

基于 SDNA-GA 优化的模糊神经网络控制

宋春宁, 刘少东

(广西大学 电气工程学院, 南宁 530004)

摘要: 针对最新的生物 DNA 研究, 病毒中同一 DNA 碱基顺序可以编码出 2 条或者 3 条不同的多肽链. 在此基础上分析与模仿了重叠基因和重叠密码的机理, 得到一种新的基于重叠基因编码框架, 从而提高了问题求解的效率; 同时, 得到一种移码解读框架的 DNA 遗传算法 (SDNA-GA) 计算模型, 并将其应用于一类广义隶属度型 T-S 模糊神经网络控制器 (GTS-FNNC) 的优化设计, 实现了 GTS-FNNC 的在线学习.

关键词: 重叠基因编码框架; 自适应变异; SDNA 编码框架; 遗传算法; T-S 模糊神经网络

中图分类号: TP18

文献标志码: A

Optimal design of fuzzy neural network controller based on SDNA genetic algorithm

SONG Chun-ning, LIU Shao-dong

(College of Electrical Engineering, Guangxi University, Nanning 530004, China. Correspondent: SONG Chun-ning, E-mail: 914579661@qq.com)

Abstract: According to the latest biological research of DNA, the virus in the same DNA base sequence can encode two or three different polypeptide chains. Further analysis and imitation of overlapping genes and password mechanism are inspired by using the above mechanism, and a new overlapping gene encoding framework is found, thereby improving the efficiency of problem solving. A frame shift interpretation framework based on this mechanism of DNA genetic algorithm (SDNA-GA) computing model is applied to a class of the generalized membership type T-S fuzzy neural network controller (GTS-FNNC) optimization design to achieve GTS-FNNC online learning.

Key words: overlapping gene encoding framework; adaptive mutation; SDNA encoding framework; genetic algorithm; T-S fuzzy network

0 引言

模糊控制在许多实际工程中越来越受到人们的关注, 但是模糊控制需要的设计参数众多, 尤其是复杂系统的设计解析过程比较困难, 不能得到满意的模糊规则. 采用 DNA 编码方法来表达模糊系统隶属度函数中的设计参数及模糊规则, 是解决模糊系统优化问题的主要途径. DNA 编码质量的好坏直接影响优化算法的速度和效率, 编码序列优化采用的主要方法有神经网络方法^[1]、粒子群遗传算法^[2-3]、混沌遗传算法^[4]等. 由于神经网络容易陷入局部最优, 采用神经网络设计的模糊控制器往往不是全局最优. 遗传算法虽然具有全局搜索能力, 但在局部搜索解空间时不是很有效, 并且对于 0-1 编码模型的遗传操作, 染色体长度较长, 计算速度慢. 文献 [5] 提出了重叠基因的 DNA

编码, 用于解决模糊系统的优化问题. 本文重点分析基于重叠基因和重叠密码的机理, 设计一种移码解读框架的 DNA 遗传算法 (SDNA-GA) 的模型, 并在广义隶属度型 T-S 模糊神经网络控制器 (GTS-FNNC) 的优化设计中进行验证, 以提高搜索效率.

1 DNA 编码模糊控制器的控制规则学习

DNA 编码模型每个染色体由 4 种碱基 A、T、C、G 的结合体构成, 可以表示多个基因. 每 3 个连续碱基组成的密码子表示待解决问题的某个设计参数.

图 1 给出了采用 DNA 编码方法得到的一个带有重叠基因的染色体, 其中基因 2 重叠于基因 1 和基因 3 中, 每个碱基都起始于 ATG, 终止于 TAG.

重叠 DNA 编码模型的遗传操作算子如下:

收稿日期: 2012-12-17; 修回日期: 2013-04-01.

基金项目: 国家自然科学基金项目(60924002).

作者简介: 宋春宁(1969—), 男, 副教授, 从事智能控制、嵌入式控制等研究; 刘少东(1984—), 男, 硕士生, 从事智能控制的研究.

