

文章编号: 1000-6788(2004)09-0085-05

智能组卷系统的建模与算法研究

王友仁, 张砦, 崔 江, 姚 睿, 储剑波

(南京航空航天大学自动化学院, 江苏 南京 210016)

摘要: 根据教育考试理论与统计学, 分析了自动组卷目标要求, 建立了智能组卷系统的数学模型. 提出了一种基于免疫遗传算法的智能组卷算法, 提出了分段编码、段内遗传操作、误差加权适应度评估方法. 通过应用实例分析, 实验结果表明所提出新组卷算法的组卷成功率高、组卷速度快.

关键词: 教学专家系统; 自动组卷系统; 组卷目标; 数学模型; 免疫遗传算法

中图分类号: TP319

文献标识码: A

Mathematical Model and Algorithm of Intelligent Test Paper Auto-generation System from Item Pool

WANG You-ren, ZHANG Zhai, CU I Jiang, YAO Rui, CHU Jian-bo

(College of Automation and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract With the help of examination theory and education mathematics, the requirements of the test paper are analyzed in details. The mathematical model of intelligent test paper generation system is set up. The method of test paper auto-generation based on immune genetic algorithm (IGA) is proposed. Segmented coding, crossover and mutation in segment, evaluation function of genetic are developed. Instance analysis to intelligent system of test paper generation from test item pool about analog electronic technology is present. The experimental results show that the IGA algorithm can solve the test paper auto-generation problem in test item pool effectively, and has better performances of quick convergence to genetic algorithm.

Key words education expert system; intelligent system of auto-generating test paper; characteristics of test paper; mathematical model; immune genetic algorithm

1 前言

目前计算机辅助教学(Computer Assistant Instruction-CAI)已经成为现代高等教育的一个重要组成部分, 作为其主要研究方向之一, 电子试题库技术与自动组卷系统发展迅速. 自动组卷系统是实现无纸化考试、考试标准化与个性化、远程网络学习辅导与在线测试等的基础. 要想设计一份能够比较全面、准确地测试出学生掌握有关知识情况的试卷, 通常都需要由具有较高水平的教育专家(组), 经过一定时间的研究与设计才能完成, 试卷的质量与教育专家个人的经验、知识水平、认真程度、个人见解等有关. 所谓自动组卷系统, 是将人工智能技术与人类教育专家的组卷知识和经验结合起来, 由计算机来完成试卷内容的设计, 且使得由计算机所生成的试卷达到专家级水平.

自动组卷系统研究中的主要难题是如何保证生成的试卷能最大程度满足用户的不同需要, 并具有随机性、科学性、合理性. 而在网络交互式环境下, 用户对组卷速度的要求较高. 因此, 高效、科学、强壮的智能组卷算法是实现自动组卷的关键技术, 也是目前CAI领域的研究热点之一. 目前的智能组卷算法^[1-4]主要是盲目随机选取法、回溯试探法、遗传算法, 这些算法往往因为约束条件的局部满足而导致组卷失败, 同时组卷过程所需时间开销大, 无法满足实时应用. 因此, 迫切需要根据试卷生成问题的特点, 提

收稿日期: 2003-11-03

作者简介: 王友仁(1963-12-), 男(汉), 南京市人, 副教授, 博士. 目前主要研究方向为计算机辅助教学系统、仿生硬件理论与技术等, Email: wangyrac@nuaa.edu.cn

出全局寻优和收敛速度快的智能组卷算法。

2 组卷问题的描述

根据用户输入的组卷要求, 搜索试题库中特征参数相匹配的试题, 生成满足要求的试卷。自动组卷系统理论基础涉及试卷评价指标与组卷要求、数学模型、智能组卷算法、试题库中试题属性及编码方法。

2.1 组卷目标分析

计算机根据用户的组卷要求自动生成试卷, 这就涉及试卷的质量评价指标问题。传统的试卷评价指标^[1, 2]主要以试卷的难度、区分度、信度和效度来作为衡量的标准。试卷难度的计算公式为

$$D = \frac{1}{P} \cdot \sum_{i=1}^M d_i \cdot p_i = 1 - \frac{\bar{Y}}{P}, \quad (1)$$

