第 18 卷 第 6 期　　电子科学学刊　　Vol.18 No.6
1996 年 11 月　　JOURNAL OF ELECTRONICS　　Nov. 1996

遗传算法中突变算子的数学分析及改进策略 [1]

张良杰　　毛志宏　　李衍达

(清华大学自动化系 北京 100084)

摘 要　本文在简要介绍遗传算法的基础上，通过引入 i 位改进子空间的概念，对不同情形下突变概率的最优选取进行了分析，然后采用模糊推理技术来确定选取突变概率的一般性原则。良好的仿真结果显示了本文所提改进策略的有效性。

关键词　遗传算法 (GA), i 位改进子空间，模糊推理

中图号　TN-052, TP872

1 引 言

遗传算法 (Genetic Algorithm) 简称 GA，是一种全局意义上的自适应搜索技术，它的得来是受启发于自然界的选择陶汰规则及遗传学原理。其突出特点在于包含了与生物遗传及进化很相似的步骤，如选择 (selection)，复制 (reproduction)，交配 (crossover) 及突变 (mutation) 等，这些称为遗传算子。达尔文的生物进化及适者生存等为人熟知及承认的进化理论及客观存在的进化现象是 GA 可行性的最好保证。由于突变算子的存在，保证了 GA 能搜索到问题解空间的每一点，并且只要遗传搜索中每一代都包含上一代的最优个体，则搜索便会收敛到全局最优解 [1]。同时，由于 GA 所特有的“内含并行”特性，在每一代，尽管它仅计算 G 个个体，而实际处理模式数却是 $O(G^3)$[2]，所以它是一种比较迅速的学习算法，已经在许多领域中得到了成功的应用 [3]；如机器学习，模式识别，神经网络结构参数优化及权重学习 [4-6]，模糊逻辑控制器的设计 [7] 等等。并且与其它寻优算法的优点相结合的 GA 也相应而生 [8]。

但是目前对于 GA 的原理的理解及应用还处于初期，有许多亟待完善之处 [2,3-9]。例如，常规 GA 中，群体大小的选择常凭经验；交配方式中交叉点的选择没有依据，造成有时进化效率很低；突变概率的确定也缺乏指导等，这些都限制了 GA 的收敛速度。许多研究者对 GA 的改进展开了一些研究工作。例如，J. H. Holland[2] 描述了一类 GA 的框架，并给出了 GA 的几种形式。这些研究主要集中于由以下 6 个参数描述的一类特定的 GA，即群体大小 (Population size)、交配率 (Crossover rate)、变异率 (Mutation rate)、代间隔 (Generation gap)、比例窗 (Scaling window)、选择策略 (Selection strategy) 等。J. J. Grefenstette[9] 在 1986 年试图对具有全局搜索特性的 GA 确定其最佳控制参数，从而采用了对 GA 的参数空间进行全局搜索的实验研究方法，来寻找对特定问题而言的有效 GA 等等。但以上改进方法均缺乏清晰的数学模型及足够的数学分析给予搜索指导。

鉴于此，我们展开了对常规 GA 的诸算子进行分析及改进的工作。初步的研究结果在文献 [7] 中进行了介绍。本文在文献 [7] 的基础上，对突变算子进行了必要的数学分析，给出了相应的实现策略。

[1] 1995-04-05 收到，1995-07-05 定稿
国家攀登计划认知科学（神经网络）重大关键项目和教委博士点基金资助课题

2 常规的遗传算法

一般来说，GA 总可把待解问题转化为如下一个寻优问题，即

$$\min\{f(C)|C \in IB^N\}. \tag{1}$$

这里假设对 $\forall C \in IB^N = \{0,1\}^N$，$0 < f(C) < \infty$ 且 $f(C) \neq \mathrm{const}$ 。利用 GA 解决此类优化问题的主要步骤如下：首先要将待解问题的侯选解按照一定的规则编码成字符串 (通常是 0,1) 组成，一条字符串则类似于一条基因串，其中蕴含着该问题解的信息。然后，再给出评价各字符串 (基因串) 性能的办法，通常是定义一适合度函数值 f，以评价基因串的优劣。GA 在开始时随机地产生一定数目的个体，通过运用遗传算子对基因串进行组合变化，得到新的个体群；通过计算其中各基因串的适合度函数值 $f_i, (i = 1, 2, \cdots)$，得到对这些基因串优劣的评价。再根据优劣分配繁殖机会，优的个体分配的机会多，产生的子代机会也就多；劣的个体分配的机会少，产生的子代少，甚至被性能更优的子代代替。被选个体又会形成新的一代，这样搜索按照一代一代逐步朝着更好的解的方向进化。此过程一直重复，直到终结条件得到满足。