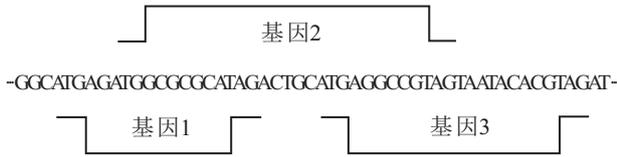


图 1 基因重叠

1) 交叉. 对于被选中的用于繁殖的每一对 DNA 链染色体, 将其部分内容进行互换. 通过交叉, 凭借交叉点产生新的基因, 使基因得到极大的改变.

2) 变异. 染色体中 DNA 链的变化有碱基的替换、丢失和嵌入. 在 DNA 编码模型中, 也采用这些变异. 图 2 是染色体中的 1 个碱基由 G→C 的变异例子, 基因 2 变成了基因 2', 并且产生了新的终止密码子 TAG, 从而基因 3 消失.

3) 倒位. 染色体中两个随机选择的位置之间的某些碱基序列的顺序发生了倒位.

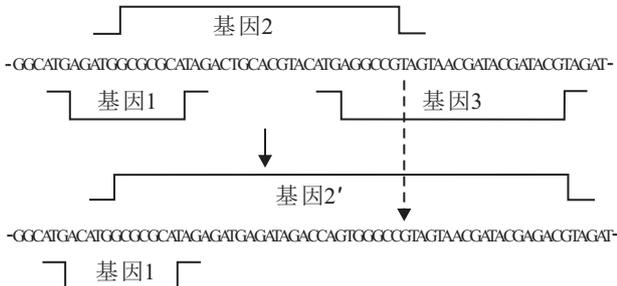


图 2 变异

2 广义隶属型 T-S 模糊神经网络控制器

将重叠 DNA 编码方式应用于 GTS-FNNC 的优化设计, 实现了该网络的在线学习, 增强了网络的自适应能力.

T-S 模糊控制器的两个输入变量为偏差 $e(n)$ 和偏差变化率 $r(n)$. 系统的设定输出为 $SP(n)$, 系统的实际输出为 $y(n)$. T-S 模糊控制器采用以下带有线性后项形式的 T-S 模糊规则:

If $e(n)$ is E_i and $r(n)$ is R_j Then

$$\Delta u(n) = p_{i,j}e(n) + q_{i,j}r(n). \quad (1)$$

其中: $\Delta u(n)$ 是这条规则对 T-S 模糊控制器输出变化的贡献, $p_{i,j}$ 和 $q_{i,j}$ 是规则后项设计参数. 两个输入变量分别被 N_1 和 N_2 个输入模糊集合模糊化, 用于模糊化 $e(n)$ 或 $r(n)$ 的模糊集表示为 E_i 或 $R_j, i = 1, 2, \dots, N_1, j = 1, 2, \dots, N_2$, 其对应的隶属度函数为 $\mu_i(e)$ 或者 $\mu_i(r)$. $\mu_i(x)$ 的数学表达式为

$$\mu_i(x) = e^{-|a_i + xb_i|c_i}, x = e, r, \quad (2)$$

其中 a_i, b_i, c_i 为设计参数. 通过选择合适的设计参数 a_i, b_i, c_i , 式 (2) 逼近三角形、梯形和高斯型等隶属度函数. 每条规则的适用度采用 Zadeh 模糊逻辑 AND

操作, 且规则后项对应输出的隶属值为

$$\mu_{i,j}(\Delta u) = \min(\mu_i(e), \mu_j(r)), \quad (3)$$

重心反模糊器的输出为

$$\Delta u(n) = \frac{\sum_{i=1}^{J_1} \sum_{j=1}^{J_2} \mu_{i,j}(\Delta \mu) * (p_{i,j}e(n) + q_{i,j}r(n))}{\sum_{i=1}^{J_1} \sum_{j=1}^{J_2} \mu_{i,j}(\Delta \mu)} = k_i e(n) + k_p r(n), \quad (4)$$