其中, P 为试卷满分值, d_i 为第 i 道试题的难度值, p_i 为第 i 道试题的参考分值, M 为试卷的题目总数, \bar{Y} 为该份试卷全部考生的平均得分值。试卷难度值可以根据各道试题的难度值 d_i 来估算, 而由考生的平均分数又可验证试卷的难度系数是否正确。但 d_i 是需要大量统计数据来估算的, 无法精确测量。

试卷区分度的计算公式为

$$F = \frac{1}{P} \cdot \sum_{i=1}^M F_i \cdot p_i = a \cdot \delta + b \quad (2)$$

其中, F_i 为第 i 道试题的区分度, a 、 b 为常数(由经验确定), δ 为试卷考试成绩的标准差。由 δ 可验证试卷区分度的正确性。试卷的区分度以试题的区分度 F_i 为基础, 而 F_i 一般难以精确计算。试题库中试题的区分度初值由专家根据经验确定, 再根据考试情况进行调整。

试卷的信度与效度一般只能定性评估。由于传统的试卷性能评价需要试题属性值, 而统计测量与专家经验都难以获得精确的试题属性值。因此, 必须结合定量与定性方法, 综合考虑到一般试卷质量指标、教学要求、特定用户需要, 来确定合适的试题属性值, 建立对试卷的科学评价准则, 形成智能组卷目标要求。在自动组卷系统中, 试题分数、估计用时为定量表示, 试题所属知识点、试题类型、知识掌握要求用符号表示, 试题难度、区分度用模糊分级表示。

组卷目标是生成一份试卷的性能要求与特征参数, 所设计的组卷目标主要包括: 试卷总分、各章节分值、各题型分值、试卷难度、考试时间、知识点满足、能力层次满足、试卷区分度、试卷形式。

2.2 数学模型

计算机自动组卷过程为在一定题量的试题库中搜索满足组卷目标要求的一组试题组合。根据组卷目标分析来确定每道试题的主要属性代码, 对于试题库中全部试题可以预先建立目标状态矩阵 A

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,C} \\ a_{2,1} & a_{2,2} & \dots & a_{2,C} \\ & & \vdots & \\ a_{N,1} & a_{N,2} & \dots & a_{N,C} \end{bmatrix} \quad (3)$$

其中, N 代表试题库中全部试题的总数, C 为每道试题的属性代码项数(也对应生成试卷时最多的组卷目标要求, 即 $C = 9$)。在矩阵 A 中每行代表一道试题及其属性, 每列代表试题库中全部试题的某个属性。组卷过程问题求解为: 根据用户输入或缺省设置的组卷目标要求, 在矩阵 A 中寻找满足组卷目标要求的行组合, 即搜索确定矩阵 A 中全部行状态的多变量组合优化 (K_1, K_2, \dots, K_N)。这里 K_i 为第 i 行的状态变量(即第 i 道题), 若 $K_i = 1$ 表示第 i 行被选中, 而 $K_i = 0$ 表示第 i 行没有被选中。

每个组卷目标要求都对应着一个约束条件, 则全部 C 项组卷目标要求分别对应着如下约束条件:

试卷总分: $\sum_{i=1}^N K_i \cdot a_{i,1} = P$, P 是试卷总分, $a_{i,1}$ 为第 i 道试题分值;

章节分数: $\sum_{i=1}^N K_i \cdot b_{i,s} = P_s$, 第 s 章节试题分数为 P_s 。当试题章节编号 $a_{i,2} = s$ 时有 $b_{i,s} = a_{i,1}$, 否则为零;

题型分数: $\sum_{i=1}^N K_i \cdot g_{i,m} = W_m$, W_m 为第 m 种题型试题的分数. 当试题题型编号 $a_{i,3}$ 属于第 m 种题型时有 $g_{i,m} = a_{i,1}$, 否则 $g_{i,m}$ 等于零. 试题共分成 7 种题型 (即选择填空 A、问答题 B、改错题 C、证明题 D、分析判断题 E、计算题 F、综合题 G);

试卷难度: $\sum_{i=1}^N K_i \cdot a_{i,1} \cdot a_{i,4} = P \cdot D$, $a_{i,4}$ 为第 i 道试题的难度系数, D 为指定试卷的难度值.

