

学习分类器在绩效寻优中的应用及其组织决策意义

冯彦杰, 王浣尘

(上海交通大学 管理学院, 上海 200054)

摘要: 首先简单介绍了学习分类器算法及其应用情况, 其次通过一个博弈问题比较了学习分类器、分类器、反馈系统及随机游走策略这 4 种比较常见的组织决策模式在绩效寻优方面的效率, 指出了学习分类器算法在解决复杂系统的绩效寻优方面具有比较高的效率, 对组织的决策模式具有很大的启发意义。

关键词: 寻优; 分类器; 遗传算法

中图分类号: F406 **文献标识码:** A

Application of learning classifier in performance optimization and decision meaning

FENG Yan-jie, WANG Huan-chen

(Management School, Shanghai Jiao Tong University, Shanghai 200054, China)

Abstract: Firstly the algorithm of learning classifier system and its application are introduced, then through a game question, a comparison is made of four optimizing strategy: learning classifier system, classifier system, feedback system and walking system. Finally, the conclusion is reached that the learning classifier system is the best one to solve optimizing questions of complexity system.

Key words: optimizing; classifier system; GA

1 引言 (Introduction)

学习分类器系统是在传统的分类器基础上, 将基于信任分配的学习机制和基于遗传的规则发现机制有机结合在一起的机器学习系统. 它通过简单单位串规则的学习、发现, 更有效地引导 agent 在一个随意环境中的行为. 国内目前可以见到的文献一般用于集成制造系统^[1]、智能机器人系统^[2]、模式识别等领域中. 但是国内文献中往往并不是以学习分类器的名称出现, 有称为分类器的进化算法的也有称为自适应的分类器算法等等, 但是国外的文献一般统一称为学习分类器系统 (learning classifier system, LCS), 并广泛应用到了除工程领域外的其他领域, 例如在组织管理领域等^[3,4].

有关学习分类器系统的工作原理请参考相关文献^[5], 本文不再赘述.

在复杂系统中, 往往由于问题空间太大, 绩效寻优往往是在很多条件未知的情况下进行的. 以下, 本文通过一个简单的战略对策问题, 对学习分类器在绩效寻优方面的效率进行说明, 从而从一个侧面说

明了建立学习型组织的意义.

2 问题的描述及基本假设 (Problem statement and basic assumptions)

这个战略对策问题包含两个参加者, 他们为了提高各自的绩效进行博弈. 在每一个回合每个参加者的支付是由他们各自采取的行动共同决定的, 每个人得到的利润是每个回合的支付值的和. 每个参加者都由一个人工适应 agent 来代表, 他们每次从包括学习分类器系统在内的 4 种适应性策略中选择一种.

博弈的目标是通过选择三位长的二进制字符串 (比如 101, 111, 000) 使支付最大化. 博弈者的支付函数如下:

如果两个参加者第一位数字之和为 0, 则

$$P_i = -0.1 * \sum_{j=2}^3 r_{ij}, \quad j = 2, 3, \quad i = 1, 2; \quad (1)$$

如果两个参加者第一位数字之和为 1, 则

① 如果 $r_{i1} = 1$, 则

$$P_i = 1 - 0.1 * \sum_{j=2}^3 r_{ij}; \quad (2)$$

② 如果 $r_{i1} = 0$, 则

$$P_i = -0.1 * \sum_{j=2}^3 r_{ij}; \quad (3)$$

如果两个参加者第一位数字之和为 2, 则

$$\begin{cases} P_i = \sum_{j=2}^3 r_{ij} / \sum_{q=1}^2 \sum_{k=2}^3 r_{qk} - 0.1 * \sum_{j=2}^3 r_{ij}, \\ j, k = 2, 3, i, q = 1, 2. \end{cases} \quad (4)$$

其中: P_i 为第 i 个参加者的利润; r_{ij} 为第 i 个参加者第 j 个数字的值; i, q 为参加者的编号; j, k 为位置编号.