下面即为一种常规 GA 的算法格式：

```
产生初始的个体群；计算每个个体的适合度函数值；进行选择
REPEAT
    进行交配；进行突变；计算每个个体的适合度函数值；进行选择
UNTIL 满足终结条件
```

GA 各种算子在不同的问题中可能就有不同的实现方法，更具体的描述参见文献 [10] 。

3 突变概率选择的数学分析及其模糊规则实现策略

常规 GA 的突变概率 P_m(基因串中某一位发生突变的概率) 是一个常数。至于如何确定合适的 P_m，目前还没有一个统一的原则，况且固定的 P_m 往往不能同时适合寻优搜索中不同情形的要求。因此我们在讨论不同情况下搜索的快速性对 P_m 要求的基础上，自然地引入利用模糊推理动态选取 P_m 的方法。

设一个初始的串 C_0 突变为 C_1 。因为 C_1 优于 C_0 的概率越大，越有利于进化，所以对于问题 (1) 式，用 $P(f(C_1) < f(C_0))$ 的大小作为选取 P_m 的依据。在对这个概率的进一步讨论之前，先给出两个定义。

定义 1　串 $C_1, C_2 \in IB^N = \{0,1\}^N$ 之间的汉明距离 (Hamming Distance) 为

$$H(C_1, C_2) = \sum_{i=1}^{N} |B_i(C_1) - B_i(C_2)|, \tag{2}$$

其中 $B_i(C)$ 指串 C 中第 i 个位置上的码数 $\{0,1\}$，$i = 1, 2, \cdots, N$ 。

定义 2　在问题 (1) 式中，对串 C_0 定义如下几类 $\Omega_0^{(1)}, \Omega_0^{(2)}, \cdots, \Omega_0^{(i)}, \cdots$，其中 $\Omega_0^{(i)} = \{C|C \in IB^N,\ f(C) < f(C_0),\ H(C, C_0) = i\}$ 。称 $\Omega_0^{(i)}$ 为 C_0 的 i 位改进子空间，$\bigcup_{i=1,2,\cdots} \Omega_0^{(i)}$ 为 C_0 的改进空间，简记为 $\bigcup_i \Omega_0^{(i)}$ 。

可以看出，如果 $C_1 \in \bigcup_i \Omega_0^{(i)}$ ，则 C_1 优于 C_0 ；否则 C_1 相对于 C_0 没有改进。 C_1 优于 C_0 的概率等于 $C_1 \in \bigcup_i \Omega_0^{(i)}$ 的概率，即

$$P(f(C_1) < f(C_0)) = P(C_1 \in \bigcup_i \Omega_0^{(i)}). \tag{3}$$

P_m 的选取应使这个概率越大越好，这就是搜索过程中选取 P_m 的一般性原则。下面来确定这个概率具体的表达式。

从 C_0 到 C_1 ，有 $H(C_0,C_1)$ 个位发生了改变， $N-H(C_0,C_1)$ 个位未改变，则 $P(C_0 \to C_1) = P_m^{H(C_0,C_1)}(1-P_m)^{N-H(C_0,C_1)}$ ，进而可求得概率

$$P(C_1 \in \Omega_0^{(i)}) = |\Omega_0^{(i)}| P_m^i (1-P_m)^{N-i}, \tag{4}$$

其中 $|\Omega_0^{(i)}|$ 为 $\Omega_0^{(i)}$ 中元素的个数，由于 $\Omega_0^{(i)} \bigcap \Omega_0^{(j)} = \Phi,\quad i \neq j$ ，有

$$P(C_1 \in \bigcup_i \Omega_0^{(i)}) = \sum_i [|\Omega_0^{(i)}| P_m^i (1-P_m)^{N-i}]. \tag{5}$$

定理 C_0 突变为 C_1 进入 i 位改进子空间的概率 $P(C_1 \in \Omega_0^{(i)})$ 在 $P_m = i/N$ 时取到最大。

证明 (4) 式两边取对数，有

$$\ln P(C_1 \in \Omega_0^{(i)}) = \ln|\Omega_0^{(i)}| + i\ln P_m + (N-i)\ln(1-P_m). \tag{6}$$

