

群体。

⑤ 判断是否满足结束条件。如果是,则结束,并将问题的特征描述及结果元素存入记忆库;否则,重复步骤(3)~(5)。

7) 免疫算法与遗传算法不同之处

(1) 搜索目的。遗传算法以搜索全局最优解为目标;而免疫算法在搜索全局最优解的同时,还可以搜索多峰值函数的多个极值,是以搜索多峰值函数的多个极值为目标的。

(2) 评价标准。基于以上搜索目的,遗传算法以解(个体或染色体)的适应度为唯一的评价标准;而在免疫算法中,为保持群体多样性,只有那些适应度高且浓度较低的个体才是最好的,是以解(个体或抗体)的适应度(亲和度值)及解(个体或抗体)本身浓度的综合作为评价标准。

(3) 交叉与变异操作的使用。在遗传算法中交叉操作作为主要操作之一,其作用是保留好的“基因”同时又给群体带来变化,而变异操作由于其变化较为激烈,只能作为算法中的辅助操作(使用的概率很小),从而保证算法的平稳全局收敛。在免疫算法中,为维持群体的多样性从而实现多峰值收敛,操作以克隆和变异为主,很少使用或不使用交叉操作。

(4) 记忆库。在遗传算法中没有记忆库这一概念。记忆库是受免疫系统具有免疫记忆特性的启示,在免疫算法结束时,将问题最后的解及问题的特征参数存入记忆库中,以便在下次遇到同类问题时可以借用这些结论,从而加快问题求解的速度,提高问题求解的效率。

从上面的分析可以看出,免疫算法与遗传算法的根本不同之处在于:群体更新策略,遗传算法仅依据个体的适应度进行复制,而免疫算法则基于个体的适应度和浓度进行复制,从而具有极强的多样性保持能力。因此,研究免疫算法是克服遗传算法群体多样性保持能力不足、易陷入局部最优等缺点的一个有效途径。

6.3 习题解析

1. 设计遗传算法必须经过哪几个基本步骤?

参考答案:

遗传算法的基本步骤如下。

(1) 选择编码策略,把参数集合 X 和域转换为相应编码空间 S 。

(2) 定义适应值函数 $f(X)$ 。

(3) 定义遗传策略,包括选择群体大小、选择、交叉、变异方法及确定交叉概率 P_c 、变异概率 P_m 等遗传参数。

(4) 随机初始化生成群体 $P(t)$ 。

(5) 计算群体中个体的适应值 $f(X)$ 。

(6) 按照遗传策略,运用选择、交叉和变异操作作用于群体,形成下一代群体。

(7) 判断群体性能是否满足某一指标,或者已完成预定迭代次数,不满足则返回步骤(6),或者修改遗传策略再返回步骤(6)。

2. 选择机制有哪几个基本策略?

参考答案:

选择机制有以下几个基本策略。

选择。

(1) 基于适应值比例的选择有繁殖池选择、转盘赌选择、Boltzmann 选择。

(2) 基于排名的选择有线性排名选择、非线性排名选择。

(3) 基于局部竞争机制的选择有锦标赛选择 (Tournament Selection)、 (μ, λ) 和 $\mu + \lambda$ 选择。

3. 遗传算法有哪几个控制参数? 各起什么作用?

参考答案:

在遗传算法的运行过程中, 存在着对其性能产生重大影响的一组参数。在初始阶段或群体进化过程中, 需要合理地选择和控制这组参数, 使 GA 以最佳的搜索轨迹达到最优解。遗传算法控制参数有位串长度、种群规模、交叉概率、变异概率。

4. 设计一个简单的遗传算法来说明遗传算法的工作机制, 这个算法至少包括复制、交叉和变异算子。

参考答案:

设计并实现一个求解具有 30 个城市的旅行商问题的遗传算法, 30 个城市的位置如图 6.5 所示。

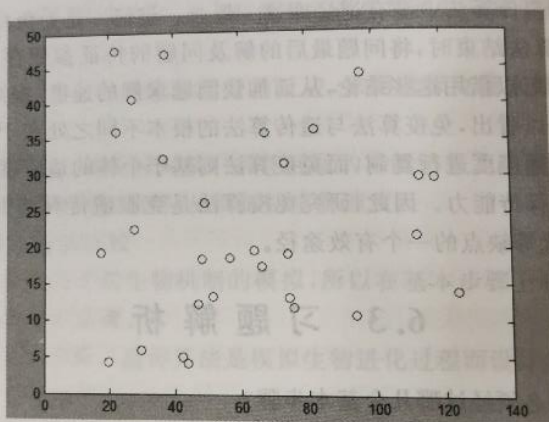


图 6.5 30 个城市的位置图

(1) 算法设计流程。遗传算法的算法设计流程如图 6.6 所示。

(2) 编码。例如, 一条周游路线为: 5—1—7—8—9—4—6—2—3—5, 那么其编码为 (5 1 7 8 9 4 6 2 3)。

$$\text{fitness} = \max\text{Dis} - \text{Distance} + 10$$

其中, $\max\text{Dis}$ 表示当前种群中的最大周游路径值, Distance 表示当前周游路径值, 为了方式适应度出现 0 的情况, 本实验在 $\max\text{Dis}$ 和 Distance 之差上加上了 10。

(3) 个体选择策略。转盘赌选择是遗传算法中用的最多的选择策略之一。在转盘赌选择中, 一个转盘被划分为 N 个扇形, 第 i ($i=1, 2, \dots, N$) 个扇形面积的大小与第 i 个个体的适应值成正比。然后转动转盘 N 次, 每次选择一个个体。

转盘赌选择与确定性选择的不同之处在于: 种群中的每个个体在转盘都有机会被选中。

(4) 杂交算子设计

对下面两个父个体的表示, 随机地选取两个交叉点“|”。例如:

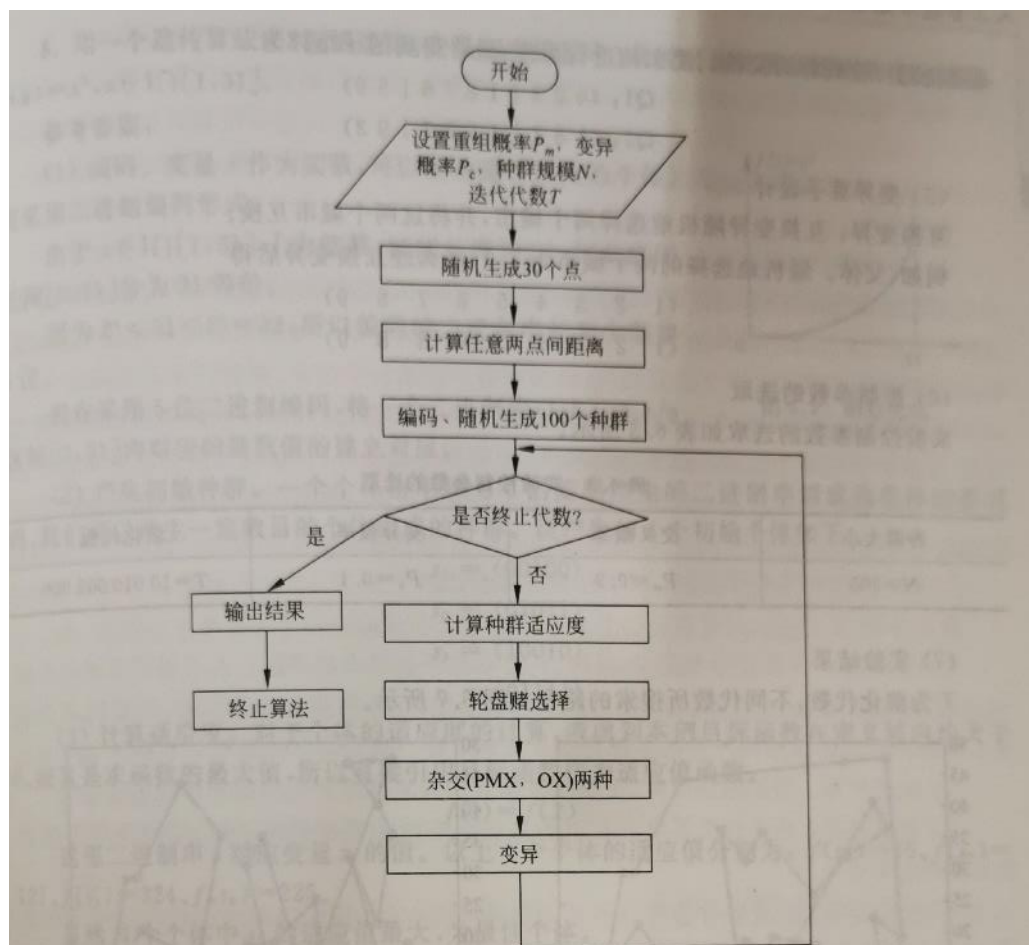


图 6.6 遗传算法的算法设计流程

P1: (1 2 3 | 4 5 6 7 | 8 9)

P2: (4 5 2 | 1 8 7 6 | 9 3)

首先,两个交叉点之间的中间段交换,得到

Q1: (x x x | 1 8 7 6 | x x)

Q2: (x x x | 4 5 6 7 | x x)

其中, x 表示暂未定码(本节不同),得到中间段的映射关系,有

$1 \leftrightarrow 4, 8 \leftrightarrow 5, 7 \leftrightarrow 6, 6 \leftrightarrow 7$

然后,对子个体1、子个体2中 x 部分,分别保留从其父个体中继承未选定城市码2、3、9,得到

Q1: (x 2 3 | 1 8 7 6 | x 9)

Q2: (x x 2 | 4 5 6 7 | 9 3)

最后,根据中间段的映射关系,对于上面子个体1的第一个 x ,使用最初父码1,由 $1 \leftrightarrow 4$ 交换得到第一个 x 为4,类似地,子个体2的第二个 x ,使用最初父码8,由 $8 \leftrightarrow 5$ 交换得到子个体1的第二个 x 为5。如果映射关系中存在传递关系,即备选交换有多个码,则选择此前

未确定的一个码作为交换。类似地进行操作,最终得到的子个体为

Q1: (4 2 3 | 1 8 7 6 | 5 9)

Q2: (1 8 2 | 4 5 6 7 | 9 3)

(5) 变异算子设计

互换变异: 互换变异随机地选择两个城市,并将这两个城市互换。

例如,父体。随机地选择的两个城市为 3、7,那么经互换变异后得

(1 2 3 4 5 6 7 8 9)

(1 2 7 4 5 6 3 8 9)