这个对策问题的支付矩阵的一部分如表 1, 其余部分可以通过公式(1)~(4)推算出来. 在支付矩阵中, 取了后两位数字值之和, 从而, 101 和 110 这两个字符串的作用是相同的. 总体上, 两个博弈者具有 $2^6 = 64$ 种可能的字符串组合.

表 1 支付矩阵

Table 1 Payoffs matrix

二进制数字的		最后两位		支 付	
首位的值		数字的和			
A	B	A	B	A	B
1	1	0	1	0.00	0.90
1	1	1	1	0.40	0.40
1	1	2	1	0.47	0.23
1	1	0	2	0.00	0.80
1	1	1	2	0.23	0.47
1	1	2	2	0.30	0.30

111 是这个游戏中的零和策略. 但是因为比有比更高的收益, 实际中并不是每个博弈者都会主动选择这个策略的.

本文利用循环赛的方法来寻找游戏的最好的策略, 这种方法最早来源于囚犯的两难境地的研究. 在每一轮循环赛中用平均在每个游戏中具有最高支付的算法作为优胜的策略.

3 策略设计及其组织决策意义 (Strategies and their organization decision meaning)

在本文的研究中, 用了 4 个适应性策略来进行前面所描述的简单的博弈的循环赛. 每个单个的游戏持续 200 个回合. 由于两个参加者都有 4 种策略可选择, 故共有 16 种组合. 16 种组合中的每一种都进行了 128 次游戏, 产生了这个模型的 2048 个执行.

3.1 学习分类器系统 (Learning classifier system)

利用学习分类器系统, 将学习分类器的规则库中包含 20 个规则(从 64 个规则中随机选出). 每个

规则的条件部分与对手前一轮的行动相匹配, 行动部分代表当前的回合中所采取的行动. 一个典型的规则是如下形式的: 假如对手前一轮的行动为 101, 则赛手采取行动 110.

所有匹配对手前一轮行动的规则库都有资格为了被激活而投标. 投标实力是规则实力加一个随机高斯误差(均值为 0, 标准方差为规则实力的平方). 具有最高的投标值的规则被允许激活其行动部分.

在一个行动的支付被确定后, 获胜的规则的实力将根据绩效进行更新. 奖励(或惩罚)是两个因素的函数, 绝对绩效水平和两个赛手的绩效差异:

$$S_{AX} = S_{AX} + \alpha P_A + 1.1(P_A - P_B). \quad (5)$$

其中: S_{AX} 是指选手 A 的获胜规则 X 的实力; α 是奖励系数, 如果 $P_A \geq 0$, α 为 1, 如果 $P_A < 0$, 则 α 为 2; P_A, P_B 分别指 A 和 B 的支付.

这种奖励机制使得不能产生正的支付的规则受到严厉的惩罚. 不能比对手表现得更好的规则也会受到惩罚.

学习分类器系统依赖于遗传算法产生新的规则来提高其规则库(一般只包含所有规则的一小部分)的质量. 在现有系统中, 每 50 个回合就利用遗传算法将最弱的 4 个规则替换掉. 被替换的规则是经过对规则库中的具有最低实力的 10 个规则进行随机抽样挑选出来的. 为了使选择倾向于较强的和适应性相对较高的规则, 用轮盘赌的方法进行父母规则的选择, 被选择的概率与每一个规则的实力是成比例的.

父母规则通过交叉产生子代规则. 在交叉时利用双交叉技术对条件部分和行动部分都进行交叉. 在交叉中, 新规则的每个比特位都有 1% 的变异概率. 子代规则的实力是这样得出的:

$$S_{子代} = P S_{父亲} + (1 - P) S_{母亲}. \quad (6)$$

其中 P 是 0 和 1 之间的一个随机数.

学习分类器模型对模拟学习型组织具有较大的实际意义^[3,6]. 在信息不充分的情况下, 组织往往都是根据过去的经验来进行战略选择的, 战略选择后根据获得的绩效来决定下一步的战略行动, 促进组织成功的战略经验就会被保留继承, 并通过遗传算法所模拟的创造过程根据过去的成功经验生成新的战略行动方案, 这就是学习型组织的知识积累、学习和创造过程.