T-S 模糊控制器在 n 时刻的输出为

$$u(n) = u(n - 1) + \Delta u(n). \quad (5)$$

T-S 模糊控制器是相当于带有可变比例增益 k_i 和可变积分增益 k_p 的 PI 控制器. T-S 模糊控制器 (4) 可根据连接主义思想用神经网络^[6-7]来构造, 见图 3. 其中: 第 1 层为输入变量, 第 2 层为用于模糊化的广义隶属函数, 第 3 层为 Zadeh 模糊逻辑 AND 操作, 第 4 层为乘积运算层, 第 5 层为求和运算层, 第 6 层为除法运算层. 通过各层运算, 图 3 所示的运算结果与式 (4) 一致. T-S 模糊控制器 (4) 是采用广义隶属函数 (2), 并利用神经网络构造的, 故称 T-S 模糊控制器为 GTS-FNNC. GTS-FNNC 作为非线性控制器, 在处理非线性系统时比线性控制器更有效, 但隶属度函数和规则后项的设计参数较多, 难以解析地设计, 因此提出采用重叠基因编码和移码解读框架算法, 用来优化设计 T-S 模糊神经网络系统.

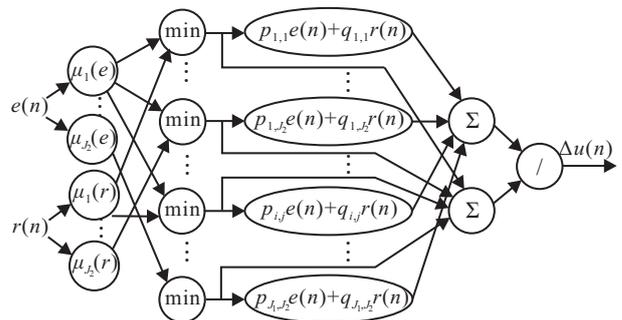


图 3 广义隶属型 T-S 模糊神经网络控制器结构

3 基于重叠基因的 DNA 遗传算法

SDNA-GA 采用重叠基因编码和移码解读框架来编码 T-S 模糊规则, 并且运用模糊控制技术来实现自适应变异, 提高了效能.

3.1 生物重叠基因与重叠密码

在一些病毒中存在一种 DNA 碱基顺序可以编码出两条不同的多肽链, 这表明每个基因与另一基因共用一段相同的碱基顺序, 但解读框架不同. 图 4 表示的是两个基因的部分碱基顺序中密码子及其编码的氨基酸. 由于移动了一个碱基, 得到两条完全不同的氨基酸序列, 若再移动一个碱基, 则有 3 种解读框架,

编码出3种不同的氨基酸序列. 这种重叠基因情况在许多病毒DNA中都被发现^[8].

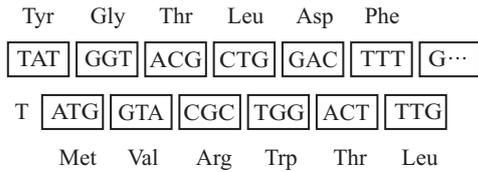


图4 基因共用相同的碱基顺序但解读框架不同

借用这种DNA重叠基因和移码解读框架来编码某一设计问题的参数, 可以增加DNA编码的效能, 减少每代的个体数, 并使SDNA-GA优化设计某一设计问题的搜索效率得到提高.

3.2 T-S模糊规则的重叠基因编码和移码解读框架

GTS-FNNC的设计参数可通过4字符集 $\sum\{A, G, C, T\}$ 来编码, 并且形成染色体. DNA码串对应T-S模糊规则库, 设计参数组成如图5所示^[5,9].

