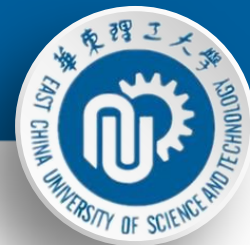


课时内容

第5章 智能搜索技术

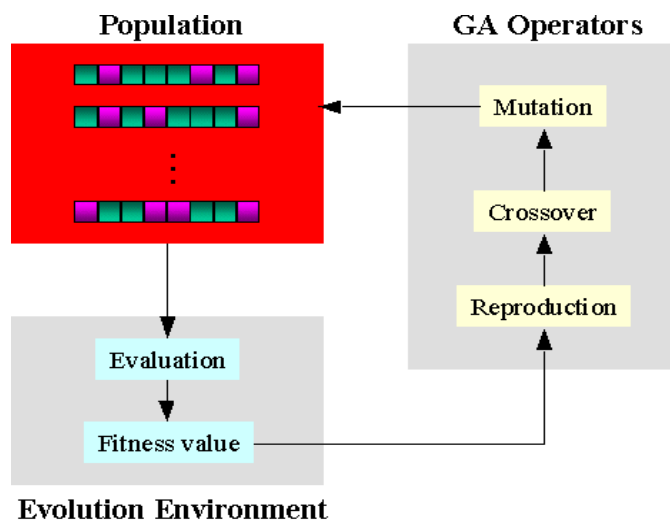




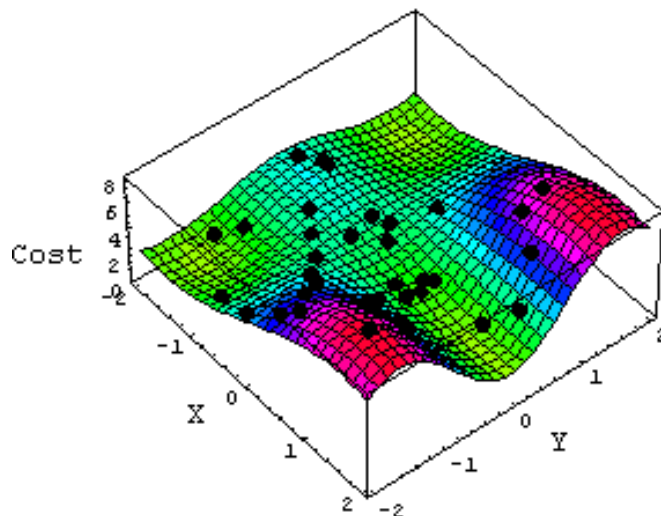
◆ 遗传算法的思路与特点

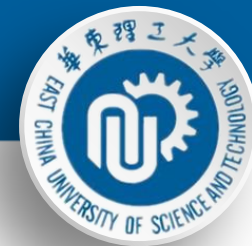
遗传算法的一般步骤:

1. 随机产生种群。
2. 根据策略判断个体的适应度，是否符合优化准则，若符合，输出最佳个体及其最优解，结束。否则，进行下一步。
3. 依据适应度选择父母，适应度高的个体被选中的概率高，适应度低的个体被淘汰。
4. 用父母的染色体按照一定的方法进行交叉，生成子代。
5. 对子代染色体进行变异。
6. 由交叉和变异产生新一代种群，返回步骤2，直到最优解产生



Genetic Algorithm Evolution Flow





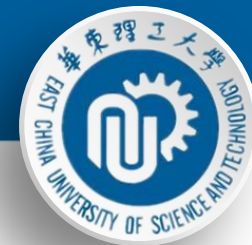
◆ 遗传算法的基本操作

1. 产生初始种群

0001100000	0101111001	0000000101	1001110100	1010101010
(8)	(5)	(2)	(10)	(7)
1110010110	1001011011	1100000001	1001110100	0001010011
(12)	(5)	(19)	(10)	(14)

2. 计算适应度

适应度值



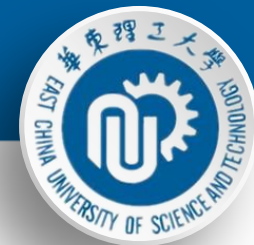
◆ 遗传算法的基本操作

■ 简单实例

3、选择

个体	染色体	适应度	选择概率	累积概率
1	0001100000	8	0.086957	
2	0101111001	5	0.054348	
3	0000000101	2	0.021739	
4	1001110100	10	0.108696	
5	1010101010	7	0.076087	
6	1110010110	12	0.130435	
7	1001011011	5	0.054348	
8	1100000001	19	0.206522	
9	1001110100	10	0.108696	
10	0001010011	14	0.152174	

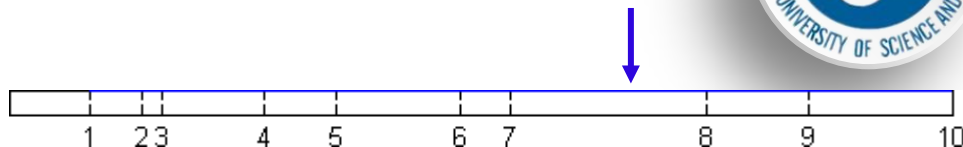
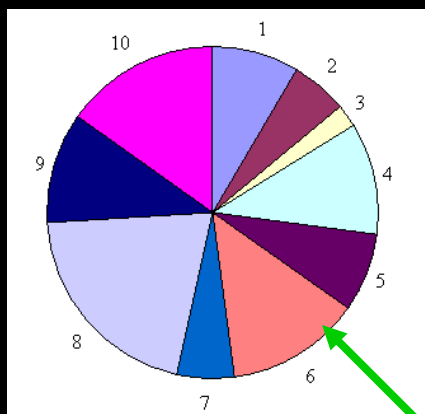
$$8/(8+5+2+10+7+12+5+19+10+14)=8/92=0.086957$$



