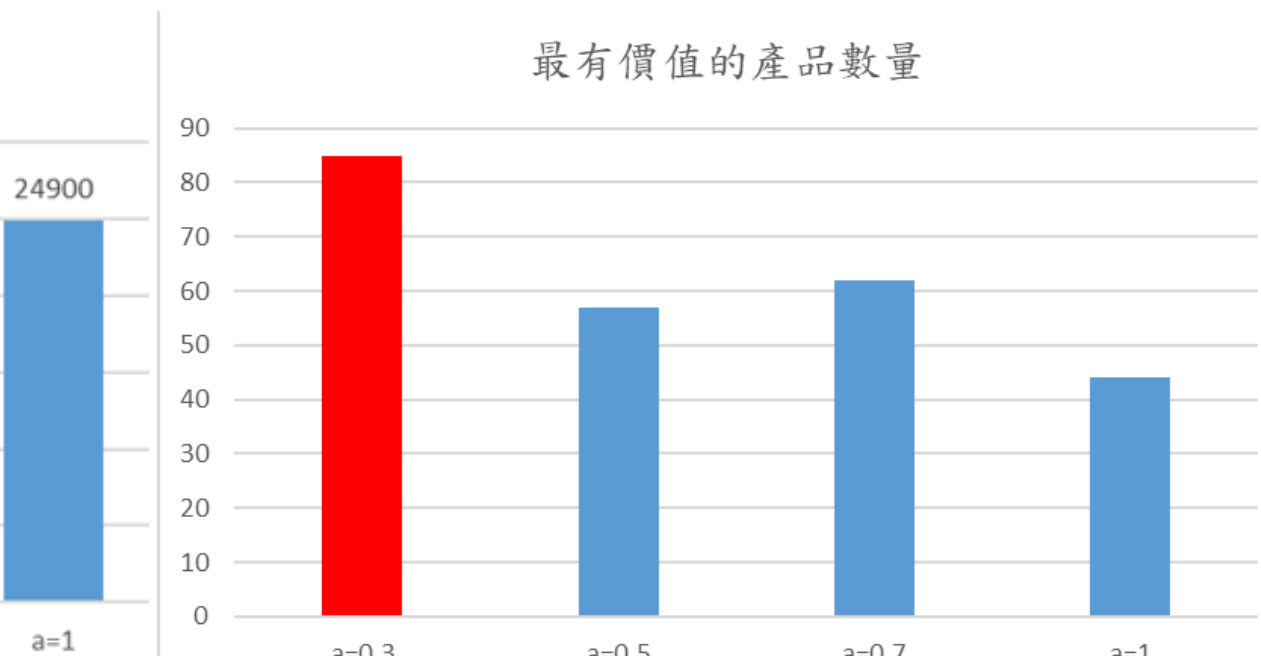
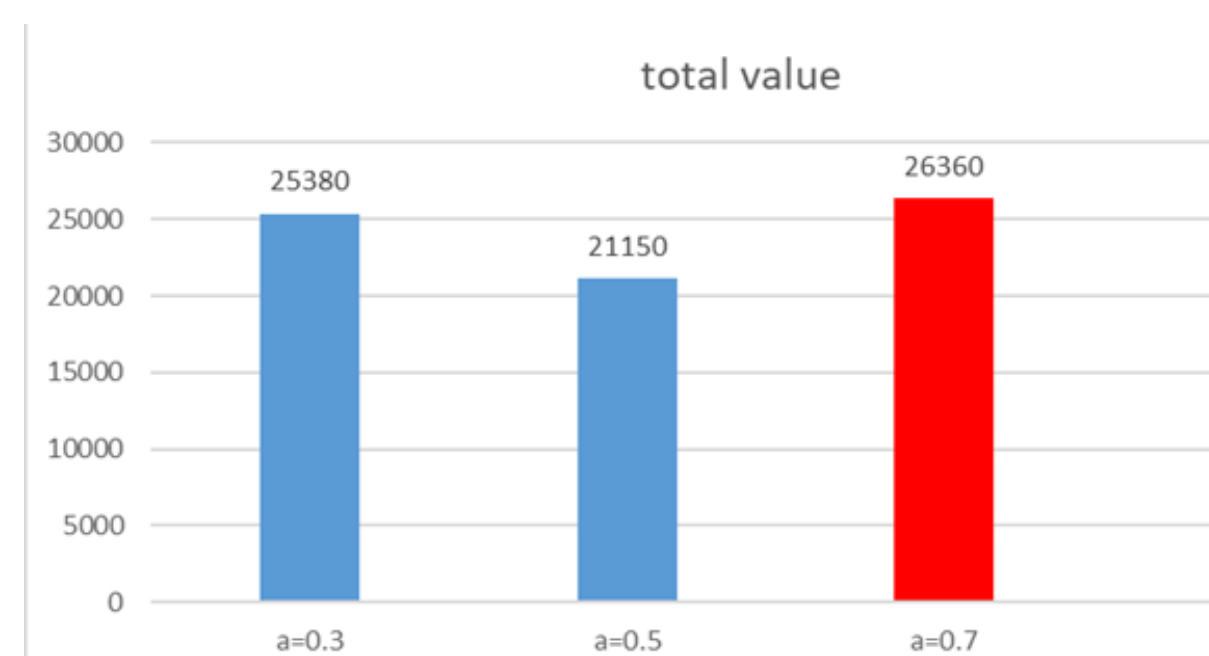
- 在a=0.7時能有較高的產出量且能有最高的總生產價值
- 權重的設計確實能引響強化學習的拉式策略且a越小能優先生產高價值產品，但總價值不一定最高

TYPE	VALUE
1	10
2	20
3	30
4	40
5	50



- a=1時的生產總量最高
- 產品價格差距小的情況下，a=0.7的生產總價值最高

TYPE	VALUE
1	10
2	50
3	90
4	140
5	190



- 在產品價格差距大的情況下，a=0.7的生產總價值最高
- a**越小**能生產出更多的高價值產品，但對總價值的增加引響不大