

基于三阶染色体结构的自适应进化规划及应用

陈得宝¹, 刘墩东², 魏 辉³(¹ 淮北煤炭师范学院物理系, 安徽 淮北 235000; ² 厦门大学自动化系, 厦门 361005; ³ 淮北市朔里职业高中, 安徽 淮北 235000)

摘 要: 针对模糊辨识中规则数多少难以选择、输入量在各规则中重要程度难以确定以及模糊规则参数难以优化等问题, 模拟生物基因之间相互联系的特点, 用三阶染色体结构对“完全”模糊规则基进行描述。根据每代不同个体的适应度信息获得不同个体的变异率, 采用自适应进化规划算法, 实现模糊规则基各级染色体引导优化, 克服了用试凑方法确定参数对辨识精度的消极影响, 算法的复杂程度大大减小, 为模糊建模提供了一种新颖的方案。通过对两类不同非线性系统的仿真, 证明了此方法的有效性。

关键词: 模糊规则基; 自适应进化规划; 模糊建模; 非线性系统

文章编号: 1004-731X (2004) 02-0322-04

中图分类号: TP13

文献标识码: A

Self-adaptive Evolutionary-programming Based on Three Layer Chromosome and Application

CHEN De-bao¹, LIU Tun-dong², WEI Hui³

¹Physical Department of Huaibei Coal Industry Teachers college, Anhui 235000, China;

²Department of Automation, Xiamen University, Xiamen 361005, China; ³ShuoLi High School of Huaibei, Anhui 235000, China)

Abstract: A new method which imitates features of relating in biology's gene is proposed in the paper, and it can solve the problems of fuzzy identify in nonlinear system, such as difficulty in choosing number of rules, weight of input variable in different rules and difficult in optimizing parameters of fuzzy rule bases, etc. In the paper, a three layers chromosome is used to describe perfect rule bases. Different to old methods, an adaptive evolutionary algorithm is used to realize guiding optimum based on the fitness information of different individual. Moreover, the whole chromosome is optimized simultaneously. The new method remedies the bad effect of compromise proposal in identifying of old method and the given algorithm is simple. The effectiveness of the given method is demonstrated by simulations of different nonlinear systems.

Keywords: fuzzy rule bases; self-adaptive evolutionary programming; fuzzy modeling; nonlinear system

引 言

非线性系统建模问题是近些年来研究的热点,特别是智能方法的出现,为非线性系统模型的建立提供了新思路。在 Narendra 和 Parthasarath 的获奖论文[1]中,他们用反向传播神经网络作为动态系统非线性部分的辨识器,但随着研究的不断深入,神经网络的局限性(如收敛速度慢,易陷入局部极小等)限制了它的应用。随着模糊理论不断成熟,模糊辨识得到了广泛的应用,它与神经网络相比具有两大明显的优点,一是模糊辨识的参数具有明确的物理意义,二是模糊辨识提供利用人类语言描述信息的途径。

模糊辨识中,语言模糊规则数的多少、输入量在各规则中的重要程度以及规则基参数是否优化直接影响算法的收敛速度和辨识效果。传统的模糊辨识将规则基的前件参数和后件参数分开辨识,算法复杂,收敛速度慢。近些年来,改善模糊辨识效果的方法很多,如递阶模糊辨识^[2]、模糊满意聚类建模等,但这些方法对模糊规则数的选取、输入量在模

糊规则中的重要度确定都采用试凑法,这些都影响到辨识的速度和精度。在规则基参数的优化方面,遗传算法在已知规则基结构的前提下优化规则基参数中发挥了重要的作用,但对个体结构不一样的规则基,交叉操作变成不可实现。进化规划仅用变异作为基因改变的唯一操作,在解决无先验知识的问题中发挥了重要的作用,但传统的进化规划仅对方差进行处理,且变异方法单一,严重影响了算法的收敛速度。本文在对传统进化规划改进的基础上,提出利用自适应方法对三阶染色体结构的语言模糊规则基进行引导优化,达到对规则基结构、重要度和参数同时优化的目的,大大改善了建模效果,为非线性系统建模提供了一种新颖的方法,特别在热能表系数模型中的应用,证明了此方法不仅具有很强的理论研究价值而且还具有很强的实用性。