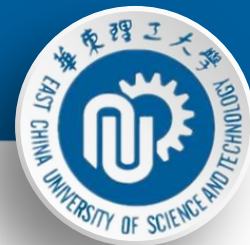
◆ 遗传算法的基本操作

■ 简单实例

3、选择



个体	染色体	适应度	选择概率	累积概率
1	0001100000	8	0.086957	0.086957
2	0101111001	5	0.054348	0.141304
3	0000000101	2	0.021739	0.163043
4	1001110100	10	0.108696	0.271739
5	1010101010	7	0.076087	0.347826
6	1110010110	12	0.130435	0.478261
7	1001011011	5	0.054348	0.532609
8	1100000001	19	0.206522	0.739130
9	1001110100	10	0.108696	0.847826
10	0001010011	14	0.152174	1.000000



◆ 遗传算法的基本操作

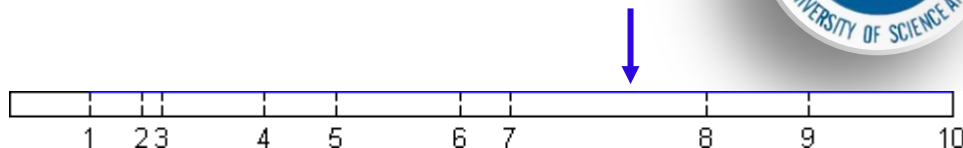
■ 简单实例

3、选择

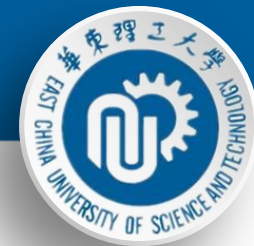
在0~1之间产生一个随机数:

0.070221
0.545929
0.784567
0.446930
0.507893
0.291198
0.716340
0.270901
0.371435
0.854641

淘汰!



个体	染色体	适应度	选择概率	累积概率
1	0001100000	8	0.086957	0.086957
2	0101111001	5	0.054348	0.141304
3	0000000101	2	0.021739	0.163043
4	1001110100	10	0.108696	0.271739
5	1010101010	7	0.076087	0.347826
6	1110010110	12	0.130435	0.478261
7	1001011011	5	0.054348	0.532609
8	1100000001	19	0.206522	0.739130
9	1001110100	10	0.108696	0.847826
10	0001010011	14	0.152174	1.000000



◆ 遗传算法的基本操作

■ 简单实例

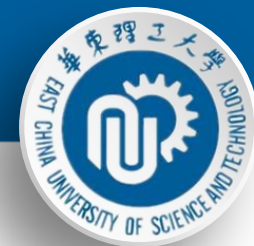
4、交叉

0001	100000	111	0010110	110000	0001	1001110100	1010101010
1110	010110	100	1011011	100111	0100	1100000001	0001010011



0001	010110	111	1011011	110000	0100	1001110100	1010101011
1110	100000	100	0010110	100111	0001	1100000001	0001010010

随机把其中几个位于同一位置的编码进行交换，产生新的个体

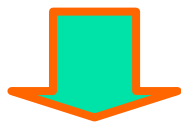


◆ 遗传算法的基本操作

■ 简单实例

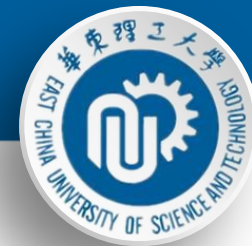
5、变异

0001010110	1111011011	110000 0100	1001110100	1010101 011
1110100000	1000010110	1001 <u>1</u> 0001	1100000001	0001010 010



0001010110	1111011011	110000 0100	1001110100	1010101 011
1110100000	100 0010110	1001 <u>0</u> 0001	1100000001	0001010 010

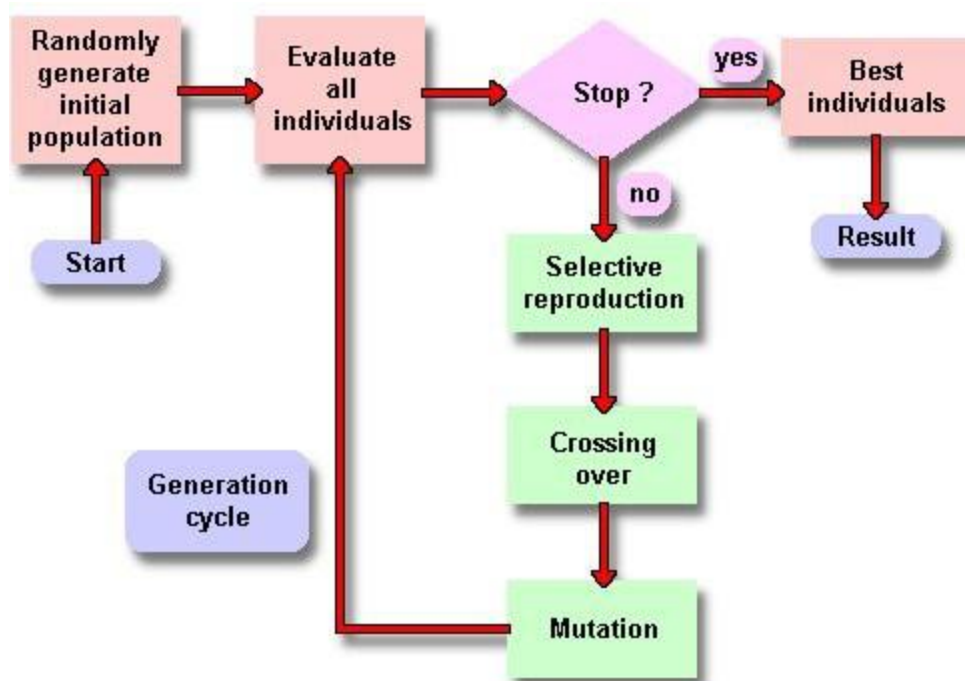
将个体染色体编码串中的某些基因座上的基因值用该基因座上的**其它等位基因**来替换，从而形成新的个体