(6) 控制参数的选取

实验控制参数的选取如表 6.3 所示。

表 6.3 实验控制参数的选取

种群大小	交叉概率	变异概率	演化代数
$N=100$	$P_c=0.9$	$P_s=0.1$	$T=10\ 010\ 005\ 000$

(7) 实验结果

T 为演化代数,不同代数所搜索的路径如图 6.7 所示。

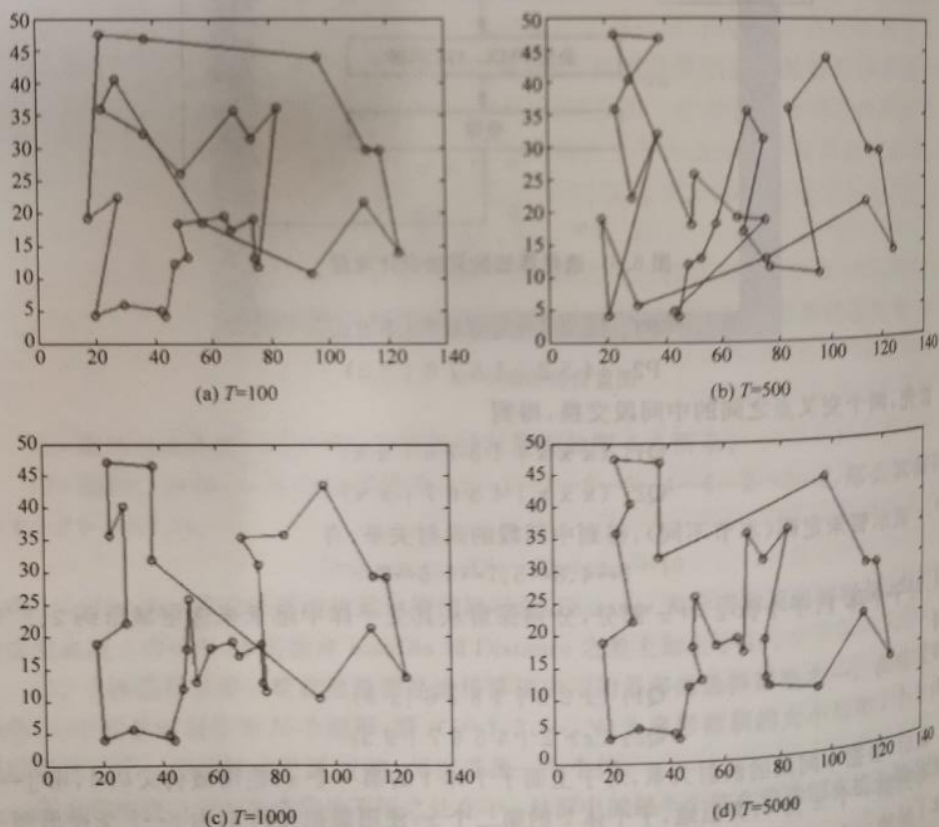


图 6.7 不同演化代数的效果

5. 用一个遗传算法来(包括选择、交叉和变异算子)求解图 6.8 所示的函数优化问题:
 $f(x)=x^2, x \in I \cap [1, 31]$.

参考答案:

(1) 编码。变量 x 作为实数, 可以视为遗传算法的个体。现采用二进制编码形式。

由于 $x \in I \cap [1, 31]$, I 为整数, 区间长度为 31, 因此将闭区间 $[0, 31]$ 分为 31 等份。

因为 $2^4 < 31 < 2^5 = 32$, 所以编码的二进制串长至少需要 5 位。

现在采用 5 位二进制编码, 将一个二进制串 $(b_5 b_4 \cdots b_0)$ 与区间 $[0, 31]$ 内对应的整数值建立对应。

(2) 产生初始种群。一个个体由串长为 5 的随机产生的二进制串组成染色体的基因码, 我们可以产生一定数目的个体组成的种群。设产生的 4 个初始个体如下。

$$s_1 = \langle 00100 \rangle$$

$$s_2 = \langle 01011 \rangle$$

$$s_3 = \langle 10010 \rangle$$

$$s_4 = \langle 01111 \rangle$$

(3) 计算适应度。对于个体的适应度的计算, 考虑到本例目标函数在定义域内均大于 0, 而且是求函数的最大值, 所以直接引用目标函数作为适应值函数。

$$f(s) = f(x)$$

这里二进制串 s 对应变变量 x 的值。以上 4 个个体的适应值分别为: $f(s_1) = 16$, $f(s_2) = 121$, $f(s_3) = 324$, $f(s_4) = 225$ 。

显然, 4 个个体中 s_3 的适应值最大, 为最佳个体。

(4) 遗传操作。设按转盘赌方式选择子个体, 选中的个体为 s_3 和 s_4 。

对 s_3 和 s_4 进行交叉操作, 选择一个交叉点, 如第 2 位与第 3 位之间的位置。交叉后产生新的子个体, 即

$$s'_3 = \langle 10111 \rangle$$

$$s'_4 = \langle 01010 \rangle$$

这两个子个体的适应值分别为

$$f(s'_3) = f(23) = 529$$

$$f(s'_4) = f(10) = 100$$

交叉后个体 s'_3 的适应值比其父个体的适应值高。

对于变异操作, 假设已经以一小概率选择了 s_3 的第 4 个遗传因子变异, 遗传因子由原来的 0 变成 1, 产生新的个体为

$$s'_3 = \langle 10000 \rangle$$

计算该个体的适应值为 $f(s'_3) = f(16) = 256$, 发现个体的适应值比其父个体的适应值减少了, 但是如果选择第 2 个遗传因子变异, 产生的新个体为

