

---

# Genetic Algorithm

**Jong Yih Kuo**

**[jykuo@ntut.edu.tw](mailto:jykuo@ntut.edu.tw)**

**Department of Computer Science and  
Information Engineering**

**National Taipei University of Technology**

# Introduction

## □ 最佳化問題的解法

- 特定型：目標函數 (object function) 滿足某些特性，如線性、非時變、可微分等。微分法及梯度法屬此類。
- 廣義型：適用各種目標函數。隨機搜尋及基因演算法屬此。

## □ 基因演算法由 Holland 於 1975 年提出，基於自然選擇過程的一種最佳化搜尋法。

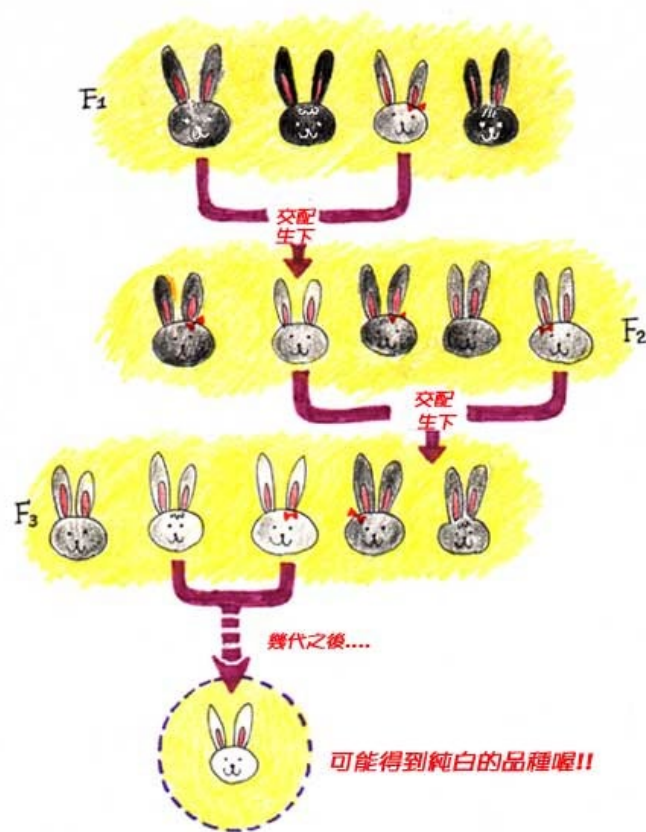
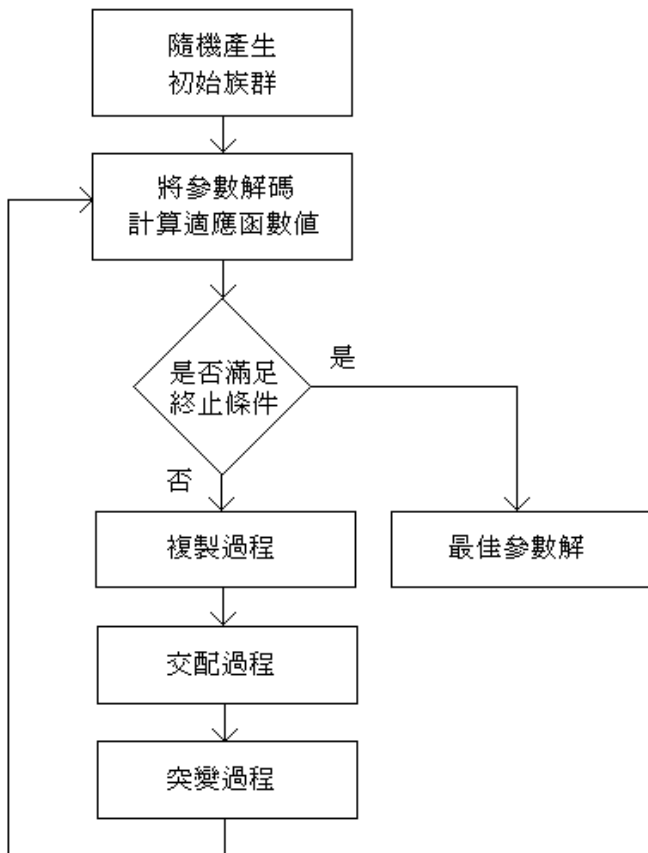
- 其搜尋速度比隨機搜尋有效率；
- 特定型最佳化法則效率高但僅能處理特定系統；
- 廣義型法則應用範圍廣但效率低。