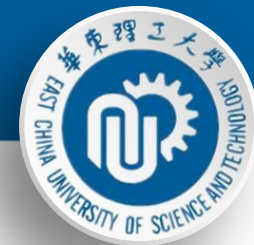


◆ 遗传算法的基本操作

■ 简单实例

6、至下一代，适应度计算→选择→交叉→变异，直至满足终止条件。



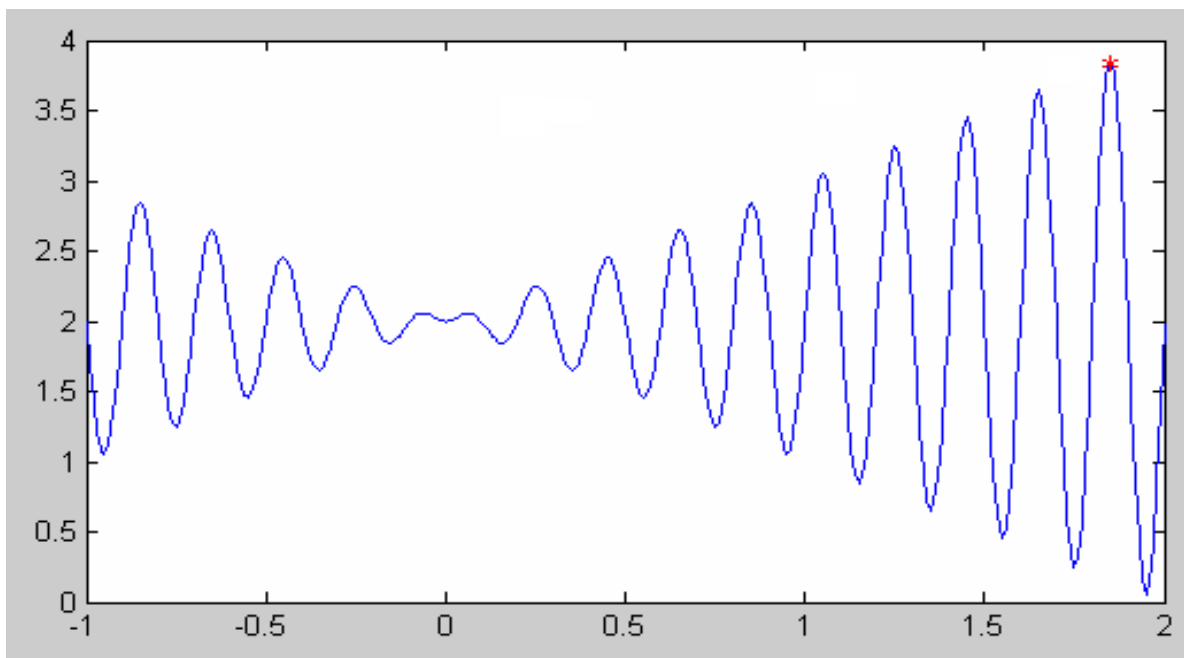


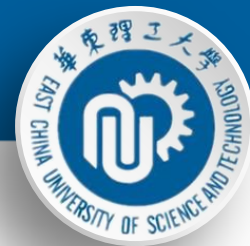
◆ 简单函数优化实例

■ 问题的提出

一元函数求最大值:

$$f(x) = x \sin(10\pi * x) + 2.0 \quad x \in [-1, 2]$$





◆ 简单函数优化实例

■ 问题的提出

用微分法求取 $f(x)$ 的极值点:

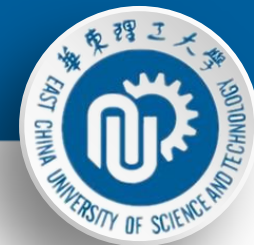
$$f'(x) = \sin(10\pi * x) + 10\pi * x * \cos(10\pi * x) = 0$$

$$\text{即 } \tan(10\pi * x) = -10\pi * x$$

多个解:

$$\begin{cases} x_i = \frac{2i-1}{20} + \varepsilon_i, i = 1, 2, \dots \\ x_0 = 0 \\ x_i = \frac{2i+1}{20} + \varepsilon_i, i = -1, -2, \dots \end{cases}$$

ε_i ($i = 1, 2, \dots$ 及 $i = -1, -2, \dots$) 是一接近于0的实数递减序列。



◆ 简单函数优化实例

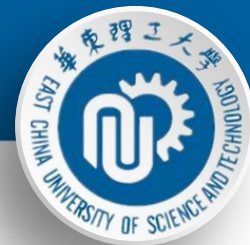
■ 问题的提出

当 i 为奇数时 x_i 对应局部极大值点, i 为偶数时 x_i 对应局部极小值。

x_{19} 即为区间 $[-1,2]$ 内的最大值点:

$$x_{19} = \frac{37}{20} + \varepsilon_{19} = 1.85 + \varepsilon_{19}$$

此时, 函数最大值 $f(x_{19})$ 比 $f(1.85)=3.85$ 稍大。



◆ 简单函数优化实例

■ 编码

表现型: x

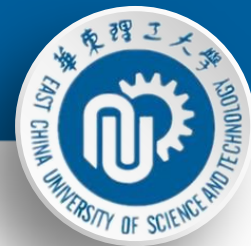
基因型: 二进制编码 (串长取决于求解精度)

串长与精度之间的关系:

若要求求解精度到6位小数, 区间长度为 $2 - (-1) = 3$, 即需将区间分为
 $3/0.000001 = 3 \times 10^6$ 等份。

$$2097152 = 2^{21} < 3000000 < 2^{22} = 4194304$$

所以编码的二进制串长应为22位。



◆ 简单函数优化实例

- 产生初始种群

产生的方式： 随机

产生的结果： 长度为22的二进制串

产生的数量： 种群的大小（规模）， 如30, 50, ...

11110100111100001011000

1100110011101010101110

1010100011110010000100

1011110010011100111001

0001100101001100000011

0000011010010000000000

.....