3.2 分类器系统 (Classifier system)

分类器系统利用与学习分类器系统相同的投标

和奖励系统但没有用遗传算法产生新规则,为了便于比较结果在规则库中放入了所有可能的,64个规则.这种策略可以模拟实际决策问题中信息比较完全的情况.可以预测这个分类器策略会优于学习分类器策略,因为它不需要通过搜索来寻找最佳战略.这个分类器系统会以最快的方式找出最好的规则而不需要去创造最好的规则.当然如果学习分类器系统的规则库中恰好包含最优的规则,那它就会和这个分类器系统表现得一样好.

在问题空间较大时,由于所有可能的规则的数量可能会超过记忆和计算容量,上述分类器系统在实践中是不可行的.这其实提醒了组织在进行决策时应尽量掌握比较充分的信息,竞争情报对组织的绩效是有直接影响的.

3.3 反馈系统(Feedback system)

在反馈系统中,agent通过比较期望值和绩效来调整自己的行动.如果绩效超过了期望值,agent就会坚持自己的行动模式并将内在期望值提高,如果绩效低于期望值,agent就会改变自己的行动模式.

在本文的研究中,利用反馈系统策略的选手的初始行动为000,期望值为0.1.如果绩效低于期望值,那么行动字符串中的3个比特位每1个都有33%的概率产生跳跃(比如从0到1,或从1到0).最终,总会找到一种使绩效超过初始期望值的行动状态.之后,期望值通过下列算法进行调整:

$$AI_{t+1} = \alpha P_t + (1 - \alpha) AI_t, 0 < \alpha < 1. \quad (7)$$

其中:AI为期望值, α 为调整系数,取0.05, P 为绩效值.

经过一段时间的运算随着期望水平的提高和随机搜索产生一个具有较高的绩效的系统.如果环境中的变化使得系统的绩效低于系统的期望值,系统就会立即搜索新的解决方法,从而系统的运行相对稳定并在绩效差时会立即作出反应.

利用反馈系统来激励人工适应agent的方法在组织管理中是非常普遍的,比如董事会考核企业领导班子时往往是定下一个目标,在考核时如果达到就再增加期望值,如果不能达到就进行领导替换.

3.4 随机游走策略(Walking system)

随机游走策略在这个演示系统中是一种缺省策略.以行动000开始,每一回合每个比特位有5%的跳跃概率(从0到1,或从1到0).从而,利用这种策略在不考虑绩效的情况下在所有可能的行动中进行

一种随机的运行.这种策略与实际中“拍脑袋”等一些简单的、随意的决策模式有极大的相似性.

4 模拟结果及其分析(Result and analysis)

每进行200个回合的运行后,每个策略的绩效的差异通过A和B的累计绩效的差异值计算出来.下表中列出了每种策略组合的累计绩效的差的平均值.

表2中的数值为B的累计绩效减A的累计绩效,策略1为学习分类器系统,策略2为分类器系统,策略3为随机游走策略,策略4为反馈系统.

表2 各种策略的绩效比较

Table 2 Comparison of the performance

策略编号	1	2	3	4
1	1.38	-30.47	78.36	10.59
2	33.19	-5.01	105.99	84.67
3	-84.22	-107.95	-2.62	-69.40
4	4.89	77.07	68.44	0.39

由表中可以看出分类器系统策略的绩效远远高于其他的策略,学习分类器系统平均绩效比反馈系统要高.随机游走策略的绩效比其他所有的策略的绩效都要低.

组织尤其是以盈利为目的的企业时时刻刻都处在竞争中,只不过其博弈形势比起上面的博弈情况还要复杂得多.相信对4种绩效寻优策略的效率的比较对企业的战略决策模式具有非常大的启发意义.