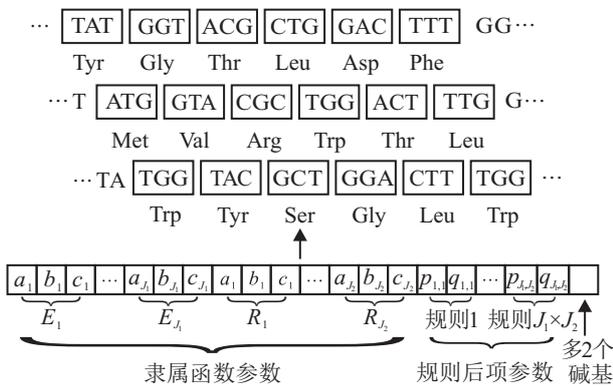


图5 对应T-S模糊规则库参数的移码解读框架

3.3 自适应变异操作

在SDNA-GA中, 除采用自适应变异操作外^[10], 其他操作算子的操作方式同DNA操作算子. 根据变异概率 P_m 对SDNA-GA性能影响的关系, 考虑用自适应变异概率 P_m 对DNA群体 P_{t+1} 随机地进行变异, 对于被选中的DNA链个体, 随机地选取某位进行DNA链中碱基序列的变化.

变异概率 P_m 的选取对SDNA-GA的性能影响较大^[11], 在SDNA-GA随机搜索的过程中, 在接近最优解的邻域时, P_m 应取小, 否则接近最优解的积木块会因变异而遭到破坏; 为防止过早地出现收敛现象, 维持群体的多样性, 在未达到最优解附近时, P_m 应取大. 为使 P_m 随群体中个体的多样性程度而自适应地调整, 本文根据 P_t 代和 P_{t+1} 代所有个体的平均适合度差值的大小, 采用模糊逻辑技术自适应地调整变异概率 P_m . 定义 P_t 代和 P_{t+1} 代所有个体的平均合适度的差值为

$$F(P_t, P_{t+1}) = |\bar{F}_{fit}(P_t) - \bar{F}_{fit}(P_{t+1})|, \quad (6)$$

其中 $\bar{F}_{fit}(P_t)$ 和 $\bar{F}_{fit}(P_{t+1})$ 分别为 P_t 代和 P_{t+1} 代所有个体的平均合适度. 模糊规则系统采用两条模糊规则

$$\begin{cases} \text{If } F(P_t, P_{t+1}) \text{ is Big Then } P_m \text{ is Small,} \\ \text{If } F(P_t, P_{t+1}) \text{ is Small Then } P_m \text{ is Big.} \end{cases}$$

Big和Small如图6所示, 用以描述 $F(P_t, P_{t+1})$ 和 P_m 的变化. 图中最大的变异率为0.1, 最大平均合适度为 F_m .

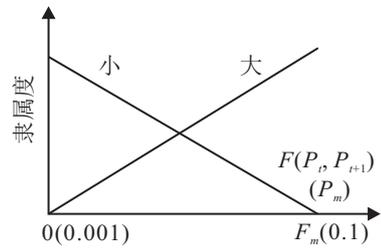


图6 $F(P_t, P_{t+1})$ 和 P_m 的隶属函数

3.4 仿真研究

本节采用SDNA-GA优化设计GTS-FNNC控制器, 用于控制一个非线性对象, 并验证其有效性. 非线性对象的模型描述如下:

$$y(k) = 0.79y(k-1) + 0.012y(k-1)y(k-2) - 0.005y^2(k-2) + 0.15u(k) + 0.8u(k-1). \quad (7)$$

GTS-FNNC的结构设计为: $e(n)$ 和 $r(n)$ 均被5个广义隶属函数型输入模糊集模糊化, Zadeh模糊逻辑AND和重心解模糊法^[12]用于计算GTS-FNNC的输出. GTS-FNNC共有25条模糊规则, 需要设计的参数有输入模糊集参数30个, 规则后项参数50个, 共80个参数. 由于用一个DNA密码子来编码一个设计参数, 并且为了移码解读DNA链, 需要增加两个碱基, 共需要采用242位DNA码串. SDNA-GA中, 每代个体为30, 交叉率为0.9, 初始变异率为0.01.