◆ 简单函数优化实例

■ 计算适应度

不同的问题有不同的适应度计算方法

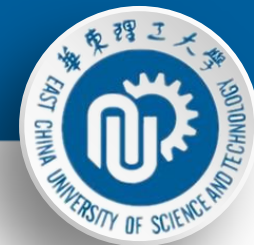
本例：直接用目标函数作为适应度函数

①将某个体转化为 $[-1,2]$ 区间的实数：

$$s=<1000101110110101000111> \rightarrow x=0.637197$$

②计算 x 的函数值（适应度）：

$$f(x)=x\sin(10\pi x)+2.0=2.586345$$



◆ 简单函数优化实例

■ 计算适应度

二进制与十进制之间的转换：

第一步，将一个二进制串 $(b_{21}b_{20}\dots b_0)$ 转化为10进制数：

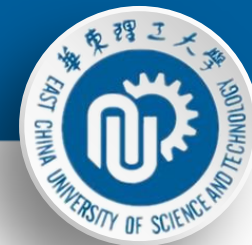
$$(b_{21}b_{20}\dots b_0)_2 = \left(\sum_{i=0}^{21} b_i \cdot 2^i\right)_{10} = x'$$

第二步， x' 对应的区间 $[-1,2]$ 内的实数：

$$x = -1.0 + x' \cdot \frac{2 - (-1)}{2^{22} - 1}$$

(000000000000000000000000)→-1

(111111111111111111111111)→2



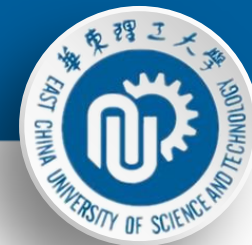
◆ 简单函数优化实例

■ 遗传操作

选择： 轮盘赌选择法

交叉： 单点交叉

变异： 小概率变异



◆ 简单函数优化实例

■ 模拟结果

设置的参数:

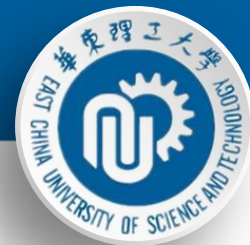
种群大小50; 交叉概率0.75; 变异概率0.05; 最大代数200。

得到的最佳个体:

$$s_{max} = \langle 1111001100111011111100 \rangle;$$

$$x_{max} = 1.8506;$$

$$f(x_{max}) = 3.8503;$$



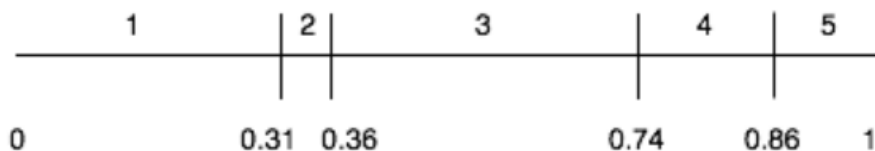
◆ 遗传操作——选择

■ 常用选择方法

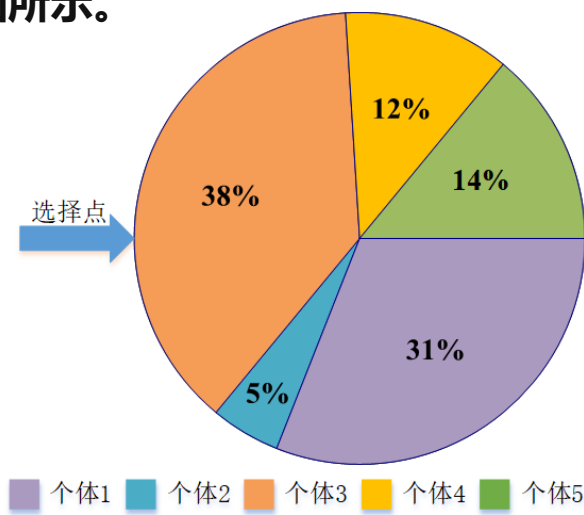
➤ 轮盘赌选择法 (roulette wheel selection)

首先将个体适应度值映射到轮盘中，个体的**适应度值越大**，其在轮盘中**分配到的角度就越大**，因此**被选中的概率就越大**。

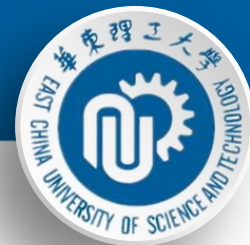
假设有5个个体，各适应度值分别为 $[6.82, 1.11, 8.48, 2.57, 3.08]$ ，总适应度值为22.06。分别计算**各个适应度值占总适应度值之比**，即为 $[0.31, 0.05, 0.38, 0.12, 0.14]$ 。
适应度值占比示意图和轮盘赌选择示意图分别如下图所示。



适应度值占比示意图



轮盘赌选择示意图



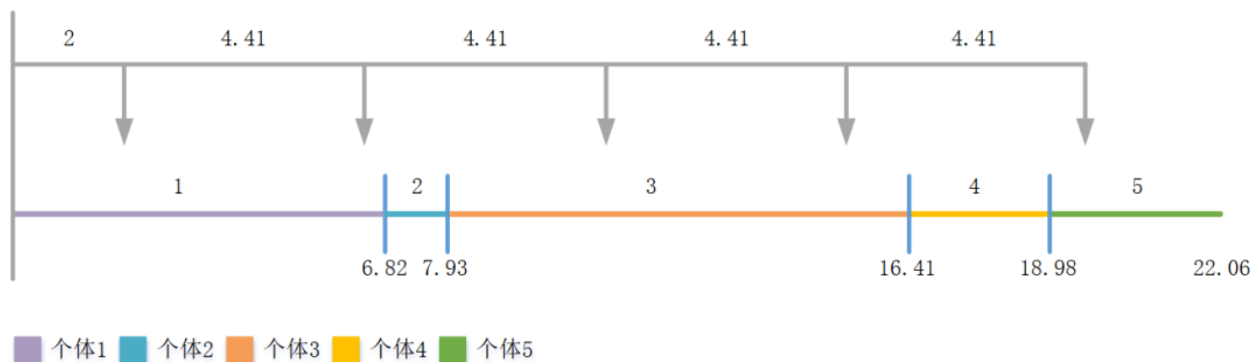
◆ 遗传操作——选择

■ 常用选择方法

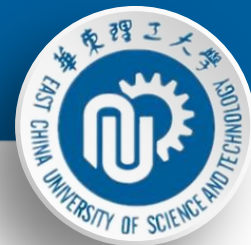
➤ 随机遍历抽样法 (stochastic universal sampling)