考试时间: $\sum_{i=1}^N K_i \cdot a_{i,5} = T$, T 为用户要求的考试时间, $a_{i,5}$ 为第 i 道题的预计答题时间.

知识点满足: $\sum_{i=1}^N K_i \cdot u_{i,n} = Z_n$, 第 n 个知识点试题分值 Z_n . 当试题知识点属性编号 $a_{i,6}$ 等于第 n 个知识点时有 $u_{i,n} = a_{i,1}$, 否则为零;

能力层次满足: 教学内容掌握的层次要求分为一般了解、熟悉理解、熟练掌握和灵活运用四个层次, $a_{i,7}$ 为第 i 道试题的教学内容掌握层次要求编码. 显然, 所选择的试题应满足指定的能力层次要求;

试卷区分度: $\sum_{i=1}^N K_i \cdot a_{i,1} \cdot a_{i,8} = F$, $a_{i,8}$ 为第 i 道题的区分度;

试卷形式: $a_{i,9} \in C$ (用户指定试卷形式), $a_{i,9}$ 为每道题的形式属性代码. 分为客观形式和主观形式两种, 前者支持客观标准做答, 后者支持过程主观输入做答 (计算题需人工阅卷).

3 智能组卷问题求解算法

3.1 算法的描述

免疫遗传算法是一种引入免疫机制的遗传算法^[5-8], 与免疫算法相似, 把待求解的问题 (目标函数和约束条件) 作为抗原, 问题的解对应为抗体. 与免疫算法相比, 免疫遗传算法是在免疫算法中增加了遗传算子; 与遗传算法相比, 增加了抗原识别、记忆和调节功能. 这种新算法既保留了遗传算法的鲁棒性, 又克服了遗传算法的局部搜索效率较差、可能出现早熟等缺点, 兼顾了搜索速度、全局搜索能力和局部搜索能力. 免疫遗传算法的一般过程为:

抗原识别 (包括要解决的问题、问题的最佳解、抗体编码等); 初始抗体群产生; 抗体适应度和亲和力计算; 抗体的产生; 群体更新; 判断是否满足结束条件? 若不满足结束条件, 则跳至第三步. 否则结束.

3.2 免疫遗传算法设计

智能组卷问题对应为抗原, 所生成的满足组卷目标要求的试卷对应为抗体. 算法实现过程设计如下:

1) 编码方法确定

设题库中共有 N 道题, 有 7 种题型 (每种题型分别有 N_1 至 N_7 道题). 用 N 位二进制字符串来表示题库中各道题的状态, 即 $(J_1 J_2 \dots J_N)$, 其中若 $J_k = 1$, 则表示第 k 道题被选中, $J_k = 0$ 则表示没有被选中.

对于有 M 道题的试卷, 则应满足: $\sum_{k=1}^N J_k = M$.

根据题库中同类型试题的分值多数相同的特点, 提出了按照题型进行分段编码的方案. 其思路为先对各题型试题分别进行独立编码, 然后再合并成 N 位代码, 则总编码串为 $(J_{1,1} \dots J_{N_1,1}), \dots, (J_{1,7}, \dots,$

$J_{N_7,7})$. 设组卷要求的各种题型的试题数分别为 M_1, \dots, M_7 , 即 $\sum_{k=1}^7 M_k = M$, 则应该满足: $\sum_{i=1}^{N_k} J_{i,k} = M_k$, ($k = 1, \dots, 7$). 这种分段独立编码方案可缩小解的搜索空间, 明显简化了问题的求解.