1 用三阶染色体构建“完全”语言模糊规则基原理

模拟生物基因之间相互联系和相互作用的特点,结合“完全”语言模糊规则基的三大基本要求,构建三阶染色体形式的语言模糊规则基如图1,第一级称为控制级,通常用二进制“0”或“1”编码,控制级中“1”的个数为有效规则数,第二级称为重要度级,决定着输入变量在规则中的重要度,选择第二级基因的原因见本文3.3分析。第三级称为参数级,

收稿日期: 2002-11-08

修回日期: 2003-05-23

基金项目: 安徽省高等学校青年教师资助计划项目(2002jq148)

作者简介: 陈得宝(1975-),男,安徽安庆人,工学硕士,研究方向为人工智能、进化计算等;魏 辉(1971-),男,安徽淮北人,讲师,研究方向为计算机控制技术。

决定着规则基的形状, 本文采用与语言更接近的不等腰三角形形式。三阶染色体相互协调工作原理为: 当控制级某位基因为“0”时, 表示其对应的第二级、每三级基因处于休眠状态, 当控制级基因某位为“1”时, 表示与其对应的第二级、第三级基因处于激活状态, 对输出产生作用。这样设计的染色体包含更多信息, 比传统方法能处理更复杂的问题。

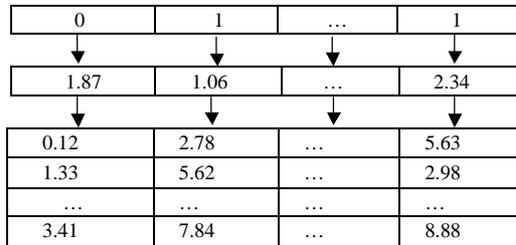


图1 三阶染色体构建模糊规则基图

2 语言模糊模型及原理

传统的模糊方法中大都包括以下几个部分(1)模糊量化(2)模糊处理(3)清晰化处理。算法复杂。本文采用不等腰三角形隶属度函数形式, 与实际更加接近, 具体原理如下:

考虑如下的多入单出模糊系统:

R_i : if u_1 is A_1^i with M_1^i and .. u_k is A_k^i with M_k^i then y is O^i

其中: R_i 为第 i 条规则, $u_j (i <= j <= k)$ 为输入变量, y 为输出变量, M_k^i 为输入变量在第 i 条规则中的重要度, O^i 为单点模糊输出, $A_j^i (1 <= j <= k)$ 为隶属度函数, 其具体形式如图 2 式 1 中个参量的意义如图 2。

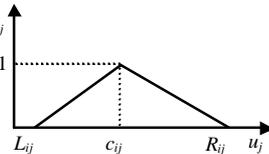


图2 式1参数意义图

$$A_j^i = \begin{cases} (u_j + L_{ij}) / (c_{ij} - L_{ij}), & \text{if } L_{ij} \leq u_j < c_{ij} \\ (-u_j + R_{ij}) / (R_{ij} - c_{ij}), & \text{if } (c_{ij} \leq u_j \leq R_{ij}) \\ 0, & \text{others} \end{cases} \quad (1)$$