如果需要**选择N个个体**，则**只需一次生成N个等间距的标记指针位置**，即可选择出N个个体。假设总适应度值为F，选择个体数目为N，则SUS具体步骤如下：

- STEP1: 计算指针的间距 $P=F/N$;
- STEP2: 随机生成起点指针位置 $Start=[0\sim P\text{之间的随机数}]$;
- STEP3: 计算各指针的位置 $Pointers=[Start+i*P(\text{其中}i=[0,1,...N-1])]$;
- STEP4: 根据各指针位置，选择出N个个体。



随机遍历抽样执行过程



◆ 遗传操作——选择

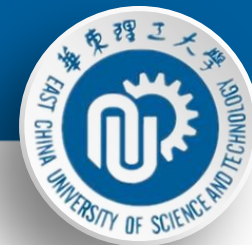
■ 常用选择方法

➤ 截断选择法 (truncation selection)

个体按适应度排列，只有优秀个体能够成为父个体，参数为截断阈值（被选作父个体的百分比）。

➤ 锦标赛选择法 (tournament selection)

随机从种群中挑选一定数目个体（竞赛规模），其中最好的个体作为父个体，此过程重复进行完成个体的选择。



◆ 遗传操作——交叉/基因重组

■ 实值重组

➤ 离散重组

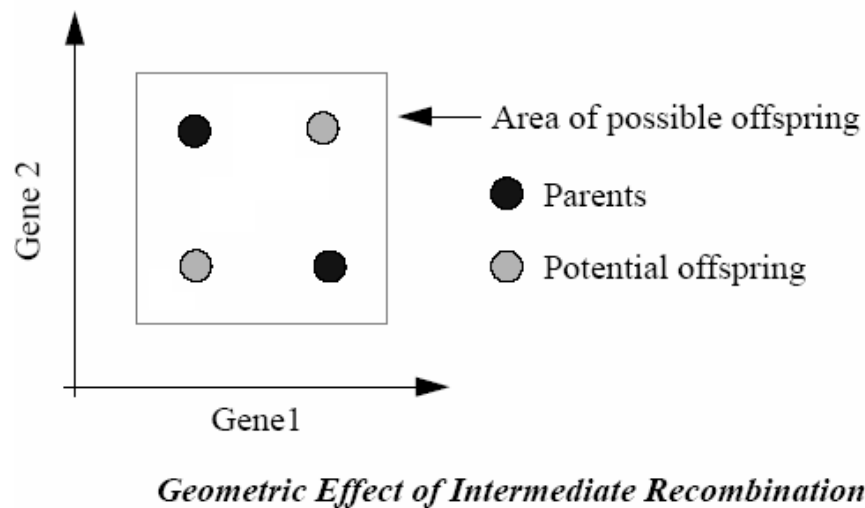
子个体的每个变量可以按等概率随机地挑选父个体。

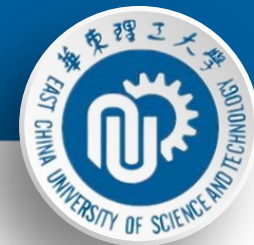
父个体1 12 25 5

父个体2 123 4 34

子个体1 123 4 5

子个体2 12 4 34





◆ 遗传操作——交叉/基因重组

■ 实值重组

➤ 中间重组

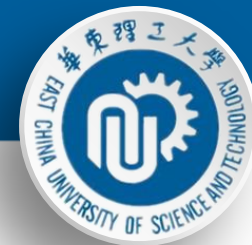
子个体 = 父个体1 + $\alpha \times$ (父个体2 - 父个体1)