○ 現有的資料支持這個學說

# Introduction

## □ 流程



# 編碼(Coding)

## □ 二進位編碼

基因	1			2			3			4			5			6		
	0	1	1	0	1	1	1	0	1	0	0	1	1	1	0	0	1	0

## □ 實數編碼

基因	1	2	3	4	5	6
	86.3	115.0	10.7	121.9	1.6	64.2

## □ 符號編碼

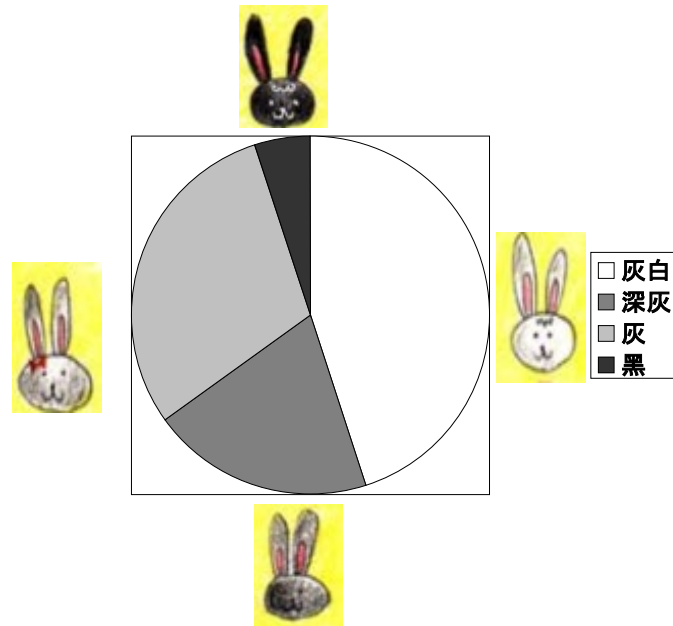
基因	1	2	3	4	5	6
	A	F	D	E	C	B

# 複製 (Reproduction)

- ❑ 依每一物種適應度決定下一子代中應被淘汰或複製的個數，其中適應程度由適應函數(fitness function)量測。
- ❑ 適應性函數
  - 決定個體的適應度(Fitness)
- ❑ 隨機取樣 (Stochastic sampling)
  - 隨機選取兩個或更多個物種作為親代。
- ❑ 明確取樣(Deterministic sampling)
  - 直接挑選適應度最高的物種。
- ❑ 混合取樣 (Mixed sampling)
  - 競爭式選擇、輪盤式選擇

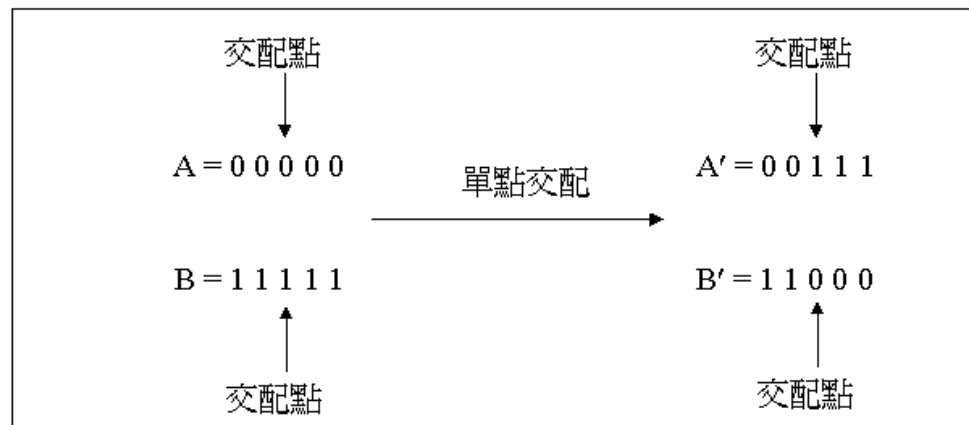
# 輪盤式選擇(Roulette Wheel)