2) 适应度函数设计

构造染色体满足组卷目标的误差函数为

$$E = \sum_{k=1}^9 \alpha_k \cdot e_k, \quad (4)$$

其中, $e_k = 0 \sim 1$ 为反映染色体满足第 k 项组卷要求程度的归一化相对误差, $0 \leq \alpha_k \leq 1$ 为相应的误差加

权系数. 全部组卷约束条件被分成为强约束、一般约束、弱约束, 强约束项包括总分值、时间、难度、题型, 一般约束项包括章节内容、教学内容水平能力、知识点, 弱约束项包括试卷形式、区分度. 强约束项对应有一定的误差加权系数, 而弱约束项对应的误差加权系数小. 另外, 根据实际组卷专家经验, 对不同的约束条件给定了不同的允许容差(1% - 5%), 只要染色体满足第 k 项组卷要求的误差在容差范围内, 即可以认为 $e_k = 0$, 这样以加快搜索到合理的解.

染色体的适应度函数定义为

$$Fit = (E + \mu)^{-1}, \quad (5)$$

其中, 取 $\mu = 0.01$ (经验值).

3) 初始种群生成与搜索结束条件

根据组卷目标中各种题型的分值要求, 在初始种群产生时, 对于每个独立代码段采用随机产生初始群体的办法, 这样就使得整个初始种群基本满足了题型分布与总分值的要求. 因题型分布、总分值已基本满足, 相当于减少了总约束项, 这对于加快进化速度有利.

搜索结束条件的判断: 一是最大染色体适应度达到 $Fit = 99 - 100$; 二是种群中最大个体适应度 $Fit > 95$, 且在连续 5 代内最大适应度值改善小于 0.1 时; 三是已进化到规定的最大代数 (现设为 1000) 时, 终止进化. 若此时最大个体适应度值 $Fit < 95$, 则认为本次问题求解搜索失败.

4) 进化算子设计

采用分段单点交叉策略: 对随机选择的两个染色体, 在每个分编码段内采用随机产生交叉位置的单点交叉, 且判断交叉后该段内的题型数不变. 若段内题型数改变, 则重新选择交叉点 (至多两次选择机会).

采用分段一次变异策略: 对染色体的各个分编码段依次进行一次段内变异操作, 各个段的变异算子是互相独立的. 先根据变异率随机确定将段内某位取反, 然后在位所在分段内, 随机确定向后或向前方向寻找与该位原值相反的位, 并将其值取反.

5) 基于浓度的群体更新策略

群体更新策略: 最大适应度个体被直接复制 (即最优模式保持策略), 并且最优个体不参与交叉、变异操作. 然后依据个体的选择概率来挑选父代进行遗传操作来繁殖产生新的种群, 每代的种群规模维持不变. 在群体更新过程中, 结合了免疫机制中抗体之间基于浓度的相互抑制作用, 通过抗体浓度因子来调节个体的生成选择概率. 在保证适应度高的个体有较大的选中概率前提下, 抑制浓度过高的抗体. 因为, 抗体的浓度若过高, 则可能引起进化早熟或局部收敛.

抗体的浓度计算公式为: $G = \frac{(0.85 \cdot Fit_{max} \sim Fit_{max}) \text{ 之间抗体数目}}{\text{抗体总数目}}$, 这里 Fit_{max} 为当前代最大适应度值.

基于浓度因子的第 k 个个体的选择概率 h_k 为

$$h_k = (1 - \alpha_k \cdot G_k) \cdot \frac{Fit_k}{\sum_i Fit_i}, \quad (6)$$

其中, $\alpha_k = 0 \sim 1$ 的常系数 (经验值), Fit_k 为第 k 个个体的适应度值.

算法实现步骤为: 抗原接受; 生成初始抗体群; 计算抗体适应值; 计算浓度与选择概率; 进化算子操作; 群体更新; 判断是否满足结束条件? 若不满足, 则返回第三步. 否则结束进化过程.

4 实例分析

下面以模拟电子技术基础试题库为例, 来分析自动组卷系统的实现过程及实验结果.

4.1 组卷要求确定

在自动组卷系统中, 要求用户输入的组卷目标项有: 试卷总分、题型分值、章节分值、试卷难度、考试时间、高级设置项 (各章节知识点、知识能力掌握、试卷形式、全卷区分度). 用户可以根据实际需要来设定组卷目标要求, 也可使用系统默认值来进行组卷.

