計算型智慧-作業二 RBFN參數基因最佳化

授課教授：蘇木春　教授

學號：104522065

學生：張翔珳

**1.使用環境**

作業系統：windows 7 or windows 8.1 64位元

使用語言：C

編譯平台：Dev C++ 5.11

**2.基因演算法架構**

m：0~30、：0~10、w：-40~40

目前測試結果較佳參數如表所示。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 參數項目 | 迭代次數 | 神經元數量 | 基因數量 | 交配機率 | 突變機率 | 突變雜訊大小 | 初始值 |
| 參數數值 | 1000 | 20 | 500 | 0.5 | 0.1 | -0.2~0.2 | 1 |

其主要流程一般基因演算法流程一樣，如圖1所示：

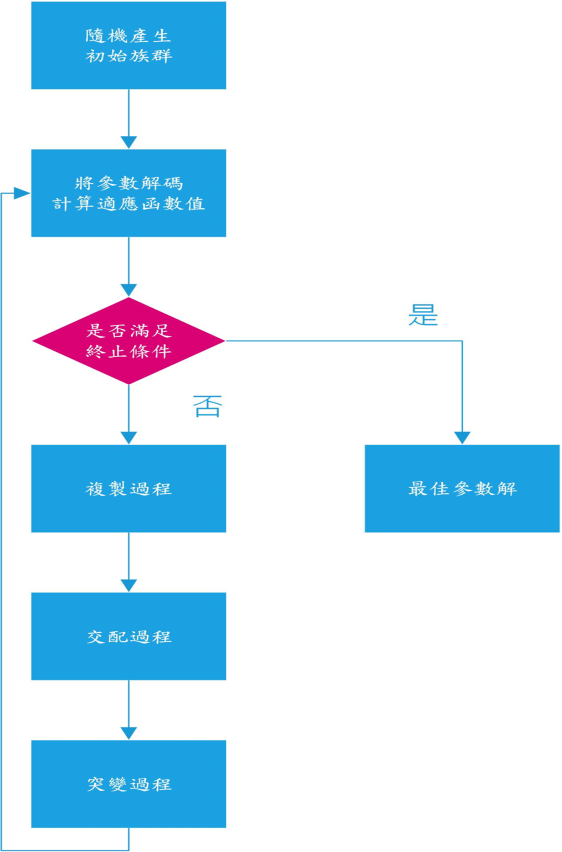


圖1 基因演化流程

採用實數型基因演算法，類神經網路RBFN是以20顆神經元的架構。

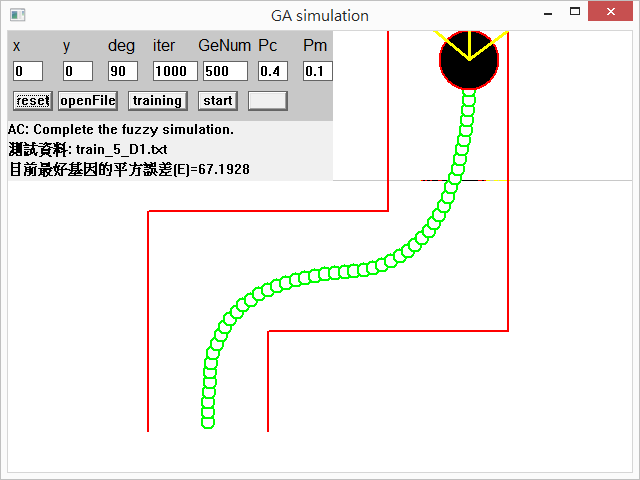
複製過程是取父代前50%的基因，再加上競爭選擇的50%基因，對每個基因加入-1~1的隨機雜訊，並乘上(1-目前迭代次數/總迭代次數)，讓基因複製會隨著迭代次數增加而減少加入的隨機雜訊，讓基因能夠從探索轉換到挖掘，總共為NUM\_OF\_GENE個基因(a，100%)。

交配方式是把交配池的每兩筆為一組，依照交配機率決定是否做交配，再隨機決定接近或遠離，以及隨機決定拉近或拉遠的程度(0~1)，這個階段最多交配出NUM\_OF\_GENE/2個基因(b，50%)。

突變方式除了隨機加上一亂數雜訊(-0.2~0.2)，並隨機產生NUM\_OF\_GENE/20個新的基因(c，5%)，再從a+b(至多原基因150%的數量)當中挑選出a - c個(d，95%)，最後以d + c個(100%)做下一代基因演化。

**3.如何使用本視窗程式?**

程式的執行結果如圖2所示，左上角提供參數設定，依序為起始點座標設定(x, y)、起始角度、迭代次數、基因數量、交配機率、突變機率。



動作按鈕

參數設定

目前狀態和過程中的輸入參數

訓練、測試檔案

適應函數值

圖2 程式結果

執行順序：(1)→(2)→(3)

(1)openFile：選擇訓練資料。

(2)training：依照目前設定參數檔進行訓練。

(3)start：依照訓練結果啟動自走車模擬。

(4)stop：暫停自走車模擬。

(5)reset：讓自走車回到起點，但保留訓練結果

**4.實驗參數對結果的影響**

**(1)迭代次數**(藍色為成功走完，紅色為非成功走完)

基因數量 = 500，Pc = 0.4，Pm = 0.1，使用資料集train\_5\_D1.txt

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 第幾次  迭代次數 | 1 | 2 | 3 | 4 | 5 | 6 |
| 2000 | 131.16 | 80.59 | 67.92 | 132.40 | 77.34 | 72.01 |
| 1000 | 110.35 | 70.48 | 108.68 | 78.40 | 106.83 | 126.42 |
| 750 | 261.65 | 138.50 | 82.25 | 168.85 | 129.23 | 95.11 |
| 500 | 100.65 | 152.69 | 216.24 | 189.57 | 174.85 | 95.14 |
| 300 | 175.76 | 212.21 | 151.66 | 136.82 | 125.50 | 133.17 |
| 100 | 283.29 | 414.47 | 274.51 | 440.90 | 260.78 | 548.54 |

