

黑白配—神奇的遺傳演算法

國中組應用科學科第一名

台北縣立福和國民中學

作者：鄭劭宇、蕭俊宏、陳柏甫、林以唐

指導教師：洪國政、許芳誠

一、研究動機

在六億年前，海洋中有許多不同種類的生物，經過了幾億年的演化和突變，令我們覺得好奇，為何基因經過交配和突變能產生這麼多變化，而且愈來愈優良，於是我們想利用電腦的符號來代表各種不同的基因，使其交配、突變，研究近年來興起的「遺傳演算法」學說。

二、研究目的

- (一) 瞭解遺傳演算法的原理。
- (二) 遺傳演算法的應用。
- (三) 遺傳演算法對人工智慧的意義。

三、研究設備與器材

- (一) 電腦
- (二) Microsoft Visual Basic 5.0
- (三) Paint Shop Pro4.12
- (四) Microsoft Office 97

四、研究問題

- (一) 遺傳演算法的原理是什麼？
- (二) 如何設計「黑白牌」的程式？
- (三) 如何設計「猜數字遊戲」的程式？
- (四) 如何設計「二次函數」的程式？
- (五) 如何應用遺傳演算法？
- (六) 遺傳演算法和人工智慧之間有什麼關係？

五、研究步驟

(一) 遺傳演算法的原理是什麼？

遺傳演算法是最近幾年才興起的一門學說，它是一個具有無窮潛力的演算法。它利用基因交配的特色使許多數據或文字去其糟粕，取其精華，求得正確的答案。

由於大自然的基因演化情況是「物競天擇」、「適者生存」，因此能使品種越便越優良，就像藍綠藻演化成各種生物一般。人工智慧目前最大的困難，即是電腦不能模擬人的思考方法，而只能做出條件判斷的功能。遺傳演算法使電腦自行依實際情況來判斷，模仿人類的思考。我們可用各種符號代表基因，為方便計算起見，以 0 和 1 代表兩種不同的基因。假設有兩串基因，其基因串為 10110100 和 11001001，兩串基因進行交配，如從中間切開，就變成 10111001 和 11000100，如此產生新的基因。但若是一味的讓兩組基因不斷交配，不會有什麼結果，因為大自然內的生物之所以會越來越進步，是因為產生出來壞的基因會遭到環境的淘汰，而在交配前死去，存活下來的，就是比較好的基因，其交配的時候，才會產生更好的品種，再加上突變的因素，於是產生更優良的品種。我們在作基因的演算時，也應告訴電腦那一組基因比較好，電腦再依照「適應值」的大小，來判斷抽出基因好壞的機率，再將抽出的基因加以交配，如此一代一代下去，所演化出的基因會越接近我們所需的答案，到了數代之後，電腦就會做出判斷，告訴人們正確的答案是什麼。遺傳演算法的演算步驟如下所示：

1. 電腦產生一些亂數，讓人們或是電腦給每一個亂數一個適應值，適應值越大，在稍後的基因選擇中，所被抽到的機率就越大。
2. 電腦根據每一組適應值，來決定抽到每一組基因的機率，接著抽出同樣多組的基因，放入交配池內。
3. 在交配池內，因為所被抽出的基因大部份是適應值比較高的基因，所以比第 0 代的基因優良一些，在他們之中，每一組基因隨意切割成幾小段，再和另一組基因交配，以產生新的第一代基因。
4. 新的第一代基因已經比第零代基因優良許多，再重新回到第二個步驟，重新給適應值，再度循環，經過數代之後，找出正確的答案。

(二) 如何設計「黑白牌」的程式？

1. 方法：我們利用遺傳演算法，設計一個能猜出使用者心中所想的牌的顏色及順序的程式。
2. 表單配置：

(1)Form1 (黑白牌黑白配)

黑白牌黑白配

電腦所猜出的牌

■	■	□	■
■	■	□	■
□	□	□	□
■	■	■	□
■	■	■	■
■	■	■	□
□	■	■	■
■	□	■	□

適應值(1-8)

6
2
8
3
11
4
5
7

備忘板

3
1
5
1
0
1
2
3

產生亂數(R)

設定(S)

確定(O)

離開(Q)

說明(H)

(2)Form2 (設定)

設定

交配選項

交配切開枚數：
隨機

交配對象：
 隨機
 1對2, 3對4, 5對6, ... (固定)
 1對2, 3對4, 5對6, ... (但有輪替)