试卷总分默认为 100 分, 各章的默认分值为第 8、10 章是 0 分, 第 1、9 章小于 10 分, 其余各章为 11-

20 分; 各题型试题默认值为选择填空题 30 分、分析判断题 20 分、计算题 50 分; 试卷难度共分成 A、B、C、D、E 五个等级, 分别为难(A 对应难度系数为 0.75~ 1.0)、比较难(B 对应难度系数为 0.5~ 0.75)、中等(C 对应难度系数为 0.3~ 0.5)、比较容易(D 对应难度系数为 0.2~ 0.3)、容易(E 对应难度系数小于 0.2), 默认难度等级为中等; 考试时间默认为 120 分钟; 高级设置项由用户设定, 默认为没有要求。

4.2 试题属性与编码设计

试题库中每道试题的存放信息包括属性代码及题号、试题内容、参考答案、考试反馈信息(答题情况与选题次数)。试题属性代码代表一道试题的特性, 自动组卷系统根据各道试题的属性代码来进行组卷。

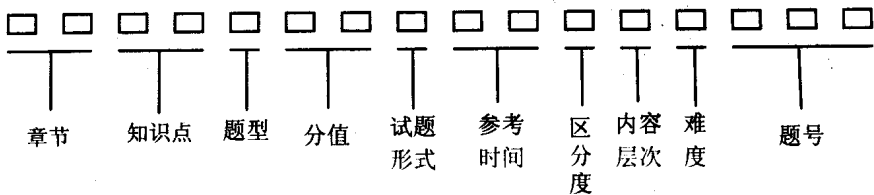


图 1 试题属性代码结构

本系统中设置每道试题的属性码为 16 位(如图 1 所示)。例如某道题属性码为 0103101001422002, 其含义为: 第一章中第三个知识点、选择题、分值 1、客观题、1 分钟、区分度为 4、难度一般、题序是 2。

4.3 实验结果分析

实验时控制参数选择为: 种群规模为 100, 交叉概率 $p_c = 0.2$, 变异概率 $p_m = 0.005$, $\alpha_i = 0.3$ 。标准遗传算法(SGA)采用赌轮选择方案, 控制参数不变。电子试题库中试题总量为 6000 道题。

表 1 为 100 次运行的实验统计结果, 没有出现连续两次组卷失败现象。表 2 和表 3 为默认组卷参数时自动组卷的典型结果。由实验可知, 智能组卷算法能获得满足组卷要求的试卷, 免疫遗传算法的组卷成功率与收敛速度要优于标准遗传算法。

在免疫遗传算法中增加了简单的抗体浓度 G_k 计算公式, 其计算量在每代进化的计算量中所占比例很小。个体选择概率 h_k 计算公式中的常系数 α_i 常根据经验与实验确定, 选取 $\alpha_i = 0.2 \sim 0.4$ 之间比较好。

表 1 算法运行结果示例

	平均收敛代数	组卷成功率
免疫遗传算法	232	99%
标准遗传算法	405	95%

表 2 试卷中试题难度分布

试题难度等级	E	D	C	B	A
对应难度等级分值	12	25	36	21	6

表 3 各章分值分布

章号	一	二	三	四	五	六	七	八	九	十
分值	7	18	12	20	15	10	10	0	8	0

5 结束语

基于免疫遗传算法的智能组卷方法已经成功用于模拟电子技术题库自动组卷系统中, 其在组卷成功率、组卷速度上得到了明显的提高。智能组卷系统设计涉及教学理论、组卷方法与计算机软件开发技术, 对该问题的深入研究将会推动高等教育的现代化发展。

参考文献:

[1] 王丽娜, 肖京, 董晓梅等. 通用智能题库系统研究[J]. 东北大学学报(自然科学版). 1999, 20(2): 155- 158
WANG Lina, XIAO Jing, DONG Xiao-mei, et al. Study on general intelligent item bank [J]. Journal of Northeastern University (Natural Science), 1999, 20(2): 155- 158

