

计算机模拟实验报告

武子越 3170104155

1. 实验内容

完成针对 TSP 问题的遗传算法程序，并用 101 城市数据进行测试。简述问题、算子、参数设置及理由、实验数据分析、结论。

2. 基本思路描述

遗传算法涉及到的细节比较多，方便起见将每个步骤都分出一个函数来写：

◆ **main.m: 用于运行和调参**

◆ **TSPGA.m: 遗传算法求解 TSP 问题的主程序。**

程序包括数据导入处理、初始化种群、遗传的主要步骤、可视化，其中为了方便计算每个个体的总距离，主程序中计算了 101 个城市的距离矩阵，对于每一步迭代的种群，记录其中最优个体值和平均值。

◆ **Gene_select.m: 选择函数。**

每次迭代选择种群中的两个个体，其中使用了条件判断防止同一个个体选取两次。

◆ **Gene_cross.m: 交叉操作函数。**

随机产生两个交叉位 a, b, 根据交叉概率, 两个染色体交叉 Gca 到 Gcb 之间的位置, 在交叉的过程中, 判断染色体上是否有相同编码的情况, 即路径上有两个相同的城市, 如果有, 则该编码不参与交叉的操作。

◆ **Gene_mute.m: 变异函数**

随机产生两个变异位 a, b, 根据变异概率, 两个染色体变异 Gca 到 Gcb 之间的位置, 将该位置的染色体片段进行倒转。

◆ **Genefit_value.m: 计算每个个体的适应度（距离）**

◆ **Genefit_value_all.m: 计算种群的总适应度**

由于距离越短的解越好，因此通过取倒数的方式实现距离小的适应度更大。同时该函数计算累计概率，适应度大的将会有更大的概率被选中，为了提升算法的效率，使得优秀的个体更容易被选中，可以将优劣个体的概率差距拉大（这里采用了距离相反数的 T 次方，实验中 T 取 16）

◆ **plotcities.m: 绘制城市图（沿用上一次作业的绘图程序）**

◆ **Whether.m: 判断是否交叉或变异的函数。**

输入 0-1 的概率值，返回 0 或 1。

3. 参数的选择

在本程序当中有四个输入参数，分别为初始化种群大小，迭代次数，交叉概率，变异概率，为了探究参数对遗传算法表现的影响，每次固定其他参数，改变一个参数观察结果的变化，对于每一组参数做五次实验，记录平均结果，同时记录所有参数中最好的结果。

3.1 交叉概率和变异概率

在遗传算法实施中，改变交叉概率 cross 和变异概率 mute，考虑生物进化的实际情况，交叉概率应当适中，从而使得比较容易找到更多片段的组合，而变异概率稍小，用于跳出局部最优解的情况进行新的尝试，而不至于将已有的较好的解大量破坏，因此这里选取的交叉概率为 0.3-0.75，变异概率为 0.01-0.25，控制初始种群数量为 100，即开始时随机生成 100 组 TSP 城市路线方案，每个个体为长度 101 的向量。同时每组进行 1000 次迭代，五次实验

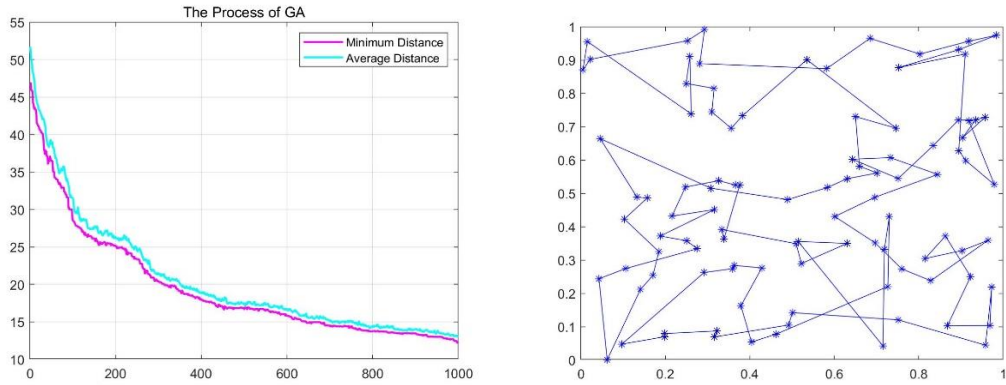
求得平均值，同时输出最好的结果及其对应参数。

实验结果如下，第一列为交叉概率，第二列为变异概率：

CROSS/MUTE	0.01	0.05	0.75	0.1	0.15	0.2	0.25
0.3	28.2344	19.4493	15.9028	16.3226	14.2322	14.724	13.452
0.4	27.1313	18.9567	15.9308	14.9379	14.5832	14.1218	14.3715
0.5	25.8333	18.066	16.5064	16.2716	13.7253	14.0702	12.5663
0.55	24.0399	17.5548	16.0098	15.5846	14.1467	13.9323	13.7447
0.6	24.1932	17.6499	16.9799	14.2784	13.4733	13.2595	13.275
0.65	21.4249	18.4336	16.499	15.0351	14.428	13.0953	14.3629
0.7	25.2668	18.0286	16.6383	15.5965	14.9915	13.4153	15.3083