- 適應函數值越大的物種在輪盤上面積越大，選中機率越大
- 隨機選取輪盤上的一點，即為親代。



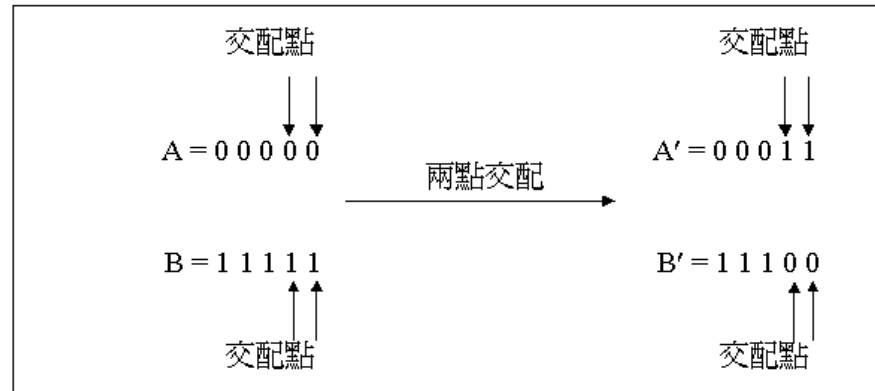
# 交配 (Crossover) 1

- 隨機選取兩個親代物種，彼此交換位元資訊，組成另外兩個新子代物種。
  - 發生機率由交配機率控制。
- 單點交配：選出兩字串中，隨機選取一交配點，交換兩字串中此交配點後所有位元。

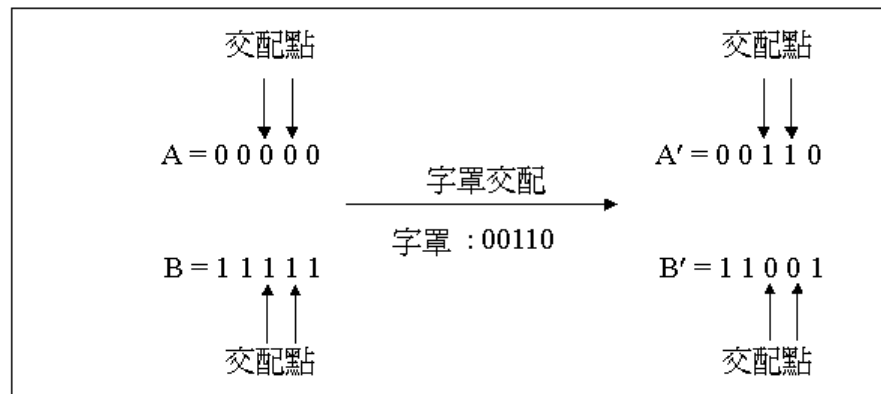


# 交配 (Crossover) 2

- 兩點交配：隨機選取兩交配點，交換交配點間所有位元。



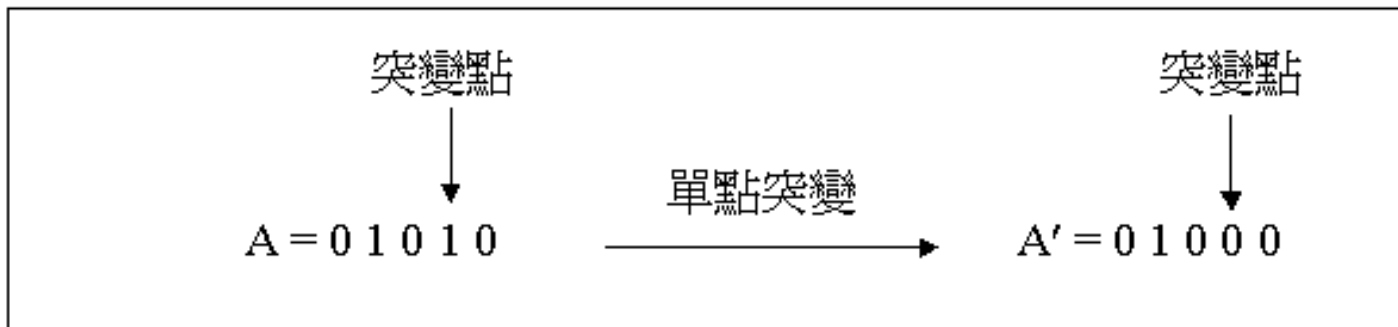
- 字罩交配：產生與物種字串長度相同的字罩，字罩隨機由 0 與 1 組成，1 是兩物種字串交換位元資訊的位置。