突變選項

突變形態：
 黑變白, 白變黑
 任選一個位置突變

突變基因組個數：
5

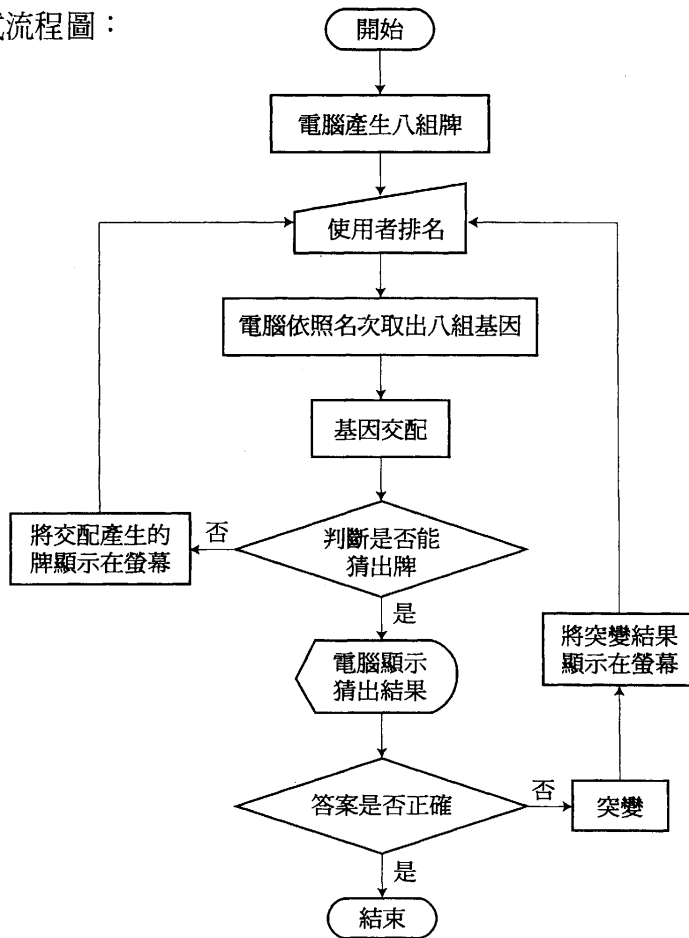
猜出選項

猜出：
 當有四組牌相同時, 就猜出答案
 全部的牌相同時才猜出答案

確定

取消

3. 程式流程圖：



4. 實驗過程：我們以 0 代表白，1 代表黑，使用者先想出五張牌，依照電腦猜出牌的正確率給適應值，再讓八組牌去交配，最後求出答案。我們心中先設定“黑白黑白黑”。

表一：第零代

基因串編號	電腦以亂數決定的牌	與使用者的牌差距數目	適應值	被選入交配池的機率
1	10010	3	4	4/36
2	01101	2	7	7/36
3	00110	3	4	4/36
4	11101	1	8	8/36
5	10110	2	7	7/36
6	01101	2	7	7/36
7	01001	3	4	4/36
8	01011	4	1	1/36

表二：取隨機數

隨機數	1	2~3	4~6	7~10	11~15	16~21	22~28	29~36
對應的基因串編號	8	1,3,7	1,3,7	1,3,7	2,5,6	2,5,6	2,5,6	4

表三：以隨機數取出的基因串

隨機數	17	11	13	28	5	35	33	22
選出的基因串編號	2	5	6	2	3	4	4	6
選出的基因串	10010	10110	01101	01101	00110	11101	11101	01101