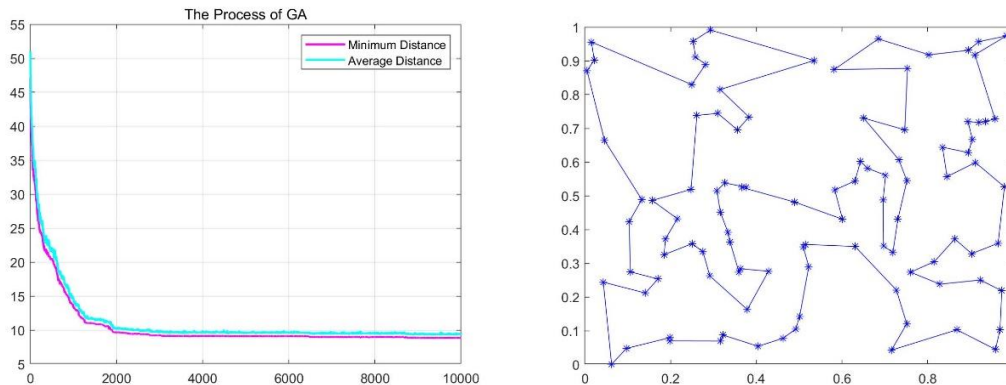
可以看到，在 1000 次迭代的情况下，交叉概率在 0.5 的时候得到的平均结果较好，不过各组结果差距不大，另外变异概率较小时，解的大小普遍在 20 多，增大变异概率，能够收敛到更好的解。

其中，在交叉概率等于 0.5，变异概率等于 0.25 时得到一组最好的解，其距离为 12.1730：



左图为遗传算法求得该解的过程，右图是实际的城市的分布情况，可以看到开始时总距离下降较快，后来逐渐变慢，最终收敛到一个局部最优解。粉色为每一代种群中最好的解，而蓝色为每一代解的平均值，可以看到二者的趋势基本一致。

另外考虑迭代次数的影响，这里迭代了 1000 次，但是从图上来看还有进一步下降的趋势，考虑到遗传算法也为一种穷举的方法，因此对于 $(num, cross, mute)=(100, 0.5, 0.25)$ 的参数情况进行一万次迭代，可以得到更好的解，距离为 8.7353：

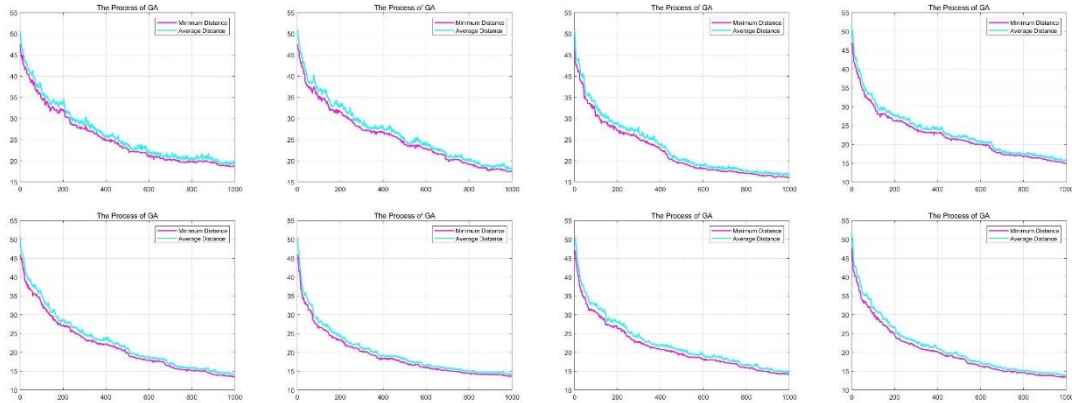


3.2 初始种群数量

上一步选出了一个较好的参数,即交叉概率 0.5,变异概率 0.25。保持这两个参数不变,改变初始种群数量进行 1000 次迭代,兼顾效果以及程序运行的效率,这里选择 30-100 的初始种群大小,进行五次实验求平均值,具体结果如下:

大小	30	40	50	60	70	80	90	100
结果	19.1009	17.4264	16.0073	14.9506	13.4101	14.4026	14.1642	13.3719

将其中最好的一次情况求解图画出,可以见到求解过程分别如下:



可以看到,初始种群较小的时候,折线更不平缓,变异对于种群整体的影响较大,随着种群数量的增加,更容易找到更好的解,求解过程也更为平缓。当然,种群大小过大也会导致每轮迭代的时间增加,计算开销将更大。

4. 结论

遗传算法中,初始种群数量,迭代次数,交叉率和变异率都会影响最后的结果,当然也和交叉算子,变异方式等选取有关,本实验选择跳过重复的编码的方法进行交叉,选择位点间片段翻转的方法进行变异,从上面的实验分析中我们可以看出:

◆ 初始种群数量

对于初始种群数量,从 30 到 70 左右时,解随着种群数量的增加在平均意义上逐渐变好,进一步增加种群数量则对于解的改善没有明显的效果,原因是迭代次数仍然小,较大种群数量理论上可以找到更好的解,但是时间开销也比较大,可能需要上万次迭代才能显示出改进的效果。

◆ 交叉概率

对于交叉概率,通常遗传算法会选择 0.5-0.99 之间的数字,实验中我们看到,遗传概率中等(在 0.5-0.6)左右的时候得到的解相对较好。

◆ 变异概率

通常的遗传算法变异概率设置较小,然而增大变异概率有助于尝试更多种组合,也有机会得到更好的解,然而变异概率过大会使得图上产生较多锯齿形,也容易丢到已经找到的比较好的路线组合。

◆ 迭代次数

迭代次数,迭代次数越大越有可能找到更好的解,然而在一万次之后再增加迭代次数会大大加大时间开销,对于解的改善反而不够明显,实际应用中 also 显得不够经济。

5. 参考文献

Genetic Algorithms Demystified [M] Peter Leow. 2016