令 $\partial(\ln P(C_1 \in \Omega_0^{(i)}))/\partial P_m = i/P_m - (N-i)/(1-P_m) = 0$ ，得 $P_m = i/N$. $\partial^2(\ln P(C_1 \in \Omega_0^{(i)}))/\partial P_m^2 = -i/P_m^2 - (N-i)/(1-P_m)^2 < 0$, 故 $\ln P(C_1 \in \Omega_0^{(i)})$ 在 $P_m \in (0,1)$ 上是凸的，最大值在驻点达到，即 $P(C_1 \in \Omega_0^{(i)})$ 在 $P_m = i/N$ 时取到最大。 证毕

在 GA 的许多仿真实验中，我们注意到这样的现象：在搜索的后期，若 P_m 很小，收敛的速度就会很慢；相反若 P_m 取得较大，基因串中同时有多位发生突变，搜索就不易收敛；有限搜索会遇到局部最优，这时若 P_m 取小了，则不利于跳出局部最优向更好的方向搜索。在这种情况下如何选取最优的 P_m ，可通过进一步的推导得到启示。

(1) 在搜索后期，一代中的每个串已经很接近于最优解，对于其中之一的 C_0 ，属于它的改进空间的元素已经不多，即 $|\bigcup_i \Omega_0^{(i)}|$ 不大，一般来说， $|\Omega_0^{(1)}| \neq 0$ ，则

$$\begin{aligned} P(C_1 \in \bigcup_i \Omega_0^{(i)}) &= \sum_i [|\Omega_0^{(i)}| P_m^i (1-P_m)^{N-i}] \\ &= |\Omega_0^{(1)}| P_m (1-P_m)^{N-1} \left[\left(1 + \frac{|\Omega_0^{(2)}|}{|\Omega_0^{(1)}|}\frac{P_m}{1-P_m} + \frac{|\Omega_0^{(3)}|}{|\Omega_0^{(1)}|}(\frac{P_m}{1-P_m})^2 + \cdots\right)\right]. \end{aligned} \tag{7}$$

因通常 $P_m/(1-P_m) << 1$ ，故

$$P(C_1 \in \bigcup_i \Omega_0^{(i)}) \leq |\Omega_0^{(1)}| P_m (1-P_m)^{N-1} \left[\left(1 + \left(|\bigcup_i \Omega_0^{(i)}| / |\Omega_0^{(1)}|\right) P_m/(1-P_m)\right)\right]. \tag{8}$$

又因 $|\bigcup_i \Omega_0^{(i)}|$ 不大，故 (7) 式 $\approx |\Omega_0^{(1)}| P_m (1-P_m)^{N-1}$ ，即 C_0 突变后有所改进的概率近似等于使 C_0 突变后进入 1 位改进子空间的概率，由定理可知， $P_m = 1/N$ 时最优。这也就确定了在搜索后期选取 P_m 的原则。

(2) 在遇到局部最优时，(设 C_0 达到局部最优)，对 C_0 进行为数不多的位数改变不能使子代有所改进，即在 i 不大时，$|\Omega_0^{(1)}|=0$ 。此时不妨设 $|\Omega_0^{(r)}|\neq 0, |\Omega_0^{(i)}|=0, (i<r)$ ，则

$$P(C_1\in\bigcup_i\Omega_0^{(i)})=\sum_{i\geq r}[|\Omega_0^{(i)}|P_m^i(1-P_m)^{N-i}]. \tag{9}$$

由前面的定理可知，当 $P_m\leq r/N$ 时，$|\Omega_0^{(i)}|P_m^i(1-P_m)^{N-i}, (i\geq r)$ 都是递增的，因此 $P(C_1\in\bigcup_i\Omega_0^{(i)})$ 取最大时的 P_m 必大于等于 r/N 。这个结论并没有给出最优 P_m 的精确取值，却确定了遇到局部最优时选取 P_m 的范围。

通过上面的分析，可以看出，搜索的快速性要求 P_m 有一定的“伸缩性”，即 P_m 是一个条件概率，在寻优搜索中，当然可以用一个具体函数，比如 $g(*)$ 来描述 P_m ，但是要确定 $g(*)$ 的形式，使它能够反映一些不易精确表达的结论，还是有很大难度的，即便它能反映 P_m 的选取规律，其精确的函数形式也是不易理解的，更不易于调整及修正。为了克服这个困难，我们采取了模糊推理技术，用它来刻画 P_m 则十分合适而自然，因为一方面它能够比较容易地反映前面理论分析的结论 (P_m 的选取规律)，又能较为方便地融入对具体问题的一些已有的经验，同时其推理形式亦易于理解。