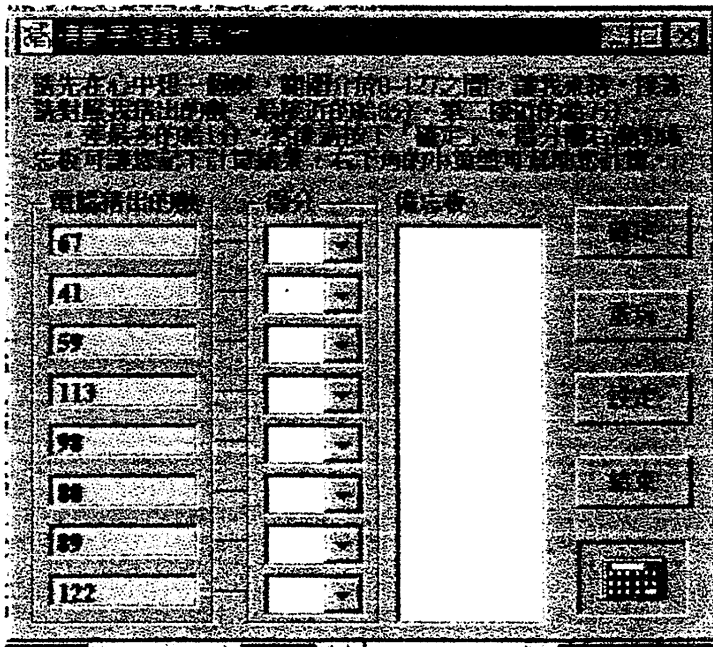
表四：交配池

交配池編號	交配池	交配對象 (隨機產生)	交配對象 (隨機產生)	子代族群
1	10010	4	1,3	11110
2	10110	5	1,3	10110
3	01101	7	1,3	01101
4	01101	1	1,3	00001
5	00110	2	1,3	00110
6	11101	8	1,3	11101
7	11101	3	1,3	11101
8	01101	6	1,3	01101

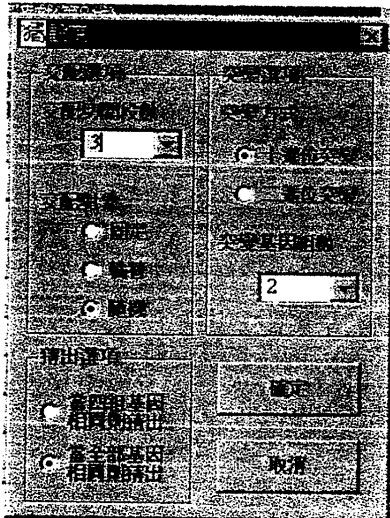
如此循環下去，即可算出正確的牌。經過這個實驗後，我們又改變了交配切開段數、交配對象、突變基因組數……等變因，觀察電腦猜出正確答案需要幾代。

(三) 如何設計「猜數字遊戲」的程式？

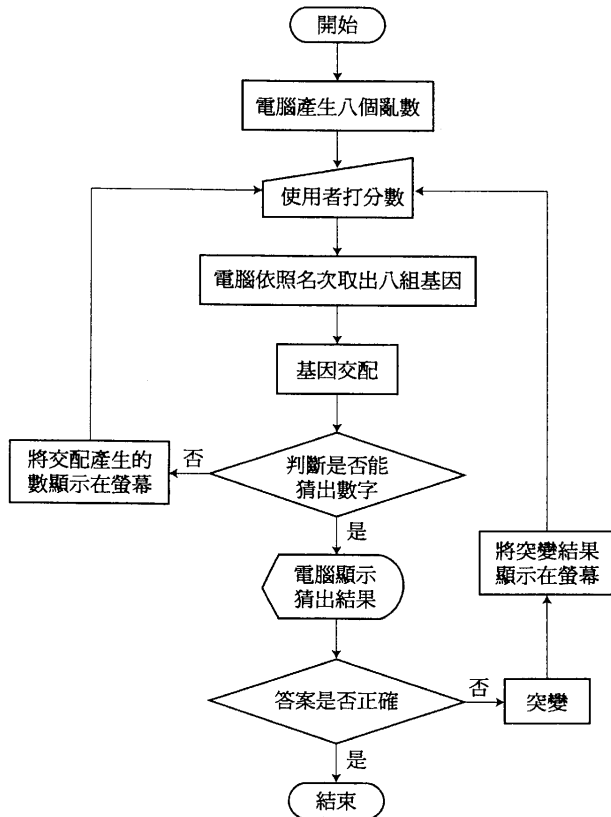
1. 方法：我們將黑白排程式延伸，改成能夠猜出使用者心中所想數字的程式，使用者心中字定一數，範圍為0~127，由二進位碼代表基因。
2. 表單配置：
 - (1)Form1 (使用遺傳演算法的猜數字遊戲)



(2)Form2 (設定)



3.程式流程圖：



4. 實驗過程：我們心中自行設定一數，範圍由零到一百二十七，假設為 100，然後由電腦猜。

表一：第零代

基因串編號	初始族群 (隨機產生)	十進位值	與正確答案的差距 (使用者自行計算)	適應值	被選入交配池的機率
1	0101010	42	58	3	3/36
2	0000000	0	100	1	1/36
3	1111000	120	20	5	5/36
4	1010101	85	15	7	7/36
5	1101101	109	9	8	8/36
6	1111111	127	27	4	4/36
7	1010001	81	19	6	6/36
8	0000100	4	96	2	2/36
合計				36	1(36/36)