对给定的 n 组输入输出数据 $(x_{n1}, x_{n2}, \dots, x_{nk}, y_k)$, 按式 2 计算满意度:

$$W_i = C^i * A_1^i * M_1^i * A_2^i * M_2^i * \dots * A_k^i * M_k^i \quad (2)$$

其中 C^i 为控制级基因。

采用加权平均法按式 3 计算模糊模型的输出为^[3]:

$$y_k^* = \sum_{i=1}^r W_i * O^i / \sum_{i=1}^r W_i \quad (3)$$

其中 r 为规则数, 由分析知: 只要模糊规则基合适, 则模糊模型与实际模型能无限逼近。

3 自适应进化规划算法设计

进化规划算法以其仅用变异作为基因改变的唯一方法, 在解决无先验知识或先验知识不完备的问题中发挥了重要作用^[4], 但传统方法处理的仅为 χ^2 问题且新生代群体的产生是父代群体所有个体进行变异和选择的结果, 运算量很大。型式定理为分析进化算法提供了理论依据^[5,6], 为自适应进化规划提供了理论保证。本文对传统进化规则进行改

进, 采用自适应方法, 将计算量减小了一半。

3.1 个体的表示

以一个两输入单输出系统为例, 假设初始规则数为 r , 则个体的形式如图 3。

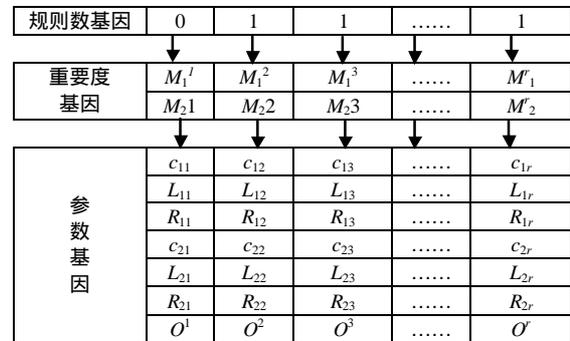


图3 个体基因表示形式

3.2 群体规模的选择

群体规模的大小, 对算法的收敛速度影响较大, 过小的群体规模使算法不收敛, 过大的群体规模使算法收敛速度变慢。通常群体规模选为 70~100 之间。

3.3 群体的初始化

针对三阶染色体的特点, 用三个数组对群体进行初始化, 控制级数组 A, 采用二进制“0”或“1”编码, 随机初始化为 $1 \times r$ 的数组, 其中 r 为初始假定规则数。因为输入变量的“重要度”不依赖于模糊模型的结构, 传统方法一般选取重要度在 2~4 之间, 这有一定盲目性; 另一方面, “重要度基因”与“参数基因”差距很大, 参数基因针对不同的输入有不同的初始化方法, 这正是将“重要度基因”与“参数基因”分为两级的原因。鉴于以上分析, “重要度基因”采用无需译码的十进制数初始化数组 B, B 维数为: 输入个数 \times 规则数, 初始化方法为 $B = R * [\text{random}(32768)/32767.0]$, R 为初始化系数, 通常取为 1 到 5 之间, 本文取 $R = 4.0$ 。参数数组 C 采用十进制数, 根据不同输入而定, 形式为 $C = \text{MaxInput} * \text{random}(32768)/32767.0$, MaxInput 为输入最大值。

3.4 适应度计算

为使产生的模糊语言模型更好地逼近实际模型, 定义误差函数为: $E = \frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2$

式中: n 为输入输出数据对的个数, y_i 为实际输出, y_i^* 为语言模糊模型输出, 当 E 满足要求时, 认为模糊模型最优, 根据分析, 定义适应度函数如下: $F = 1.0/E$, 为防止变异过程中 $L_{ij} > c_{ij}$ 或 $c_{ij} > R_{ij}$ 现象的发生, 采用作者在文献[3]中使用的惩罚函数来限制误收敛。

3.5 选择过程

利用 q 竞争方法, 首先计算群体的适应度, 选择得分高

的 50% 个个体，复制到下一代，再根据不同个体的适应度，计算个体的变异率，方法为： $p_m = k(f_{\max} - f) / (f_{\max} - f_{\text{av}})$ ，其中， p_m 为变异率， k 为变异系数，一般取 0.1~0.5， f_{\max} ， f ， f_{av} 分别为当代群体的最大适应度、个体适应度、平均适应度。