α 是比例因子，由 $[-d, 1+d]$ 上均匀分布地随机数产生。

$d=0$ 时为中间重组，一般取 $d=0.25$ 。

子代的每个变量均产生一个 α 。

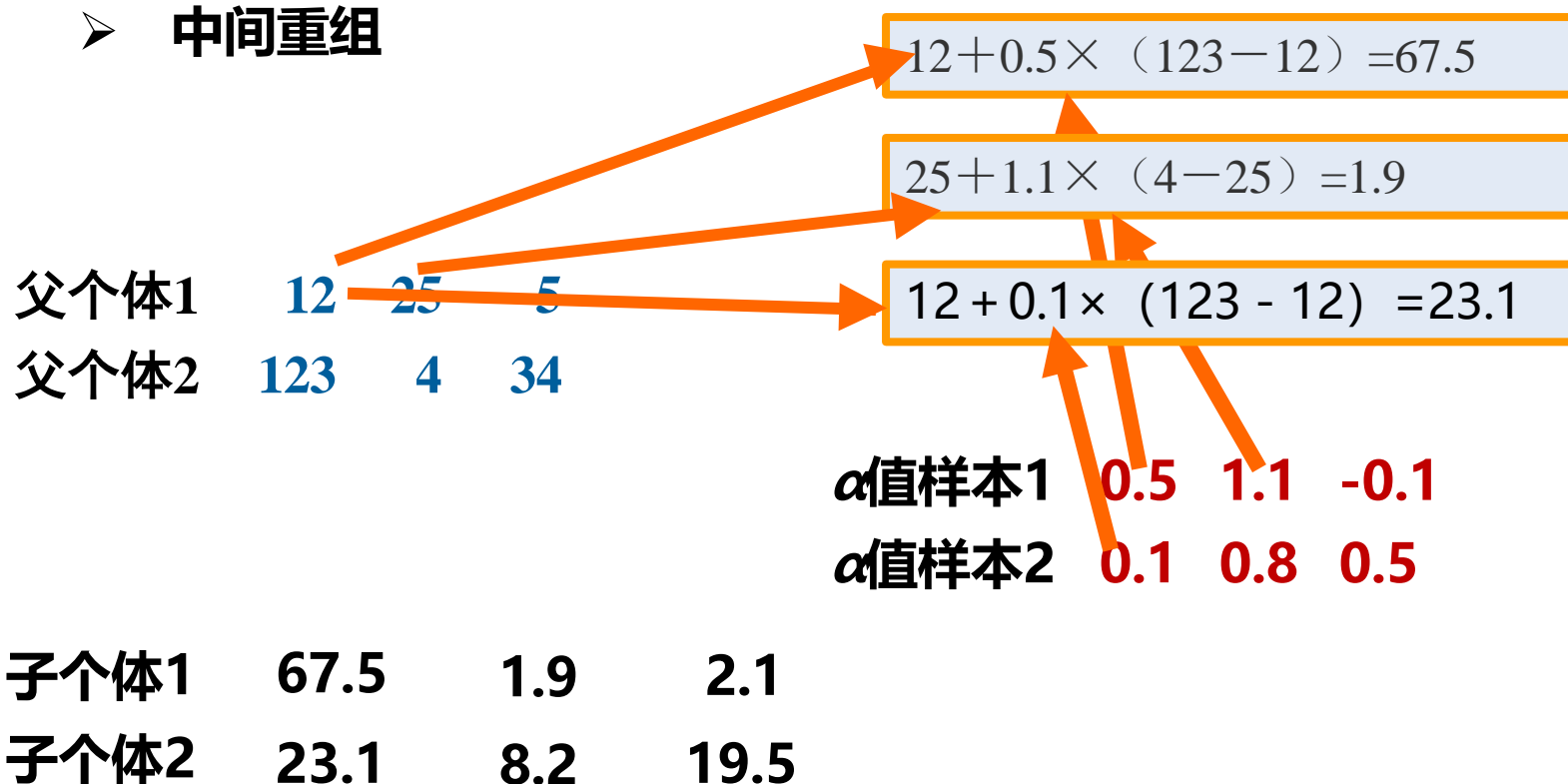


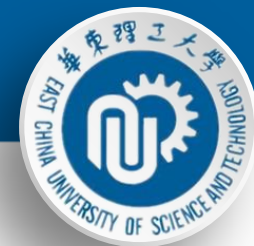


◆ 遗传操作——交叉/基因重组

■ 实值重组

➤ 中间重组

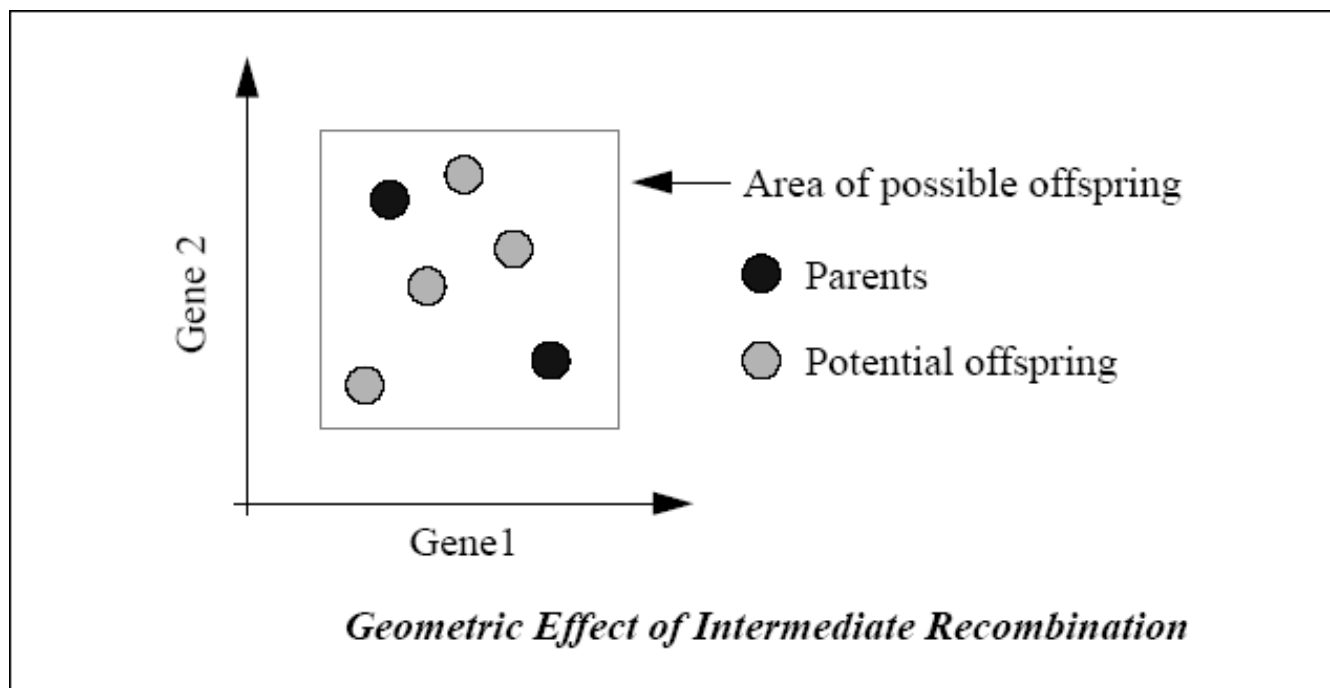


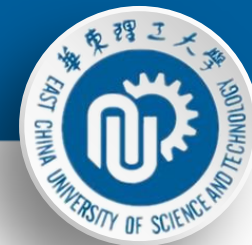


◆ 遗传操作——交叉/基因重组

■ 实值重组

➤ 中间重组

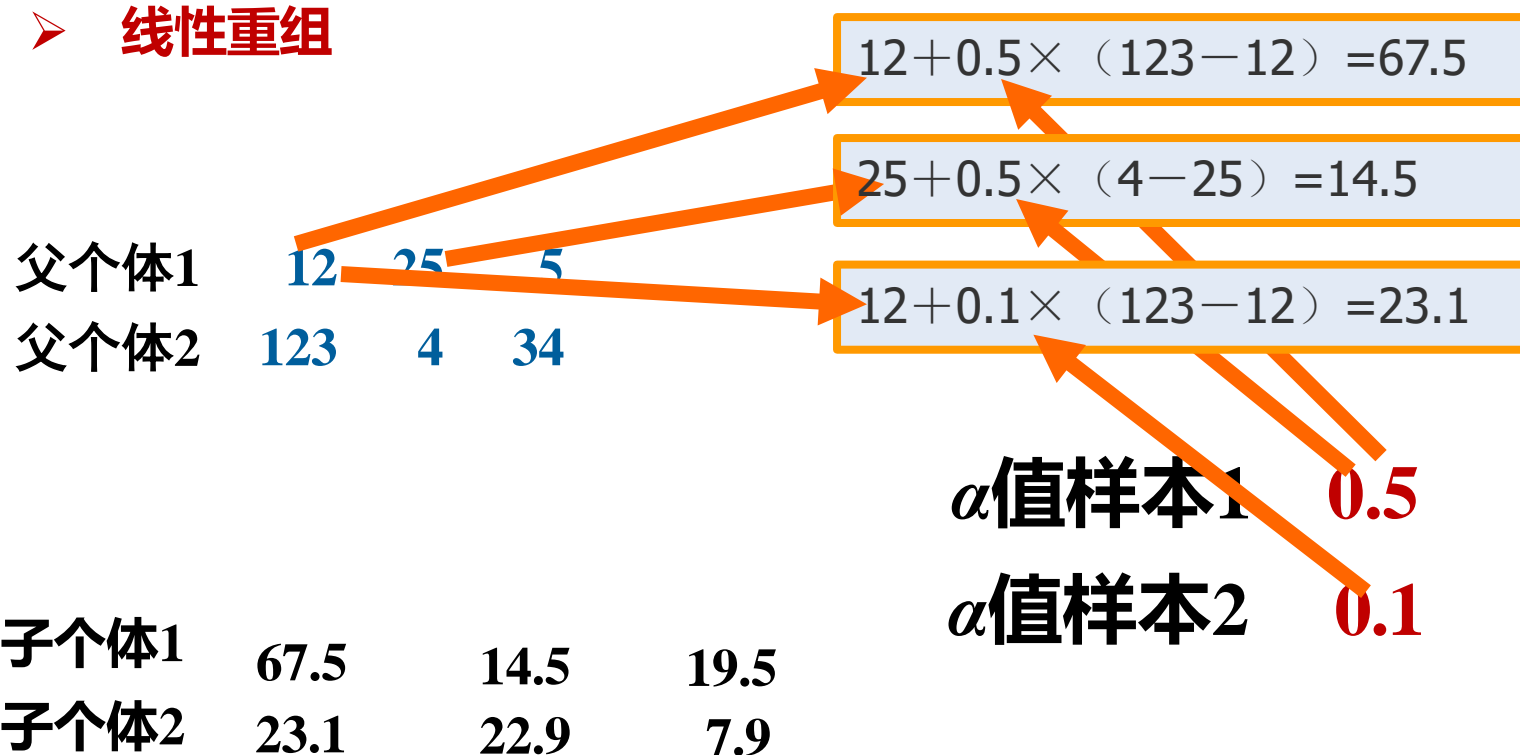


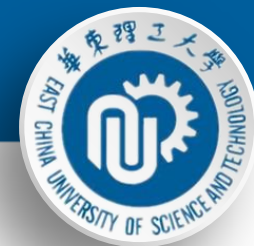


◆ 遗传操作——交叉/基因重组

■ 实值重组

➤ 线性重组

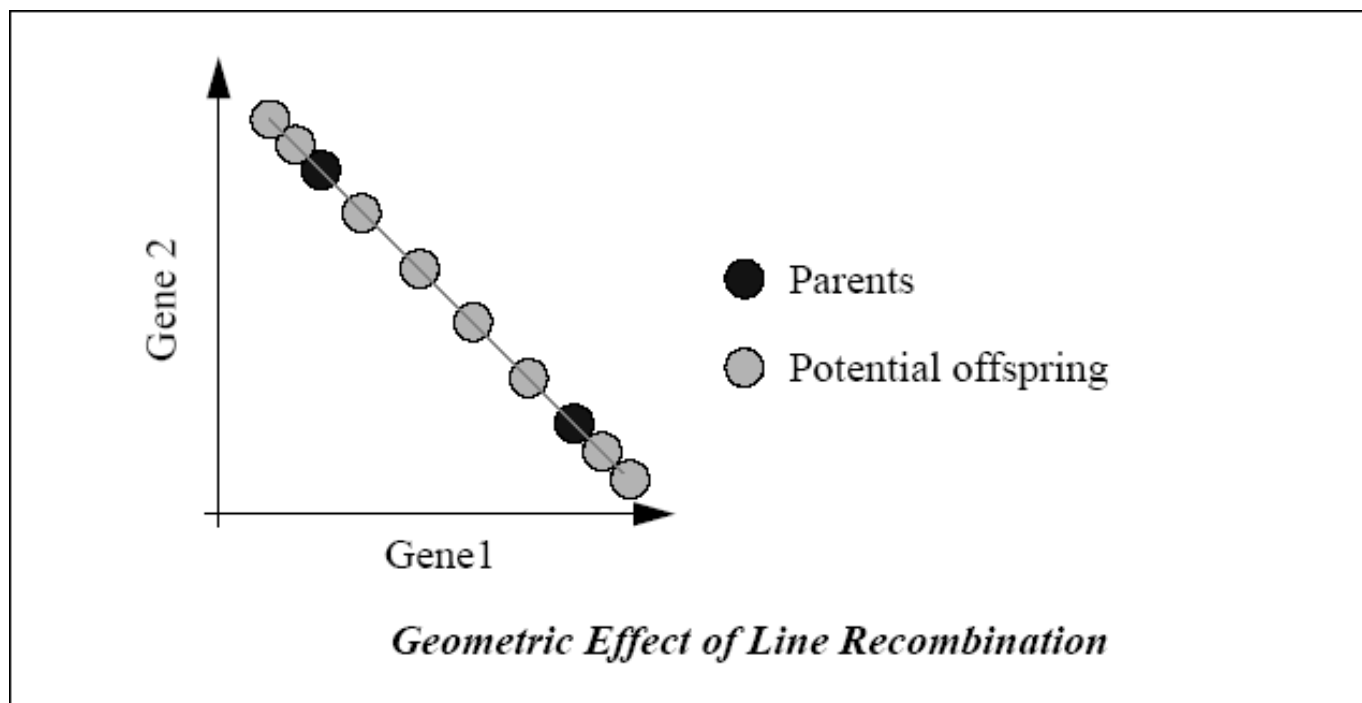