$$s''_3 = \langle 11010 \rangle$$

$$f(s''_3) = f(11010) = 676$$

显然, 这个个体的适应值比其父个体的适应值提高了。

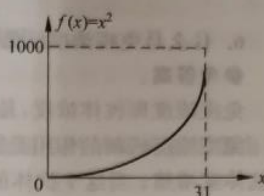


图 6.8 函数优化

(5) 模拟结果。按照标准的遗传算法,在运行到第若干代时获得最佳个体为

$$s_{\max} = \langle 11111 \rangle$$

$$f(x_{\max}) = 961$$

6. 什么是免疫浓度? 免疫浓度机制的作用是什么?

参考答案:

免疫浓度即抗体浓度,是指某抗体及与其相似抗体在群体中所占的比重。

免疫浓度机制的作用是控制和调节抗体的增殖:当抗体针对抗原具有较高的亲和度时,该抗体就增殖;当这个抗体的增殖浓度过高时就被抑制。这个过程能确保搜索方向的多样性,以防局部极大化。

免疫浓度机制可以让个体在一个特定的生存环境中进化,如果以抗体的相似度估计为标准,当群体中的某个抗体占据了相当规模,而又不是最优解时,就极易导致过早收敛。为此,当有些抗体的规模达到一定程度后,就要对其进行限制,预防个体同一化,防止其过早收敛而陷入“早熟”。

7. 免疫系统有哪几种功能? 其具体作用是什么?

参考答案:

免疫系统是一个具有分布式、自组织、具有动态平衡能力的系统。人工免疫系统就是研究、借鉴、利用生物免疫系统的原理和机制而发展起来的各种信息处理技术,计算技术及其在工程和科学中的应用而产生的多种智能系统的统称。其具体作用有以下3种。

(1) 免疫防御:即抗感染免疫,主要指机体针对外来抗原的免疫保护作用。

(2) 免疫自稳:机体免疫系统存在极为复杂而有效的调解网络,借以实现免疫系统功能的相对稳定性。

(3) 免疫监视:由于各种体内外因素的影响,正常个体的组织细胞不断发生畸变和突变,机体免疫系统可识别此类异常细胞并将其清除。

8. 免疫算法与遗传算法的基本步骤中,主要的差别是什么?

参考答案:

免疫算法与遗传算法不同之处有以下几点。

(1) 搜索目的。遗传算法以搜索全局最优解为目标;免疫算法在搜索全局最优解的同时,还可以搜索多峰值函数的多个极值,是以搜索多峰值函数的多个极值为目标的。

(2) 评价标准。遗传算法以解(个体或染色体)的适应度为唯一的评价标准。在免疫算法中,为保持群体多样性,只有那些适应度高且浓度较低的个体才是最好的,是以解(个体或抗体)的适应度(亲和度值)及解(个体或抗体)本身浓度的综合作为评价标准。

(3) 交叉与变异操作的使用。在遗传算法中交叉操作作为主要操作之一,其作用是保留好的“基因”同时又给群体带来变化,而变异操作由于其变化较为激烈,只能作为算法中的辅助操作(使用的概率很小),从而保证算法的平稳全局收敛。

在免疫算法中,为维持群体的多样性从而实现多峰值收敛,操作以克隆和变异为主,很少使用或不使用交叉操作。

(4) 记忆库。在遗传算法中没有记忆库这一概念。记忆库是受免疫系统具有免疫记忆特性的启示,在免疫算法结束时,将问题最后的解及问题的特征参数存入记忆库中,以便下次遇到同类问题时可以借用这些结论,从而加快问题求解的速度,提高问题求解的效率。

从上面的分析可以看出,免疫算法与遗传算法的根本不同之处在于:群体更新策略,遗传算法仅依据个体的适应度进行复制,而免疫算法则基于个体的适应度和浓度进行复制,从而具有极强的多样性保持能力。

9. 在生物免疫系统中,有哪几种淋巴细胞? T 细胞的作用是什么?