# 突變 (Mutation)

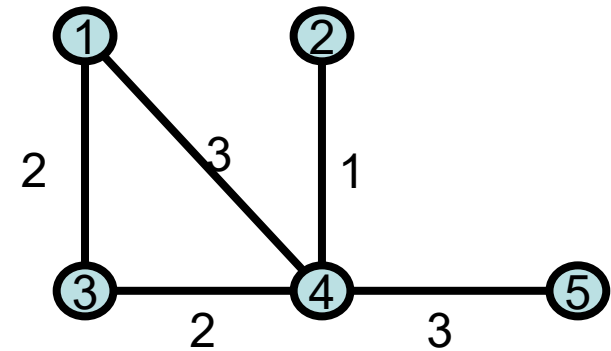
- ❑ 隨機選取一物種字串，隨機選取突變點，改變物種字串位元資訊。
- ❑ 發生機率由突變機率控制。
- ❑ 突變過程可針對單一位元、或整個字串進行突變運算、或以字罩突變方式為之。
- ❑ 對二進位字串，是將字串 0 變 1，1 變 0。



# 路徑規劃（推銷員旅行問題）

## □ 任兩點間的路徑成本

	1	2	3	4	5
1		4	2	3	6
2	4		3	1	4
3	2	3		2	5
4	3	1	2		3
5	6	4	5	3	



## □ 基因編碼

- 路徑排列組合數目有  $4! = 24$  種可能路徑。
- $24 \leq 32 = 2^5$
- 用二進位編碼，染色體上需要5個基因。
- 如果出現11111或00000....，視為無效編碼丟棄

# 基因編碼

路徑	編碼	路徑	編碼	路徑	編碼
12345	00001	13425	01010	14523	10010
12354	00010	13452	01011	14532	10011
12435	00011	13524	01100	15234	10100
12453	00100	13542	01101	15243	10101
12534	00101	14235	01110	15324	10110
12543	00110	14253	01111	15342	10111
13245	00111	14325	10000	15432	11001
13254	01001	14352	10001	15423	11010

# 適應函數值(Fitness)

- ❑ 決定群組數量N，令 $N = 4$
- ❑ 亂數產生個體(必須為有效編碼)，組成第0代群組
- ❑ 適應函數為計算路徑起點到終點的路徑成本總合。

路徑	編碼	路徑成本
12435	00011	$4+1+2+5=12$
12534	00101	$4+4+5+2=15$
14523	10010	$3+3+4+3=13$
15432	11001	$6+3+2+3=14$

# 複製 (Reproduction)

## □ 此處採輪盤式取樣

路徑	編碼	路徑成本	取樣機率
12435	00011	$4+1+2+5=12$	0.28
12534	00101	$4+4+5+2=15$	0.22
14523	10010	$3+3+4+3=13$	0.26
15432	11001	$6+3+2+3=14$	0.24