3.6 变异

针对三阶染色体的模糊模型，变异方法为：

1) 控制级变异，按变异率，若选择到的控制级基因某位为“0”，则变为“1”，否则，变为“0”。

2) 重要度级与参数级，按不同的变异率，随机选择三种变异算子中的一种进行变异操作。a. 互换同一个体同行上的两元素；b. 互换同一个体同列上的两元素；c. 按指数规律变异，通过增加高斯随机数，按式 4 进行变异。

$$c_{ij} = c_{ij} + \Delta c_{ij} \tag{4}$$

$$\Delta c_{ij} = a_p * \exp[(f_{\max} - f_k) / f_{\max}] * g_{ij}$$

式中 c_{ij} 为参数改变量， a_p 为变异系数，一般取 0 到 0.5 之间的常数， f_{\max} ， f_k 分别为该代中最大适应度和第 k 个个体适应度， g_{ij} 为均值 0，方差 1 的高斯随机数。

按上述步骤反复进行，若不满足要求，则返还到 3.4，直到满足要求为止。

	0	1	1	1	1	0	0	1	0	0	1	0
M^i	4.717495	2.130023	4.890744	2.964263	1.289265	4.040064	0.971953	1.613680	0.475252	0.361380	1.976608	2.369016
c_{ij}	1.998265	0.036325	0.930387	2.073913	0.574021	4.279372	4.393854	0.546430	3.566600	4.046705	2.578764	5.105167
L_{ij}	-0.809980	-2.991110	-2.82894	1.497418	-0.015801	1.915154	3.689652	1.339881	3.486238	2.38763	0.275498	0.298002
R_{ij}	2.575692	2.744680	1.888516	3.799821	4.161769	8.162842	4.968956	3.120787	6.335311	8.672916	7.194670	9.020856
O^i	2.608713	2.338766	0.387733	4.014066	6.405387	4.485421	4.597616	2.772556	0.721347	1.509042	4.692167	3.280587

图 5 $y=x^2$ 的最优语言模糊规则基

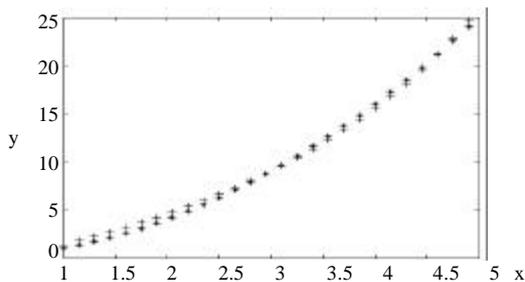


图 6 模糊模型与实际模型比较曲线
“+” —fuzzy model “*” —real model

表 1 训练结果对比

训练代数(代)	最优个体的误差		个体的最大适应度	
本文方法	文献[3]方法	本文方法	文献[3]方法	本文方法
948	3683	0.0659	0.0908	15.15419
				11.012

4.2 热能表系数模型的建立

热能表系数准确与否，直接决定着测量精度，本文在进水温度 (20 ~ 160)，出水温度 (10 ~ 150) 且进水温度大于出水温度的条件下每隔 10 均匀选取 105 组样本，

4 仿真实验

4.1 对单入单出系统的仿真

$$y=x^2 \quad (1 \leq x \leq 5)$$

训练的初始参数为：25 组数据对，初始规定最大代数为 1000，群体规模 100， $p=0.06$ ， $q=70$ ，变异率求取系数 $k=0.4$ ， $maxinput=5.0$ ，初始规则数为 12，训练结果如表 1，表中同时给出了文献[3]的训练结果，从表 1 可看出本文方法具有更好的效果。训练过程的误差曲线如图 4，第一代最好个体的误差为 67.937454。最优模糊模型如图 5，为验证模糊模型的泛化能力， x 在 1 到 5 之间每隔 0.15 均匀取值，模糊模型的输出与实际模型输出比较如图 6。