仿真前, 先通过经验公式及人工调整方法设计PI控制器, PI控制器的比例增益为0.1555, 积分增益为0.04. 然后根据GTS-FNNC和PI控制器增益的关系确定后项参数的变化范围, 即 $[q_{i,j}^{\min}, q_{i,j}^{\max}] = [0, 0.3]$, $[p_{i,j}^{\min}, p_{i,j}^{\max}] = [0, 0.06]$. 由此变化范围可以确定为从表转译后项设计参数的转化关系. 广义隶属函数中的设计参数的变化范围和转化关系式可通过计算机仿真得到. 在得到各个设计参数的变化范围后, 进行计算机仿真. 大量的仿真结果表明, 经过若干代优化, 总能得到用于控制模型的GTS-FNNC的一组优化的设计参数. 仿真时还发现, 若初始个体较好, 进化的代数可大大减少. 图7为PI控制器与采用SDNA-GA优化的GTS-FNNC控制器的仿真比较, 虚线为GTS-FNNC控制器的仿真曲线, 实线为PI控制器仿真曲线. 从图7可以看出, 经SDNA-GA优化设计的GTS-

FNNC的控制性能比PI控制器优越.表1~表3给出了SDNA-GA优化得到的GTS-FNNC的隶属函数、设计参数和规则后项参数的值.在SDNA-GA中,得到优化解时的优化代数较少,搜索效率提高,并且每代采用的个体数较少.

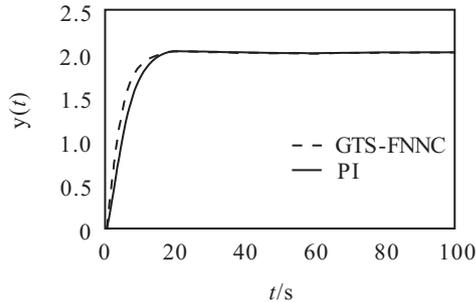


图7 GTS-FNNC控制器和PI控制器的仿真比较

表1 隶属函数 $\mu_i(r)$ 与设计参数的关系

设计参数	$\mu_1(r)$	$\mu_2(r)$	$\mu_3(r)$	$\mu_4(r)$	$\mu_5(r)$
a_i	3.5000	-4.0000	-3.5000	3.0000	2.0000
b_i	0	1.3500	-0.9000	0.1500	0.4500
c_i	6.0000	5.5566	6.0000	5.5566	2.8889

表2 隶属函数 $\mu_i(e)$ 与设计参数的关系

设计参数	$\mu_1(e)$	$\mu_2(e)$	$\mu_3(e)$
a_i	-3.5000	-1.2000	3.3333
b_i	-0.5000	0	3.3333
c_i	-4.5000	1.2000	2.8889

表3 经SDNA-GA优化的GTS-FNNC规则后项参数

隶属函数	$\mu_1(r)$	$\mu_2(r)$	$\mu_3(r)$	$\mu_4(r)$	$\mu_5(r)$
$\mu_1(e)$	0.0067	0	0.0033	0.0067	0.0067
	(0.0067)	(0.1000)	(0.0250)	(0.1125)	(0.0250)
$\mu_2(e)$	0.0200	0.0133	0.0100	0.0267	0.0467
	(0.0375)	(0.0250)	(0.2000)	(0.0165)	(0.1000)
$\mu_3(e)$	0.0067	0.0033	0.0567	0.0167	0.0367
	(0.0500)	(0.1750)	(0.0375)	(0.0250)	(0.0500)

注: 其中括号外的数字代表 $p_{i,j}$, 括号内的数字代表 $q_{i,j}$

4 结论

本文分析了生物的遗传机理和基因调控机理,构建了基于DNA编码染色体表达机制的模糊控制系统隶属函数和控制规则.采用一种新的允许基因重叠的DNA编码方法,讨论了交叉、变异和倒位等遗传操作算子,探讨了如何采用重叠DNA编码来表达模糊系统隶属函数的设计参数及模糊规则.在此基础上得到了SDNA-GA的计算模型,并将其应用于一类GTS-FNNC的优化设计,实现了GTS-FNNC的在线学习.该算法具有自适应特点的杂交概率和变异概率,使得遗传搜索的速度和效率得到显著改善.