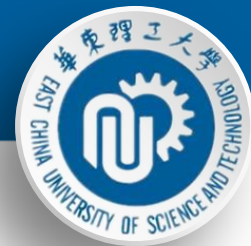


◆ 遗传操作——交叉/基因重组

■ 实值重组

➤ 线性重组



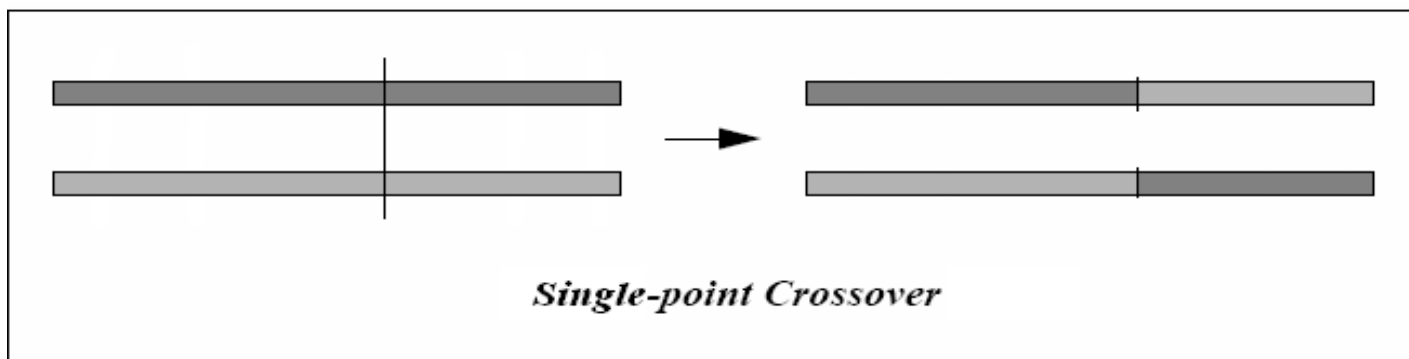


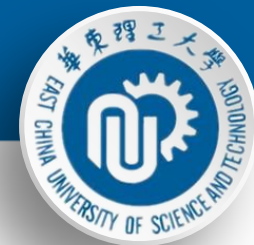
◆ 遗传操作——交叉/基因重组

■ 二进制交叉

➤ 单点交叉

单点交叉也称简单交叉，它是先在两个父代个体的编码串中随机设定一个交叉点，然后对这两个父代个体交叉点前面或后面部分的基因进行交换，并生成子代中的两个新的个体。



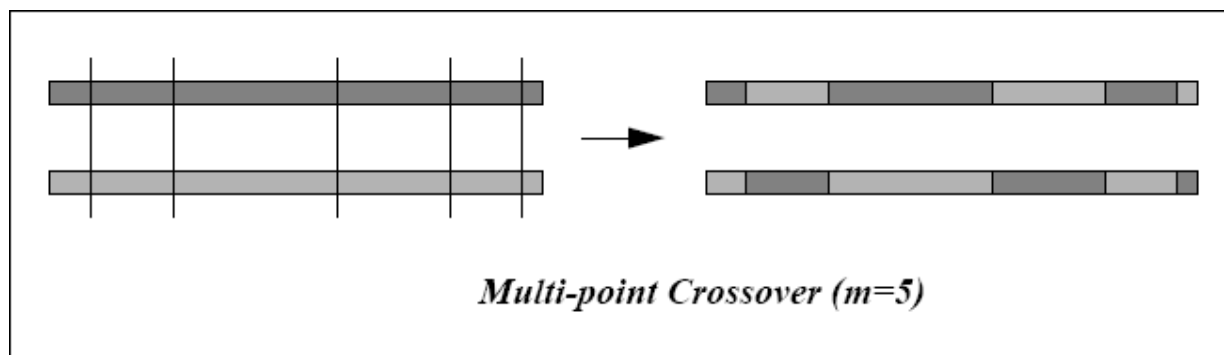


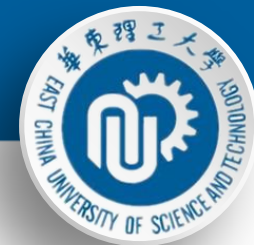
◆ 遗传操作——交叉/基因重组

■ 二进制交叉

➤ 多点交叉

多点交叉是指先在两个父代个体的编码串中随机设定多个交叉点，然后再按这些交叉点进行部分基因交换，生成子代中的两个新的个体。