表二：取隨機數

隨機數	1	2~3	4~6	7~10	11~15	16~21	22~28	29~36
對應的基因串編號	2	8	1	6	3	7	4	5

電腦會從1~36之間取八個亂數，數目如果在表二的隨機數範圍內，例如7~10，那麼電腦只要抽到7，8，9，10任一數，由於抽這些數的機率4/36，所以選擇適應值是4的6號基因放入交配池內。所以隨機數的範圍大小代表抽到那些數，也就是抽到哪些基因串的機率大小。適應值越高的，其被抽到的機率就越大，生存下來的機會也就越高。我們假設抽到的隨機數如下表：

表三：以隨機數抽出的基因串

隨機數	24	20	30	27	8	32	12	6
選出之基因串編號	1	5	6	1	2	6	7	3
選出之基因串	1000110	0101101	1010110	1000110	1111000	1010110	1110000	0010101

很明顯的，我們發現經過適應值的挑選後，所取出的基因串較優良，比較接近我們心中所想的數。

表四：交配池

交配池編號	交配池	交配對象 (隨機產生)	交配位置 (隨機產生)	子代族群	子代族群 十進位值
1	1000110	2	4	1000101	69
2	0101101	1	4	0101110	46
3	1010110	5	4	1010000	80
4	1000110	6	4	1000110	70
5	1111000	3	4	1111110	126
6	1010110	4	4	1010110	86
7	1110000	8	4	1110101	117
8	0010101	7	4	0010000	16

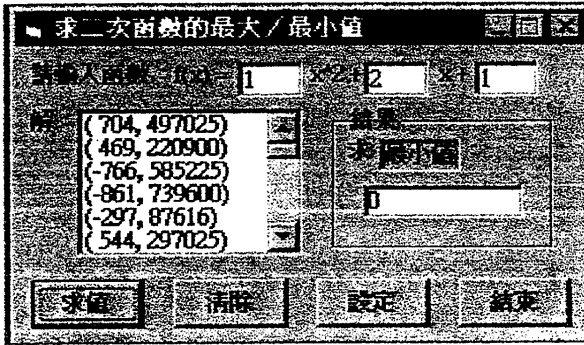
我們發現，將子代基因再度取適應值，然後交配，產生第二代的基因，如此經過數代之後，就可找出我們心理所想的數字。經過這個實驗後，我們又改變了交配切開段數、交配對象、突變基因組數……等變因，觀察電腦猜出正確答案需要幾代。

(四) 如何設計「二次函數」的程式？

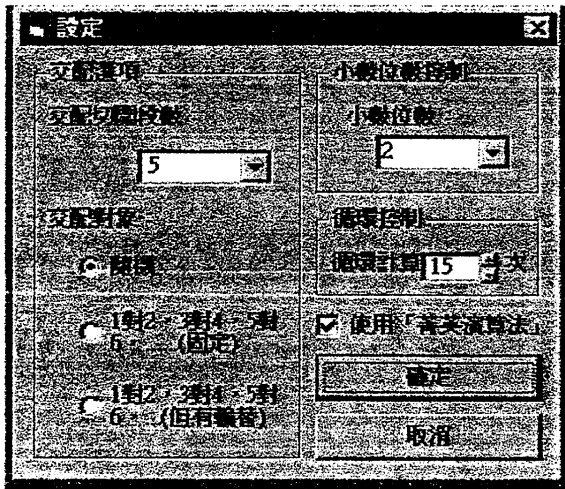
1. 方法：我們把遺傳演算法應用到解數學式子上，來解二次函數的極大、極小值，電腦一開始先用亂數代入 $f(x)$ ，求出各種不同的值。先判斷要求最大值或最小值：若要求最大值，則 $f(x)$ 愈大則適應值愈高，反之則愈低；若要求最小值，則 $f(x)$ 愈小則適應值愈高，反之則愈低。同樣利用二進位碼代表基因，至於小數部分，則將小數點之前存成一個變數，小數點之後存成另一個變數，在轉換成二進位時，將這兩個變數分別轉換。因為我們並不知道如何將小數直接以二進位表示。轉換後的二進位碼再相結合，中間以小數點分隔。例如：8.3，我們將此數分為8和3，分別轉成二進位為1000、11，再結合成1000.11。