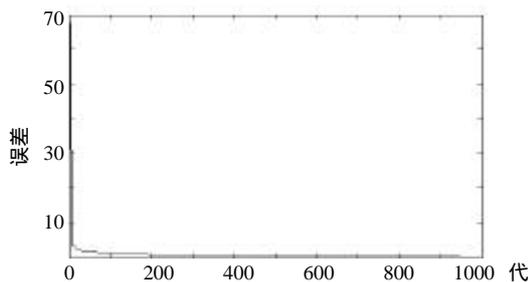


图 4 训练过程的误差曲线

见附表 2，初始规定最大代数为 2000，群体规模 100， $p=0.1$ ， $q=70$ ，变异率求取系数 $k=0.04$ ， $maxinput=160.0$ ，初始规则数为 12，经 1520 代训练后，最优个体的误差为：0.016382，最优个体适应度为：61.042612，训练过程的误差曲线如图 7，第一代最好个体的误差为 634.650721。最优模糊模型结构如图 8，为验证模糊模型的泛化能力，在满足进水温度 (20 ~ 160) 出水温度 (10 ~ 150) 且进水温度大于出水温度的条件下，随机采取不同于样本的 50 个输入，模糊模型的输出与实际模型输出比较如图 9，(为观察清晰，图中将纵坐标进行放大，间隔为 0.02)，50 组数据对比中模糊模型与实

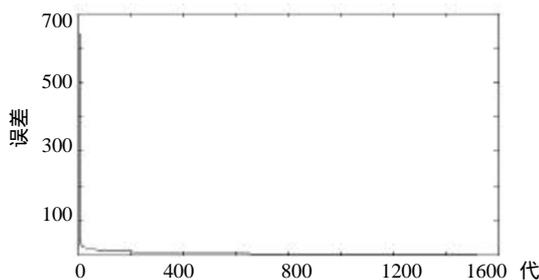


图 7 训练过程的误差曲线

	1	0	0	1	1	1	0	1	1	0	0	1
M_1^i	0.165502	1.382412	0.487643	4.183874	2.689562	3.449780	1.586971	0.100144	2.684903	0.165895	1.254032	0.305749
M_2^i	0.152903	2.439633	4.153234	2.569462	3.878924	0.054338	2.365382	3.408752	1.085439	1.934208	2.489321	3.290702
c_{ij}^1	16.686510	43.289760	121.379824	68.355191	40.562401	3.121182	25.567810	123.498825	141.468308	97.335681	29.158762	100.822090
L_{ij}^1	-52.951902	33.539772	47.635475	83.983463	-25.643602	-101.899143	13.289754	80.635155	30.629527	79.887643	18.293976	51.974914
R_{ij}^1	51.740324	128.347910	188.764252	130.375651	144.527305	150.248702	173.243827	248.470902	199.040207	187.86332	63.240815	180.639805
c_{ij}^2	2.058837	26.687642	74.768744	22.516220	82.261724	153.126905	96.53255	47.520008	56.858791	15.487670	46.688784	86.299789
L_{ij}^2	135.037439	0.057864	24.389681	22.515312	-63.230047	140.763414	59.008742	-95.443907	50.412399	0.974806	37.339673	-12.269814
R_{ij}^2	30.047191	134.467520	109.243905	54.122184	169.532642	199.915507	176.493227	165.809901	124.624826	138.867807	186.282552	134.305305
O^1	93.606381	18.87.987	28.387603	1.227573	41.257091	51.996258	128.777622	105.077334	16.538145	5.690783	78.770892	98.987068