用模糊推理技术选取 P_m 的方法是依靠一组模糊规则来实现的，这些规则形如：

$$\text{规则 } i \quad \text{IF 搜索进入了状态 } Q_i, \qquad \text{THEN } P_m=W_i, \tag{10}$$

其中 Q_i 是描述 GA 搜索状态的模糊语言变量，W_i 是描述 P_m 大小的模糊语言变量，各语言变量的选取需依具体问题而定，模糊规则中应融入前面的定理及讨论的结论，并可包含针对具体问题的已有的经验。

例如，在问题 (1) 式中，若 P_m 取为如下形式的条件概率：

$$P_m=P(^*|f,\mathrm{d}f), \tag{11}$$

这里 f 指一代中最优个体的适合度函数值，$\mathrm{d}f$ 是 f 的改变量，则通过 f ，$\mathrm{d}f$ 的大小确定 P_m 的模糊规则可取如下：

$$\left.\begin{array}{l}\text{规则 1 IF } f=\mathrm{PB}, \text{ and } \mathrm{d}f=\mathrm{NB}', \text{ THEN } P_m=\mathrm{PM}'',\\ \text{规则 2 IF } f=\mathrm{PB}, \text{ and } \mathrm{d}f=\mathrm{NS}', \text{ THEN } P_m=\mathrm{PB}'',\\ \text{规则 3 IF } f=\mathrm{PS}, \text{ THEN } P_m=\mathrm{PS}'',\end{array}\right\} \tag{12}$$

其中 PB 、PS(正大、正小) 是描述 f 的模糊语言变量，NB$'$ ，NS$'$(负大、负小) 是描述 $\mathrm{d}f$ 的模糊语言变量；PB$''$ 、PM$''$ 、PS$''$(正大、正中、正小) 则是描述 P_m 的模糊语言变量。根据 f ，$\mathrm{d}f$ 的大小，判断搜索进入了什么状态，再来选取合适的 P_m 。规则 1 是针对搜索前期的，此时 P_m 取值为 PM$''$ ；规则 2 是针对局部最优的情形的，此时 P_m 取值为 PB$''$ ；规则 3 是针对搜索后期的，这时 P_m 取值为 PS$''$ 。具体到每个模糊语言变量的确定 (包括隶属函数的形式等)，则要参照实际问题而定。

在本文的两个仿真实验中就使用了形如 (12) 的模糊规则。由于 f 都是整值，在搜索的中后期，相邻两代的 f 值相差又不太大 (以 0,–1 居多)，$\mathrm{d}f$ 不容易用 NS$'$ 或 NB$'$ 来描述，所以定义 $\mathrm{d}f$ 为当前一代的 f 与比之早 5 代的 f 的差。另外，为了简单起见，用 $P_m=2/N, P_m=5/N$,

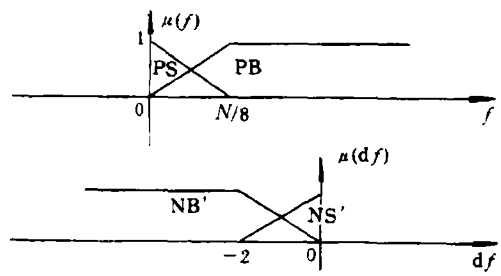

图 1 模式匹配中描述 f 和 df 的语言变量的隶属函数

$P_m=0.8/N$ 替代 $P_m=PM''$, $P_m=PB''$, $P_m=PS''$，即将 P_m 的模糊语言变量的隶属函数取为单点型隶属函数 (常数型后件)。

将经验知识用模糊推理实现后，便可在 GA 的突变概率自调整中发挥作用。实际上，这种利用启发式知识使传统算法搜索参数动态化的方法是很有潜力的。

4 计算机仿真实验

为了比较改进突变算子后的 GA 与常规 GA 在搜索中迭代次数的多少，我们设计了模式匹配及倒立摆模糊控制器的设计问题来评价改进算法的效果，这里各表中的迭代数为多次实验数据统计平均后取整得到的。

4.1 模式匹配问题

表 1 模式匹配问题仿真结果

串长 (N)	20	40	60	80	100
常规 GA($P_m=0.1/N$) 的搜索次数	20	47	72	107	174
常规 GA($P_m=2/N$) 的搜索次数	39	56	74	106	119
模糊规则确定 P_m 型 GA 的搜索次数	12	26	42	62	74

表 2 FLC 设计的仿真结果

$(\theta_0,\dot{\theta}_0)$	(0,0)	(5,0)	(5,5)	(10,0)
常规 GA 的搜索次数	13	27	39	53
改进的 GA 的搜索次数	2	7	12	38
限定步数 (f_{rs})	1	5	7	9