2. 表單配置

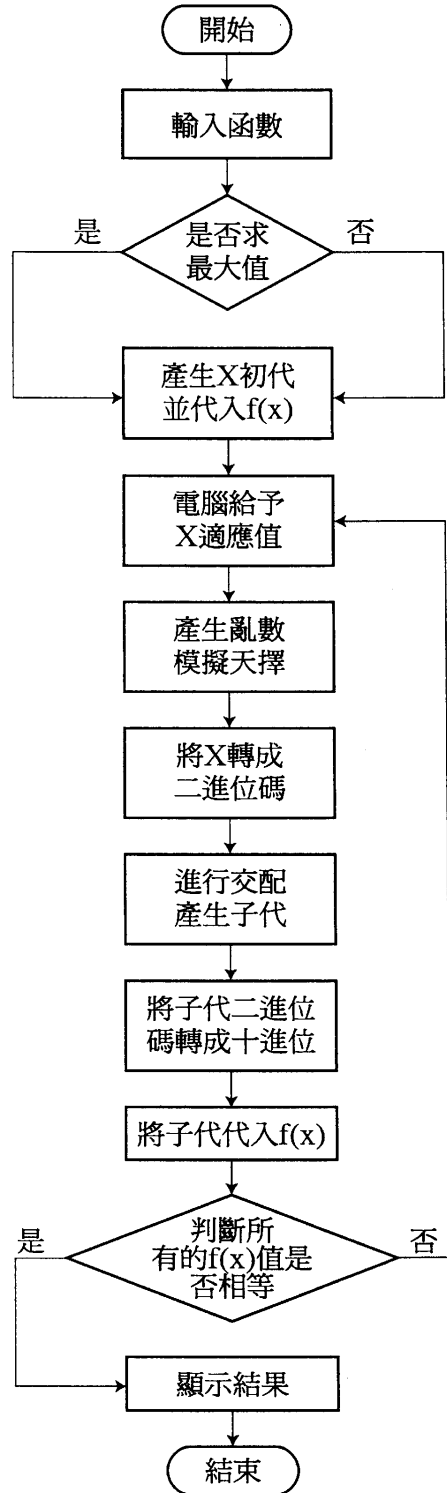
(1)Form1 (求二次函數的最大/最小值)



(2)Form2 (設定)



3.程式流程圖:



4. 實驗過程：

表一：第 0 代（我們以 $f(x)=x^2+2x+3$ 為例）

X基因編號	第0代（隨機）	十進位值	對應的 $f(x)$ 值
1	00101101	45	2118
2	-00100101	-37	1298
3	-00000101	-5	18
4	-00111010	-58	3251
5	-00001111	-15	198
6	00011001	25	678
7	-01011010	-90	7923
8	00011100	28	843

我們為了使 x 帶有正負號，因此多設了一個變數，以亂數決定 x 為正數、負數或零。之後，我們再代入函數中，求得 y 的值。在這裡為了方便起見，我們在二進位碼面前以“-”來代表負數。

表二：取適應值

X基因編號	第0代	十進位值	對應的 $f(x)$ 值	適應值	被選入交配池機率
7	-01011010	-90	7923	1	1/36
4	-00111010	-58	3251	2	2/36
1	00101101	45	2118	3	3/36
2	-00100101	-37	1298	4	4/36
8	00011100	28	843	5	5/36
6	00011001	25	678	6	6/36
5	-00001111	-15	198	7	7/36
3	-00000101	-5	18	8	8/36

在求函數值中，我們再輸入函數時已判斷求最小值或最大值者（我們以 a 值來判斷），其實可以讓電腦自己取適應值，實在沒有必要讓使用者一一輸入，浪費時間。所以我們就想出一種讓電腦可以自行判斷的方法。即是「排序」。我們採用「氣泡排序法」，依照使用者輸入的函數，由最差勁的數到最好的數，依次排進陣列“1”、陣列“2”、…、陣列“8”之中，則陣列的序號即為適應值。這樣，就可省下大半的輸入時間了。