参考文献(References):

- [1] 任守策,张蕾,刘祖照.基于自组织的决策模式及决策支持系统[J].计算机集成制造系统——CIMS,2000,6(3):66-66.
(REN Shouqu, ZHANG Lei, LIU Zuzhao. Decision-making system based on self-adaptive [J]. *Computer Integrated Manufacturing System - CIMS*, 2000, 6(3):60-66.)
- [2] 黄建军,谢维信.模糊联想分类器及其进化训练[J].西安电子科技大学学报,1998,25(1):62-69.
(HUANG Jianjun, XIE Weixin. Illegible association classifier and its evolution training [J]. *J of Xi'an Electronics Technology University*, 1998, 25(1):62-69.)
- [3] TAKADAMA K, TERANO T, SHIMOHARA K, et al. Making organizational learning operational: implication from learning classifier system [J]. *Computational and Mathematical Organization Theory*, 1999, 5(3):229-252.
- [4] YILDIZOGLU Murat. Modeling adaptive learning: R&D strategies in

(下转第476页)

参考文献(References):

- [1] KOKAME H, KOBAYASHI H, MORI T. Robust H_{∞} performance for linear delay-differential systems with time-varying uncertainties [J]. *IEEE Trans on Automatic Control*, 1998, 43(2): 223 - 226.
- [2] de SOUZA C E, LI Xi. Delay-dependent robust H_{∞} control of uncertain linear state-delayed systems [J]. *Automatica*, 1999, 35(9): 1313 - 1321.
- [3] CAO Yongyan, SUN Youxian, LAM J. Delay-dependent robust H_{∞} control for uncertain systems with time-varying delays [J]. *IEE Proc-Control Theory and Applications*, 1998, 145(3): 338 - 344.
- [4] KEEL L H, BHATTACHARYYA S P. Robust, fragile, or optimal [J]. *IEEE Trans on Automatic Control*, 1997, 42(8): 1098 - 1105.
- [5] FAMULARO D, DORATO P, ABDALLAH C T, et al. Robust

non-fragile LQ controller: the static state feedback case [J]. *Int J Control*, 2000, 73(2): 159 - 165.

- [6] YANG Guanghong, WANG Jianliang, LIN Chong. H_{∞} control for linear systems with additive controller gain variation [J]. *Int J Control*, 2000, 73(16): 1500 - 1506.

作者简介:

王 武 (1973 —), 男, 1999 年在福建农业大学机电工程系获硕士学位, 现为福州大学电气工程系博士研究生. 研究方向为时滞系统和不确定系统的非脆弱 H_{∞} 控制以及 LMI 方法应用研究等.

E-mail: ofwyang@fzu.edu.cn 或 wangwu263@163.net;

杨富文 (1963 —), 男, 1990 年在华中理工大学自动控制系获博士学位, 现为福州大学电气工程系教授, 博士生导师. 主要研究方向为 H_{∞} 控制与滤波, 迭代学习控制以及自动化工程应用研究.

(上接第 472 页)

the model of nelson & winter [EB/OL]. [2001 - 05]. Http://yildizoglu.nantesquieu.u - bordeaux. fr

- [5] 易继铠. 智能控制技术[M]. 北京: 北京工业大学出版社, 1999: 75 - 80.
(YI Jikai. *Intelligent Control Technology* [M]. Beijing: Beijing University of Technology Press, 1999: 75 - 80.)
- [6] KIM D. The link between individual and organizational learning [J]. *Sloan Management Review*, 1993, 35(1): 37 - 50.

作者简介:

冯彦杰 (1971 —), 女, 上海交通大学管理学院管理科学与工程博士研究生, 从事复杂系统, 企业管理等方面的研究. E-mail: fengyanjie@263.net;

王浣尘 (1933 —), 男, 上海交通大学管理学院教授, 博士生导师, 21 世纪发展研究院执行副院长, 管理科学与工程博士后流动站站长, IEEE 资深会员, 中国系统工程学会副理事长, 曾发表专著 3 部, 论文 300 余篇. 研究领域涉及自动控制, 系统科学与系统工程, 管理科学与工程, 经济控制论, 社会经济与企业发展战略, 可持续发展, 信息化与网络化等.