图 8 热能表系数模型的最优语言模糊规则基

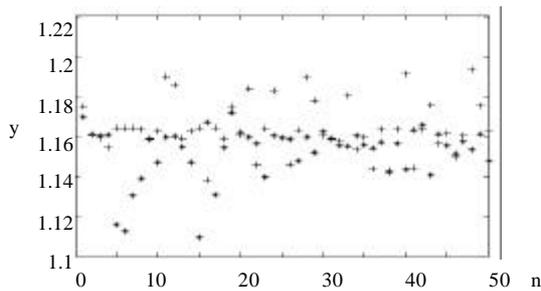


图 9 模糊模型与实际模型比较曲线
“+”——fuzzy model “*”——real model

际模型输出的最大误差为 0.0547, 平均误差为 0.011826。

5 结论

本文采用三阶染色体结构对语言模糊规则基进行描述, 使染色体包含更多的系统信息, 解决了模糊辨识中规则数的选择, 输入在各条规则中重要度的选择及参数优化等问题, 克服了 GA 算法对不同结构的染色体交叉不可实现的弱点, 用自适应进化规划方法对语言模糊规则基进行优化, 与传统

进化规划相比大大减小了运算量, 算法速度大大加快, 通过对两类不同系统的仿真, 特别是热能表模型的建立, 证明了该方法是十分有效的, 有着重要的实用价值。

参考文献:

- [1] Narendra K S, Parthasarathy K. Identification and control of dynamic systems using neural networks [J]. IEEE Trans. on Neural Networks, 1990, 1(1): 4-27.
- [2] 张平安, 李人厚, 等. 复杂系统的递阶模糊辨识[J]. 控制理论与应用, 2002, 19(1): 99-102.
- [3] 王晓兰, 陈得宝, 等. 进化规划在挖掘最优语言模糊规则基中的应用[J]. 兰州大学学报, 2001, 37(2): 71-76.
- [4] Hwang H S. Automatic design of fuzzy rule bases for modeling and control using evolutionary programming [J]. IEEE Proc-control Theory Appl, 1999, 146(1): 1-15.
- [5] Ergezingir S, Thomsen E. An accelerated learning algorithm for multiplayer perceptrons: Optimization layer by layer [J]. IEEE Trans. on Neural Network. 1995, 6(1): 31-42.
- [6] 刘芳, 李人厚, 等. 基于自适应变异规则的一种有效的进化规则[J]. 控制与决策, 2002, 17(2): 148-150.

(上接第 312 页)

4 结论

本文描述了“浅海海底管线电缆检测与维修装置”智能综合操纵和动力定位半实物仿真系统的系统构成、系统功能和仿真结果。装置于 2002 年 9 月在长江上进行了水面航行试验, 在 11 月进行了潜浮试验。水面航行试验全面测试了智能综合操纵和动力定位系统显控台上舵桨控制单元控制功能、利用智能操纵面板进行的集中手操控制功能和自动控制功能中的定速航行、定向航行和定点控位功能。潜浮试验检验了利用压载水舱控制面板来实现装置的下潜和上浮功能, 试验进行的比较成功。通过江试验验证了控制系统设计的正确性、可实现性。

参考文献:

- [1] 王丰尧. 滑模变结构控制[M]. 北京: 机械工业出版社, 1998.
- [2] 陈厚泰. 潜艇操纵性[M]. 北京: 国防工业出版社, 1980.
- [3] 浅海海底管线电缆检测与维修装置操纵性试验报告[R]. 上海交通大学, 1999.
- [4] 张亮. 浅海海底管线电缆检测与维修装置阻力试验报告[R]. 哈尔滨工程大学, 2000.
- [5] 姜秀鹏, 丁福光, 施小成. 计算机并行接口在集散控制系统中的应用[J]. 微计算机信息, 2002.
- [6] 王宏健, 施小成, 边信黔. 遥控潜器工作母船动力定位系统[J]. 中国造船, 2001.
- [7] 丁福光, 王宏健, 施小成. 潜器母船动力定位自动控制系统[J]. 船舶工程, 2001.
- [8] 常宗虎, 施小成, 边信黔. 浅海海底管线电缆检测与维修装置对线控位控制方法研究[J]. 船舶工程, 2003.
- [9] ANTHONY J HEALEY. Multivariable sliding mode control for autonomous diving and steering of unmanned underwater vehicle [J]. IEEE Journal of oceanic engineering, 1993, 18(3).