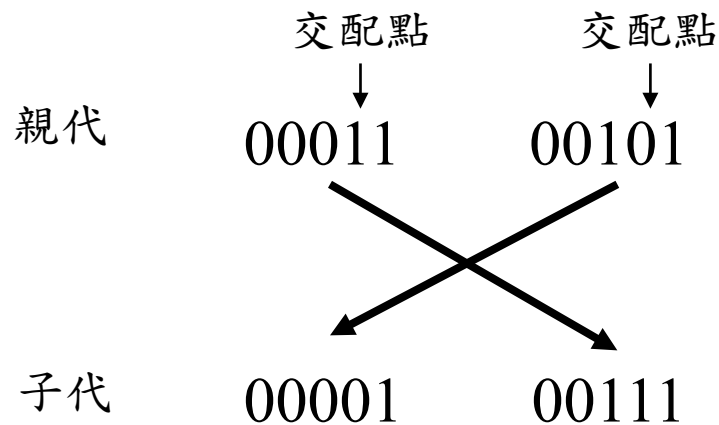
## □ 例：路徑12435，取樣機率：

$$\begin{aligned} & \left( (12+15+13+14) / 12 \right) / \left( (12+15+13+14) / 12 + (12+15+13+14) / 15 + (12+15+13+14) / 13 + (12+15+13+14) / 14 \right) \\ & = 0.279312462 \end{aligned}$$

## □ 依照輪盤式取樣選取配對。假設選中的配對為(00011, 00101)

# 交配 (Crossover)

- 依據交配率決定是否交配。
  - 假設交配率為0.5，則有一半的機率交配發生
  - 此處採用單點交配。
  - 先隨機選取交配點。然後交配產生子代。
  - 如果交配沒有發生，則直接將親代當成下一子代。



# 突變 (Mutation)

❑ 突變機率為0.05，平均每20個基因就有一個突變

○ 此處假設沒有發生突變。

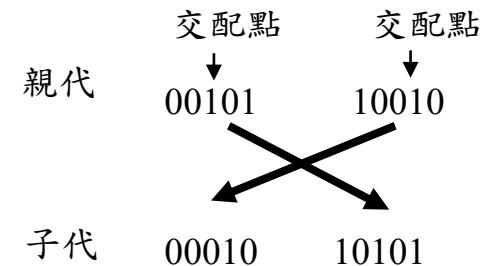
○ 以第1代的族群為：（ 00001，00111 ）

❑ 若未達族群數量，繼續選取下一組。

❑ 依照輪盤式取樣選取配對。

○ 假設這次選中的配對為(00101, 10010)

❑ 依據交配律採用單點交配。



❑ 突變機率為0.05，這次選中子代00010的第三個位元突變，成為00110

❑ 所以第1代的族群為：（ 00001，00111，00110，10101 ）

❑ 達到族群數量。

# 產生第1代族群

編碼	路徑	路徑成本
00001	12345	$4+3+2+3=12$
00111	13245	$2+3+1+3=9$
00110	12543	$4+4+3+2=13$
10101	15243	$6+4+1+2=13$