之後，我們再按照前面所講的方法：取隨機數、交配、產生子代…，即可求出所要的答案。

從這個程式中，我們發現遺傳演算法只要給予足夠的條件，我們便可以讓電

腦程式自動做出正確的判斷，為每個基因打上適當的分數（適應值），不但免去一一輸入的麻煩，也節省了不少的時間，使它的功能更趨強大。

另外，我們在這個程式中，加入了「菁英演算法」。菁英演算法，是電腦以亂數來挑選基因時，若抽到適應值為1，2，3，4的基因組，就淘汰，另外以亂數產生一組新的基因來代替。菁英演算法雖然在執行的速度上較慢，但卻較準確。

六、研究結果

(一) 猜數字遊戲：

我們改變不同的變因，看看哪一種能使猜出的速度最快。

交配代數 變因 \ 段數	2	3	4	5	6	7	
1-2, 3-4, 5-6, 7-8	28 32 24 27 21 15 17 13 19 12 22 19 18 25 23 22 27 33 21 24 15 19 21 27 22 21 20 14 17 24	26	15	21	25	21	19
1-2, 3-4, 5-6, 7-8 (輪替)	25 21 28 17 26 13 14 11 9 10 21 17 11 13 9 18 23 6 24 13 18 13 11 26 19 19 24 13 10 21	23	12	16	17	17	17
以亂數 決定對象	21 27 19 25 23 11 3 5 13 7 15 16 17 14 12 21 17 16 19 13 11 14 21 24 17 11 18 7 20 13	23	8	15	19	17	13

每五個標位下面的大格子為平均值。我們發現在切三段時，猜出的速度是最快的。

(二) 黑白牌

交配代數 變因 \ 段數	2	3	4	5	
1-2, 3-4, 5-6, 7-8	6 5 7 9 8 9 7 10 5 6 17 15 13 18 15 25 31 27 36 21	7	7	16	28
1-2, 3-4, 5-6, 7-8 (輪替)	5 6 5 4 6 5 7 6 5 6 8 9 11 10 13 17 16 20 14 19	5	6	10	17
以亂數 決定對象	3 5 4 3 3 4 6 3 4 5 6 7 9 11 7 11 8 8 10 7	4	4	8	9

每五個欄位下面的大格子為平均值。我們發現在切二段時，猜出的速度是最快的。

(三) 求二次函數極大、極小值：

(右邊的欄位：菁英演算法)

交配對象 交配代數 段數	2		3		4		5		6		7		8		9		10	
	隨機	100	17	66	13	45	12	40	11	50	12	71	14	79	15	84	17	95
固定	100	31	83	25	67	18	52	15	74	22	82	24	87	28	98	29	100	30
循環	100	27	76	21	59	15	47	13	64	14	78	19	83	23	91	26	97	28

發現菁英演算法能普遍提高演算精確度，但電腦執行的速度會減低。在切開段數為五段時，猜出的速度最快。

七、討論

(一) 如何應用遺傳演算法？

1. 利用遺傳演算法推導戰爭策略，配合地形、氣候、軍事武器、人員調配和進軍路線，評估戰略價值，研究推導出最有利的軍事策略。
2. 利用遺傳演算法分析案情，比對犯人的外型、特徵，增加破案機會。(例如劉邦友血案、彭婉如命案。)
3. 利用遺傳演算法增進網路查詢功能，使上網瀏覽更為快速便捷。
4. 利用遺傳演算法推算車輛的行駛路徑，改善日益嚴重的交通問題。(例如台北市區嚴重的交通堵塞和停車問題。)
5. 利用遺傳演算法從事室內設計，設計出使用者心目中理想的室內房屋的裝潢和配置。
6. 利用遺傳演算法，測量地底下的瓦斯濃度，迅速找出瓦斯管漏氣的地點。
7. 利用遺傳演算法，計算出最優良的房屋架構和建築材料。
8. 利用遺傳演算法推導出數學公式，解出複雜的數學問題。
9. 利用遺傳演算法配合人的審美感從事藝術創作，創作出更新奇的藝術作品。
10. 利用遺傳演算法解密碼，如軍事用途。
11. 利用遺傳演算法推算太空站應設置位置。

12.利用遺傳演算法發展人工智慧程式。

(二) 遺傳演算法和人工智慧有什麼關係？

經過以上數個實驗與討論，我們發現幾乎是每一項人們需要考慮各種情況的問題，都可以用遺傳演算法加以解決。

事實上，遺傳演算法是很神奇的，因為電腦並不知道該怎麼解出答案，但是我們在程式裡加入最奇妙的大自然基因的淘汰方法，所以能有驚人的表現，遺傳演算法再解決相當困難的問題時，花了很短的時間，更是凌駕人的思考以上。例如一個很複雜的數學函數，連用微積分也無法解出，遺傳演算法卻稍加嘗試，在短短的幾秒鐘或是幾分鐘即求出答案！