◆ 遗传操作——交叉/基因重组

■ 二进制交叉

➤ 均匀交叉

均匀 (一致)交叉, 随机产生 $\{0,1\}$ 位串, 长度与个体的相等, 称屏蔽模板, 在其“1”的基因座上, 将父代①、②的基因分别传到子个体①、②; 在其“0”的基因座上, 将父代①、②的基因分别交叉传到子代②, ①;

父代① 1 1 0 1 0 0 1 1

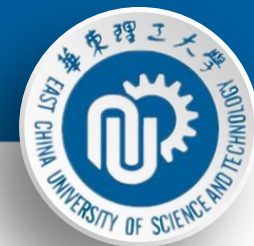
② 1 0 1 1 1 1 0 0

模板

1 0 1 1 0 1 0 1

子代① 1 0 0 1 1 0 0 1

② 1 1 1 1 0 1 1 0



◆ 遗传操作——变异

■ 实值变异

一般采用:

$$X' = X \pm 0.5L\Delta$$

$$\text{其中, } \Delta = \sum_{i=0}^m \frac{a(i)}{2^i}$$

$a(i)$ 以概率 $\frac{1}{m}$ 取值1, 以概率 $1 - \frac{1}{m}$ 取值0, 通常 $m=20$;
 L 为变量的取值范围