□ 繼續繁衍下一代直到停止條件。

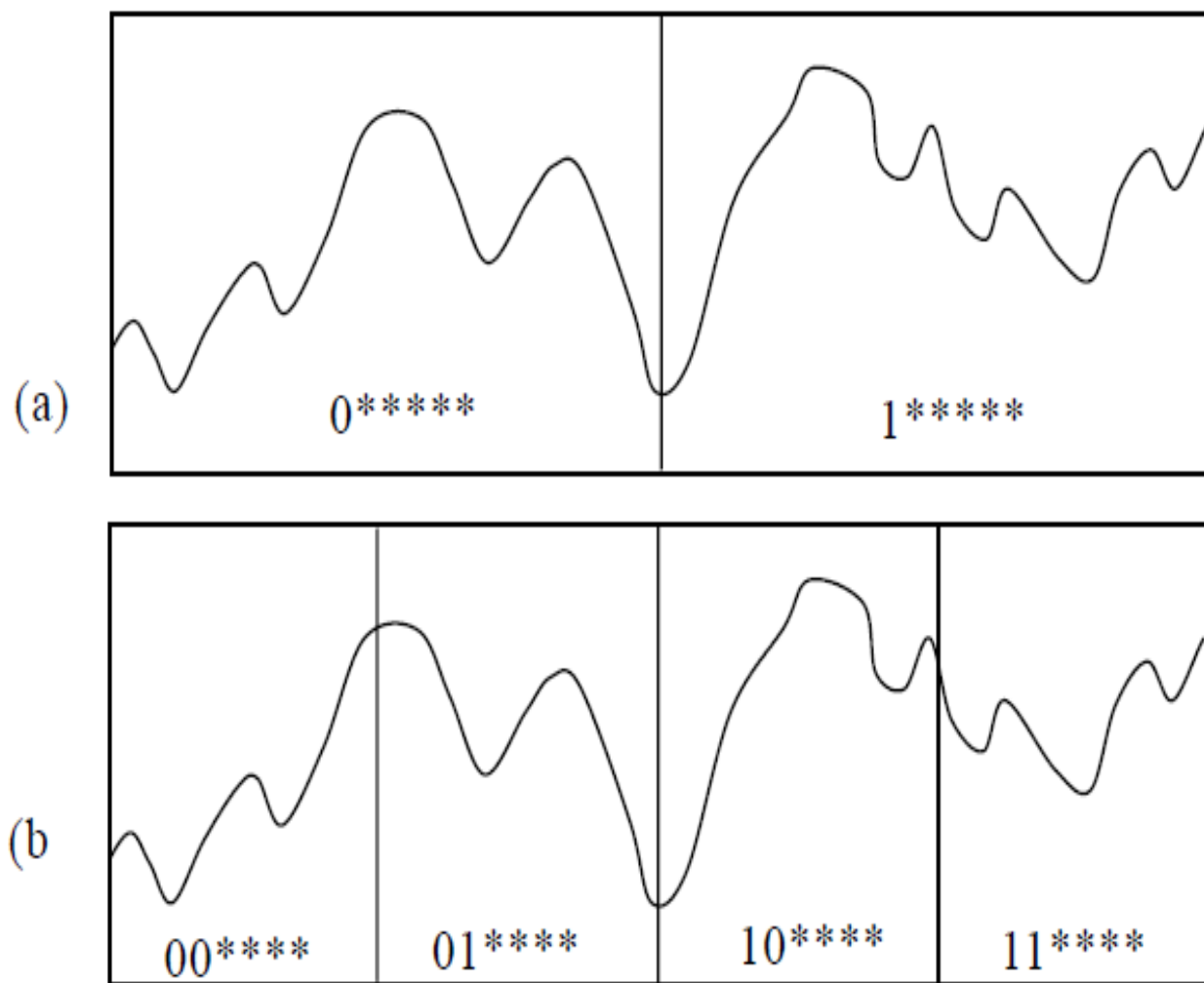


# 基因演算法特性

- 同時考慮搜尋空間上多個點而不是單一點，可較快獲得整體最佳解 (global optimum)，也可避免陷入區域最佳值 (local optimum) 的機會。
- 只使用適應函數資訊，可使用各種型態的適應函數，節省計算資源。
- 使用機率規則引導搜尋方向，而不是明確規則，較能符合各種不同類型的最佳化問題。

# 基因演算法探討

- ❑ 字串長度：長度越長精準度越高，但所須編碼、解碼運算相對增加。
- ❑ 交配機率：交配率越高，新物種進入族群速度越快，搜尋最佳值的速度也越快。
- ❑ 突變機率：突變是必須的運算，因在複製及交配過程中可能使整個族群，所有字串中的某一特定位元皆一樣。
- ❑ 避免陷入區域最佳值：視搜尋空間維度及參數範圍大小與編碼時採用的精確度(字串長度)一起考量。
- ❑ 適應函數設計原則：須能反應不同物種間適應程度差異。
- ❑ 搜尋終止之條件：對某些線上即時系統；為結省時間，當適應函數值到達系統要求後即可終止搜尋。



# Sigma截取 and 乘冪調整

- ❑ Sigma截取：Forrest 建議使用族群適應函數值的變異數資訊，在做適應函數調整前先前處理。作法是將調整前的適應函數值依下列式子減去一常數值：

$$f' = f - (\bar{f} - c \cdot \sigma)$$

- ❑ 參數  $c$  的選擇是族群適應函數值的變異數的倍數。Sigma截取可避免調整過後的適應函數值產生負值。
- ❑ 乘冪調整：Gillies 建議使用乘冪方式調整適應函數值，使調整後適應函數值是調整前的乘冪，如下所示：

$$f' = f^k$$