假如我們把遺傳演算法加深，使它有更多的基因，更像真正生物的基因構造，一定可以解決更複雜的問題。在問題六裡，我們探討過遺傳演算法的應用，有許多項目已經有人開始研究了，而且都非常成功。

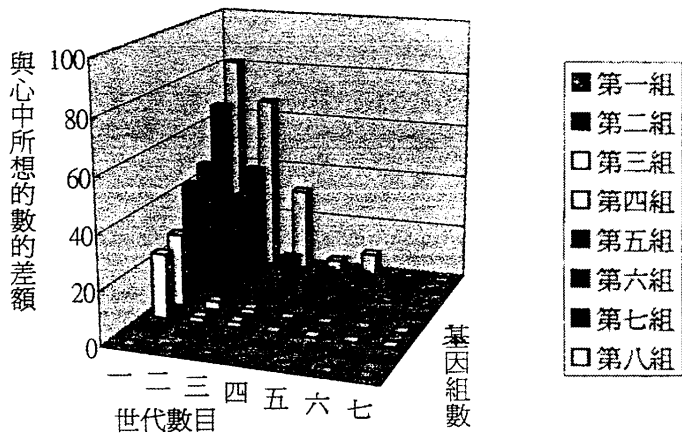
在遺傳演算法中，交配繁殖能夠讓數位染色體重組遺傳基因，整個族群會不斷產生新的基本單位，或是新的組合。適當的突變也能讓基因有更多的可能性。

遺傳演算法目前是人工智慧裡的一個分枝，它目前已經解決了許多人解答不出的函數，還有更多人正在努力應用它，以期它有一天，能夠真的模擬人類的思考。我們會再深入研究及應用，使電腦的應用領域更深一層擴展。

八、結論

(一) 遺傳演算法的演算步驟：

- 1.電腦以零和一模擬產生一些基因組，讓電腦或人們給每一組基因一個對環境的適應值。
- 2.電腦根據每一組適應值，來決定每一組基因的生存機率，接著抽出同樣多組的基因，放入交配池。
- 3.交配池內，所抽出優良的基因中，每一組隨意切成幾小段，再和另一組基因交配，以產生新一代基因。



▲ 數值的“演化”過程（以猜數字遊戲為例）

- 4.回到第二個步驟，重新給適應值，經過數代之後，找出正確的答案。
- (二) 在實驗過程中，使用菁英演算法在數學函數的計算上有較準確的結果，但速度較慢。
- (三) 在實驗過程中，我們有重大發現：若基因的長度為 n 個字元，則當 n 為奇數時，以切開 $(n-1)/2$ 段時能夠最快求出解來，當 n 個為偶數時，以切開 $n/2$ 段時能夠最快求出解來！
- (四) 我們以數值的二進位碼作為該數的基因組，是因為電腦在計算時，都是以二進位來計算。其實我們可以加以延伸，以更多符號來代表基因，例如： a 、 b 、 c …。甚至可以創造出一套編碼的方法，以增強遺傳演算法的功能。
- (五) 遺傳演算法不但能猜出人們的心思，同時也可以用來求二次函數的極大極小值、一元 n 次方程式等等。只要有足夠的條件，我們可以用排序的方式，將適應值由最差的數道最好的數，依序排進陣列中，使陣列的序數代表適應值，使電腦自動產生適應值。減少輸入時的麻煩和可能發生的錯誤，並增快運算的速度。

九、參考資料

(一)「複雜」

作者：渥德羅普

譯者：齊若蘭

出版：天下文化

(二) "Scientific American" — July 1992

作者：John H. Holland

(三)「Microsoft Visual Basic 5 視窗程式設計經典 實務篇」

作者：林永森

出版：碁峰資訊

評語

- 1.初中生能對遺傳演算法做完整的了解，設計出正確的程式，並且應用在2次方程式根的解法等俱體的實例，其數學理解能力，程式製作能力已俱有大學數學，資訊等本科系高年級之水準。
- 2.利用所設計之程式做各種實驗，觀察收斂的數據，導出一般通則，進而做出結論，此項心智活動過程顯示已能掌握科學方法的基本精神。