参考文献(References)

[1] 李银山, 杨春燕, 张伟. DNA序列分类的神经网络方法[J]. 计算机仿真, 2003, 30(2): 65-68.
(Li Y S, Yang C Y, Zhang W. The neural network

method of classifications for DNA sequences[J]. Computer Simulation, 2003, 30(2): 65-68.)

- [2] 崔光照, 李小广, 张勋才, 等. 基于改进的粒子群遗传算法的DNA编码序列优化[J]. 计算机学报, 2010, 33(2): 311-316.
(Cui G Z, Li X G, Zhang X C, et al. The optimization of DNA encodings based on modified PSO/GA algorithm[J]. Chinese J of Computers, 2010, 33(2): 311-316.)
- [3] 许世明, 张强. 基于遗传粒子群算法的DNA编码优化[J]. 计算机工程, 2008, 34(1): 218-220.
(Xu S M, Zhang Q. Optimization of DNA coding based on GA/PSO algorithm[J]. Computer Engineering, 2008, 34(1): 218-220.)
- [4] 陈霄, 王宁. 基于混沌DNA遗传算法的模糊递归神经网络建模[J]. 控制理论与应用, 2011, 28(11): 1589-1594.
(Chen X, Wang N. Fuzzy recurrent neural network modeling based on chaos DNA genetic algorithm[J]. Control Theory & Applications, 2011, 28(11): 1589-1594.)
- [5] 任立红, 丁永生. 一种采用重叠基因的DNA编码方法[C]. 中国控制与决策学术会议. 郑州, 2002: 472-475.
(Ren L H, Ding Y S. A DNA encoding method using overlapping genes[C]. Chinese Control and Decision Conf. Zhengzhou, 2002: 472-475.)
- [6] 张庆灵, 朱宝彦. T-S模糊广义系统的分析与控制[M]. 北京: 国防工业出版社, 2011: 12-15.
(Zhang Q L, Zhu B Y. The analysis and control of T-S fuzzy singular systems[M]. Beijing: National Defense Industry Press, 2011: 12-15.)
- [7] Ren L H, Ding Y S. Design of fuzzy control system by a new DNA-based immune genetic algorithm[C]. Proc of the 10th IEEE Int Conf on Fuzzy Systems. Melbourne, 2008: 244-247.
- [8] Xutao Deng, Jun Xu. Multiplicity adjustment for intersection-union test: Detecting overlapping genes from multiple microarray gene lists[C]. Computer and Computational Sciences, Second International Multi-Symposiums on. Iowa City, 2007: 52-59.
- [9] Jang J-S R. Neuro-fuzzy model and control[J]. IEEE, 1955, 83(3): 378-406.
- [10] Shin S Y, Lee I H. Multiobjective evolutionary optimization of DNA sequences for reliable DNA computing[J]. IEEE Trans on Evolutionary Computation, 2005, 9(2): 143-158.
- [11] 丁永生. 计算机智能-理论、技术与应用[M]. 北京: 科学出版社, 2004.
(Ding Y S. Computer intelligence - theory, technology and application[M]. Beijing: Science Press, 2004.)
- [12] Takagi T, Sugeno M. Fuzzy identification of systems and its application to modeling and control[J]. IEEE Trans Systems, Man, Cybernetics, 1985, 15(1): 116-132.

(责任编辑: 齐 霖)