※大約在1000次迭代次數有較佳、較穩定的結果。

**(2)族群大小**

迭代次數 = 1000，Pc = 0.4，Pm = 0.1，使用資料集train\_5\_D1.txt

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 第幾次  族群大小 | 1 | 2 | 3 | 4 | 5 | 6 |
| 1000  (約花費200秒) | 109.42 | 128.53 | 65.11 | 113.12 | 97.78 | 74.04 |
| 500  (約花費100秒) | 89.86 | 81.95 | 86.29 | 67.19 | 83.93 | 130.87 |
| 300  (約花費56秒) | 158.76 | 111.80 | 129.15 | 86.42 | 109.72 | 145.7367 |
| 200  (約花費40秒) | 117.4065 | 81.79 | 110.94 | 118.00 | 165.1426 | 74.08 |
| 100  (約花費20秒) | 92.94 | 95.26 | 150.39 | 133.06 | 113.85 | 104.25 |

※族群次數約500次數值較為穩定。

**(3)交配機率**

迭代次數 = 1000，族群大小 = 500，Pm = 0.1，使用資料集train\_5\_D1.txt

圖3是使用不同交配機率的收斂結果(橫軸為第幾代，縱軸為該代適應函數最小值)

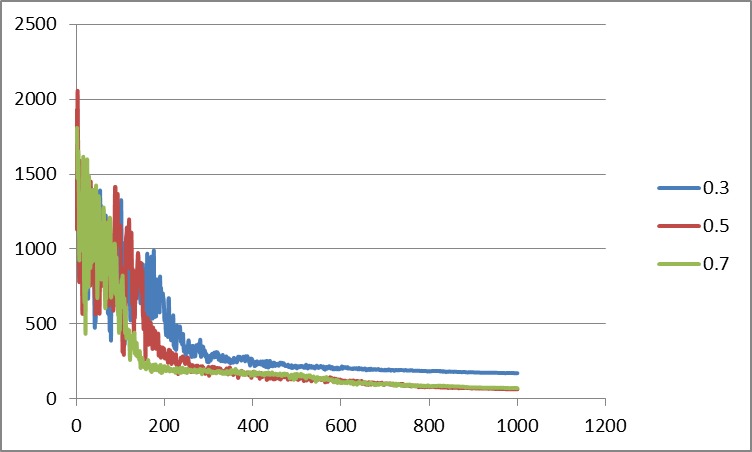


圖 3不同交配機率比較

**(4)突變機率**

迭代次數 = 1000，族群大小 = 500，Pc = 0.5，使用資料集train\_5\_D1.txt

圖4是使用不同突變機率的收斂結果(橫軸為第幾代，縱軸為該代適應函數最小值)

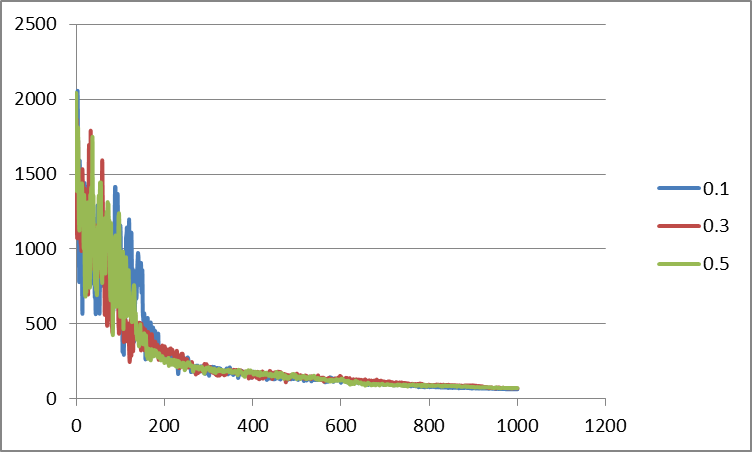


圖4 不同突變機率比較

**5.實驗遇到的問題**

(1)參數收斂慢

複製基因時加入隨機雜訊(炸彈效應)，嘗試最佳解附近的解。或是可提高交配機率，但也因此陷入區域最佳解。

(2)區域最佳解

在演化的每一代加入新的隨機解，或是在剛開始演化時，加入較大雜訊，使基因得以探索其他區域。

(3)參數是否該做邊界判斷?

若參數超出邊界，則適應函數會比其他基因大，所以會在基因演化的過程中慢慢將其淘汰，而不用額外做邊界判斷。

(4)神經元數量

神經元數量越多需要花費更多迭代次數做調整。

(5)測試資料的誤解

一開始沒有發現測試資料中的距離是以圓心開始的距離，所以訓練完後，給入的輸入資料同樣必須在加上車子的半徑，才會是神經網路訓練後的正確輸出值。

**6.心路歷程**

使用過基因演算法後，發現能夠找出某個問題的最佳參數解，但同樣需要設定和調整基因演算法的參數，只是參數的設定上來得更有意義，另外，也發現有別於以往寫程式，而是將大部分的時間用在做實驗和參數調整，感謝助教和同學熱心地解惑。