以串 C_i 到目标串 (最优解)C^* 的汉明距离作为 C_i 的适合度函数 f_i。我们对几组不同长度的目标串作了实验，在表 1 中，列出了 P_m 固定的两种情况下的结果和按模糊规则来确定 P_m 得到的结果，其中，一代群体数目维持在 G=20。可以看出在这个问题中突变算子的改进很有效果。确定 P_m 的模糊规则如 (12) 式所述，模糊规则中描述 f 和 df 的语言变量的隶属函数的形状如图 1 所示。

4.2 倒立摆系统模糊控制器的设计

这个仿真实验的详细描述 (包括倒立摆系统的数学模型、它的模糊逻辑控制器 (FLC) 的构成、用 GA 学习模糊规则及利用模糊规则选取 P_m 的部分) 请参见文献 [7]，在这里仅列出仿真结果。表 2 显示了常规 GA 与改进的 GA(其交配方式为“均匀交换”式 [7]) 的比较情况。其中 $(\theta_0,\dot{\theta}_0)$ 为倒立摆的初始角度 (与垂直位置的夹角) 及角速度，f_{rs} 为限定步数，若在这个步数范围内倒立摆能首次进入平衡态区间并能维持稳定，则搜索停止。

可见，在这个仿真实验中，改进的 GA 亦明显优于常规 GA 。

5 结　论

本文讨论了 GA 中突变算子的改进问题，并把改进后的 GA 应用到一类模式匹配问题及倒立摆模糊控制器的设计中。从仿真结果可以看出改进的 GA 相对常规 GA 而言有了明显的改善。在对突变算子的分析中，引入 i 位改进子空间的概念，对突变概率 P_m 的动态选取原则作了一般性的讨论，并针对一些特殊情形进行了具体的分析，然后采用了融入了实际操作经验的模糊推理技术来确定 P_m，从而使常规遗传算法的突变概率 P_m 的调整动态化。

最后值得指出的是，对 GA 收敛性及收敛速度的分析是进一步研究 GA 的一个重要方向，这将是指导我们改进 GA 的有力依据。

参　考　文　献

[1] Rudolph G. IEEE Trans. on NN, 1994, NN-5(1): 96–101.

[2] Holland J H. Adaptation in Natural and Artificial Systems, ANN Arbor: The Univ. of Michigan Press, 1975.

[3] 陈根社，陈新梅. 信息与控制，1994, 23(4): 215–222.

[4] Janson D J, Frenzel J F. IEEE Expert, 1993, 8(5): 26–33.

[5] Yasumasa I, Hiroaki K. IEICE Trans. Fundamentals, 1994, E77-A(4): 731–735.

[6] Tomoharu N, Takeshi A. IEICE Trans. Information & System, 1993, E76-D(6): 689–697.

[7] Zhang Liangjie, Mao Zhihong, Li Yanda. An Improved Genetic Algorithm Based on Combinative Theory & Fuzzy Reasoning and its Applications. International Conference on Neural Information Processing, Korea: 1994, 180–185.

[8] Hung S L, Adeli H. IEEE Trans on NN, 1994, NN-5(6): 900–909.

[9] Grefenstette J J. IEEE Trans. on SMC, 1986, SMC-16(1): 122–128.

[10] Goldberg D E. Genetic Algorithms in Search, Optimization and Learning. Addison Wesley Publishing Company, 1989, Chapter 4 and Chapter 5.

MATHEMATICAL ANALYSIS OF MUTATION OPERATOR AND ITS IMPROVED STRATEGY IN GENETIC ALGORITHMS

Zhang Liangjie　Mao Zhihong　Li Yanda

(*Department of Automation, Tsinghua University, Beijing* 100084)

Abstract　This paper analyzes the optimization problem of mutation probability (P_m) in genetic algorithms by applying the definition of i-bit improved sub-space. Then fuzzy reasoning technique is adopted to determine the optimal mutation probability in different conditions. The superior convergence property of the new method is evaluated by applying it to two simulation examples.

Key words　Genetic algorithm(GA), i-bit improved sub-space, Fuzzy reasoning

张良杰：　男，1969 年生，博士生，现从事模糊神经网络技术在 ATM 网络控制中的应用研究.

毛志宏：　男，1972 年生，本科生，现从事模糊神经网络技术的研究.

李衍达：　男，1936 年生，中科院院士，教授、博士生导师，现从事网络信息处理及智能控制等方面的研究.