参考答案:

免疫系统使用两种类型的淋巴细胞: T 细胞和 B 细胞。

T 细胞即 T 淋巴细胞,它在胸腺中成熟,功能包括调节其他细胞的活动及直接袭击宿主感染细胞。

B 细胞即 B 淋巴细胞,来自于骨髓淋巴样前体细胞,成熟的 B 细胞存在于淋巴结、血液、脾、扁桃体等组织和器官中。B 细胞是体内产生抗体的细胞,在清除病原体过程中受到刺激,分泌抗体结合抗原,但其发挥免疫作用要受 T 辅助细胞的帮助。

T 细胞可分为毒性 T 细胞和调节 T 细胞两类。而调节 T 细胞又可分为辅助性 T 细胞和抑制性 T 细胞。辅助性 T 细胞的主要作用是激活 B 细胞,与抗原结合时分泌作用于 B 细胞并帮助刺激 B 细胞的分子;毒性 T 细胞能够清除微生物入侵者、病毒或者癌细胞。

10. 选用如下典型的多峰值函数作为实验对象来比较免疫算法与遗传算法。

$$F(x) = \exp(-(x - 0.1)^2) \sin^6(5\pi x^{3/4})$$

其中, $x \in [0, 1]$, 存在 5 个峰值, 分别用基本的免疫算法与遗传算法进行编程计算, 并将结果进行比较。

参考答案:

下面给出两个实验, 比较免疫算法与遗传算法。

实验 1

以搜寻全局最优点为目标, 比较两种方法的寻优性能。表 6.4 所示为两种算法全局寻优性能比较。

表 6.4 两种算法全局寻优性能比较

算法	交叉率(P_c)	变异率(P_m)	浓度阈值(T)
遗传算法	0.05	0.01	
	0.5	0.01	
	0.5	0.005	
	0.8	0.005	
免疫算法	0.6	0.05	0.85

从实验结果可以看出, 免疫算法本身具有寻到全局最优点的能力, 只是寻优的效率低于遗传算法, 但是通过调整参数, 免疫算法同样可以实现快速的全局寻优。当然, 经过改进, 遗传算法已具有了较好的寻优性能, 所以, 如果问题是要求全局最优解, 选择遗传算法更有益于问题的解决。

实验 2

分别用遗传算法和免疫算法进行搜索, 搜索过程以搜寻多峰值函数的几个峰值为目标, 并比较两种方法的搜索能力。算法中, 设 $P_c = 0.6$, $P_m = 0.05$ 。读者可以尝试用以上两个基本算法进行计算。

通过计算可以看出,免疫算法具有较好的多峰值搜索能力,能够快速、高效地找到各个极值点。而遗传算法几乎无法找到所有的极值点,即使调整算法参数,也只是使最终结果限于某一个极值点而已。可见,遗传算法多峰值搜索能力要差一些。对于多峰值函数的极值点搜索问题,采用免疫算法解决最为合适。

免疫算法和遗传算法都是来源于生物系统的启示而构造出来的随机启发式搜索算法,两者在实现形式上有着相似之处,但免疫系统和遗传进化系统从功能上讲,有着本质的不同,因而,免疫算法和遗传算法同样在功能上也存在不同,免疫算法实现的是多样性的搜索,它的搜索目标具有一定的分散性、独立性;而遗传算法主要用于全局最优解的求取,它的搜索目标具有单一性、排他性。

6.4 补充习题

1. 对遗传法的选择操作:设种群规模为4,个体采用二进制编码,适应度函数为 $f(x) = x^2$,初始种群情况如表 6.5 所示。

表 6.5 初始种群情况

编号	个体串	x	适应值	百分比	累计百分比	选中次数
S_{01}	1010	10				
S_{02}	0100	4				
S_{03}	1100	12				
S_{04}	0111	7				

