

此次實作以GA+SSE的方式來解決花的種類問題

正解的SSE計算出的結果為89.2974

可見SSE與正解的關係不為完全正相關。（SSE過大的情形可能導致求出解的正確率下降）

此次實作發現與之前01問題最大的不同是，01問題進行Evaluation時只需考慮1是否更多，問題較為簡單

而這次需考慮到各個種類、各項資料的SSE總和，問題較為複雜

一開始實作在crossover時是以單一點交換的方式來進行(與之前解01問題時一樣的方式)，來解決此次問題，發現在很前面的Evaluation下就會收斂，收斂約在500-600之間。後來思考後發現若是把diversity設高，也就是更為隨機的方式進行crossover（此次實作利用兩點交換方式再加上4點隨機點交換的方式進行crossover），效果顯著。

————— 以下為PR分析

利用PR將重複多次的點鎖住不進行計算，此次實作設定可以自訂evaluation，在重複次數達到自訂的evaluation下則將該點鎖住。

接著利用Evaluation的總數不變（POP和Iteration的數量也不改變）但PR的自訂evaluation改變來探討PR在不同的evaluation設定下對時間的影響，詳見下表

Evaluation的總數為POP和iteration相乘

實驗表格以POP為600，iteration為200，Run為30的情況下進行實驗及討論  
Data Set 為IRIS

表一、不同PR Evaluation的實驗數據表

PR_evaluation	PR_Lock_point Quantity	AVG_SSE	Best_SSE	AVG_Accuracy	Time(s)	SSE_Time(s)
120	122	162	105	77.91%	111.65	89.98
360	121	164	119	77.11%	110.622	88.93
600	122	165	115	76.96%	110.03	88.44
1200	121	154	114	78.29%	111.82	90.27
3000	121	139	94	80.98%	115.48	93.81
3600	121	132	96	80.87%	119.032	97.09
4800	122	121	84	82.98%	121.65	99.74
6000	121	110	86	84.36%	127.42	105.32
7200	121	101	86	85.91%	137.10	113.23
8400	121	97	83	87.36%	144.21	119.84

9600	121	91	82	87.96%	154.88	130.30
10800	119	92	82	87.89%	167.54	142.44
12000	108	88	82	88.24%	176.69	152.95
-	0	79	78	89.6%	230.37	206.76

註解：-為沒有啟動 P R 機制（詳細作法：設定的Evaluation超過pop\*iteration，因此程式架構皆完全相同)

## 表二計算公式

下表的time(%)以上表無啟動 P R 機制的總時間來做100%，時間越少所佔百分比越小  
 $(PR\_evaluation - AVG\_Time) / PR\_AVG - time$

下表的SSE(%)為運用PR後造成 S S E 變差的比率 $(PR\_evaluation - AVG\_SSE) / 未啟用PR\_AVG\_SSE$

以 S S E 來評斷精準性原因為，SSE和花種的準確率並無完全正相關，而程式是以SSE的方式來進行評斷的，因此採用SSE為精準性，較能判斷程式的實際情形。

紅色底色的PR為評估過時間和 S S E 準確度後，較佳的採用方式

若以總Evaluation(120000)和PR\_Evaluation(6000)的比值來做判定，以evaluation的 5%時來啟動 P R 機制的效果最好(有效縮短時間，結果也不錯)

表二、無PR機制和不同PR機制下的時間及SSE差值百分比表

PR_evaluation	Time(%)	SSE (%)
120	48%	205.06%
360	48%	207.59%
600	48%	208.86%
1200	49%	194.94%
3000	50%	175.95%
3600	52%	167.09%
4800	53%	153.16%
6000	55%	139.24%
7200	60%	127.85%
8400	63%	122.78%

9600	67%	115.19%
10800	73%	116.46%
12000	77%	111.39%
12000	100%	100%

## 結論

由上圖可以看出越早啟動PR機制，SSE的值會越高，時間則會越少，但在120-6000之間的時間差並不大，但SSE值差別卻頗大，從實驗結果可以看出選擇適當的PR機制對準確度和時間是極為重要的，若啟動太早雖然節省了時間，但卻犧牲了精確度，在適合的PR機制下，也不會多花太多時間，但能成功將精確度提高，有效的兼顧了時間和精確度。

在120-6000時的時間差異並不大，但SSE卻有很大的差異，以上圖看的話若要兼顧精確度和時間，000會是較好的選擇，以evaluation的比例來看大約為5%-6%左右(6000/120000-7200/120000)