若规定选择概率为 100%,选择算法为轮盘赌算法,且依次生成的 4 个随机数为 0.42、0.16、0.89、0.71,请填写表 6.5 中的全部内容,并求出经本次选择操作后所得到的新的种群。完成相应的计算后,表格的完整内容如表 6.6 所示。

表 6.6 完成表 6.5 中的计算后的完整内容

编号	个体串	x	适应值	百分比	累计百分比	选中次数
S_{01}	1010	10	100	32.36	32.36	1
S_{02}	0100	4	16	5.18	37.54	0
S_{03}	1100	12	144	44.60	84.14	2
S_{04}	0111	7	49	15.86	100	1

本轮选择后所得到的新的种群为

$$S_{01} = 1100$$

$$S_{02} = 1010$$

$$S_{03} = 0111$$

$$S_{04} = 1100$$

2. 用遗传算法求解八皇后问题(主要练习遗传算法设计)。

参考答案:

(1) 编码。用一个长度为 8 的数组来表示,如 32107654 的 3 就表示第 0 行的皇后放在

第3列。其中,数字为0~7的排列,因为皇后不能出现在同一行列。

(2) 适应度函数。如果 $(x_i - x_j)/(y_i - y_j) = 1$, 则表示有两个皇后处于同一斜线上, 此时则将 $f(x)$ 加1, 遍历完任意两对皇后后得到的 $f(x)$ 则是该序列的适应度函数的值。由于通常适应度函数越大表示该序列越好, 因此取 $1/f(x)$ 为适应度函数的值。

(3) 操作算子。

① 选择。可以采用课件上讲过的轮盘赌选择方法。

② 交叉。先选取一段片段进行交换, 之后未交换部分应该会与交换后的部分有相同的值, 再将未交换部分的重复值对应换过来即可。例如, 有 42315706 与 53760124 两对序列进行交叉, 将前两位进行交换, 则得到 53315706 与 42760124, 将第一个序列中的第三位的 3 与第二个序列中倒数第二位的 2 交换, 同时将第五位的 5 与第二序列的倒数第一位的 4 交换, 则得到 53214706 与 42760135。

③ 变异。随机挑选自身序列中的两位进行交换即为变异。

6.5 课堂演示与实践指导

6.5.1 课堂演示

1. 遗传算法求解 TSP 问题过程展示

本演示是教材的例 6.9 用 GA 求解一元函数最大值的优化问题, 即

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

双击打开 GADemo 应用程序后, 出现如图 6.9 所示的界面。

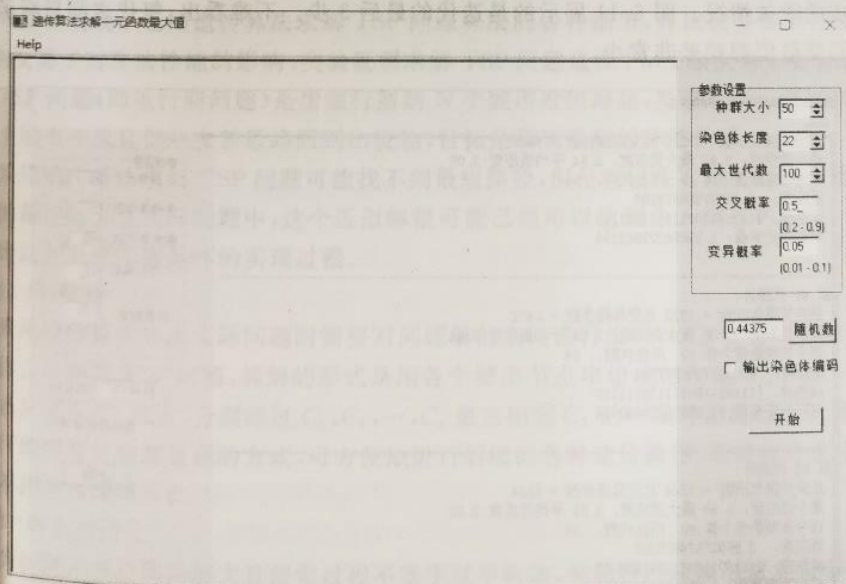


图 6.9 演示程序的初始界面