[2] 全惠云, 范国闯, 赵霆雷. 基于遗传算法的试题库智能组卷系统研究[J]. 武汉大学学报(自然科学版), 1999, 45(5): 758- 760
QUAN Huiyun, FAN Guo-chuang, ZHAO Ting-lei. Intellignet system of test paper auto-generating from test bank based on genetic algorithms[J]. J. Wuhan Univ. (Nat. Sci. Ed.), 1999, 45(5): 758- 760

(下转第 97 页)

7 结束语

本文利用有向图及有向图相关的拟邻接矩阵将层次分析积因子方法推广到一般的网络决策分析问题. 在讨论如何构造网络、计算判断矩阵量化支配关系、综合集成的基础上, 阐述了网络决策分析积因子法的步骤及特点.

参考文献:

- [1] Liu Q izhi, Wang Q in. Product method of analytic hierarchy process [J]. Asia-Pacific Journal of Operational Research, 1991, (8): 135- 145.
- [2] Liu Q izhi, Wang Q in. Product method of AHP [A]. Proceedings of International Symposium on the Analytic Hierarchy Process[C], Tianjin, China, 1988, 225- 231.
- [3] 刘奇志. 层次分析积因子方法的保序性[J]. 系统工程学报, 1995, (10), 61- 70.
Liu Q izhi. Rank preservation property of product AHP[J]. System s Engineering Journal(Chinese), 1995, (10): 61- 70.
- [4] Liu Q izhi. Product AHP and Its Properties [A]. Proceedings of the 6th International Symposium on Analytic Hierarchy Process (ISAHP) [C]. Bern-Sw itzerland, 2001. 341- 348.
- [5] Saaty T L. The Seven pillars of analytic hierarchy process [A]. Proceedings of the Fifth International Symposium on Analytic Hierarchy Process [C]. Kobe, Japan, 1999, 20- 23.
- [6] Saaty T L. A note on multiplicative operations in AHP [A]. Proceedings of International Symposium on the Analytic Hierarchy Process [C]. Tianjin, China, 1988, 82- 86.
- [7] Saaty T L. Observations on multiplicative composition in the analytic hierarchy process [A]. Proceedings of the 3rd International Symposium on Analytic Hierarchy Process [C]. Washington DC, U SA, 1994, 169- 174.
- [8] Barzilai J. On the derivation of AHP properties [A]. Proceedings of the Fourth International Symposium on the Analytic Hierarchy Process [C]. Burnaby, B. C. Canada, 1996, 244- 250.
- [9] Ramanathan R Group. Decision making using multiplicative AHP [A]. Proceedings of the Fourth International Symposium on the Analytic Hierarchy Process [C]. Burnaby B C, Canada, 1996, 262- 272.
- [10] 许树柏. 层次分析法原理 [M]. 天津: 天津大学出版社, 1988.
Xu Shubo. The Principle of AHP. Tianjin University Press (Chinese), Tianjin, 1988.
- [11] 王连芬, 许树柏. 层次分析法引论 [M]. 北京: 中国人民大学出版社, 1990.
Wang L ianfen, Xu Shubo. The Introduction to AHP (Chinese). China Renm in University Press, Beijing, 1990.
- [12] 赵玮, 姜波. 层次分析方法进展 [J]. 数学的实践与认识, 1992, (3): 63- 70.
Zhao Wei, Jiang Bo. Recent Development of AHP [J]. Mathematics in Practice and Theory (Chinese), 1992, 3: 63 - 70.
- [13] Saaty T L, L G Vargas. Decision Making [M]. RW S Publication, Pittsburgh, 1994.

(上接第 89 页)

- [3] 魏平, 张元. 一种求解组卷问题的遗传算法 [J]. 宁波大学学报(理工版), 2002, 15(2): 47- 50.
WEI Ping, ZHANG Yuan. Test paper generation with genetic algorithm [J]. Journal of N N GBO University (NSEE), 2002, 15(2): 47- 50.
- [4] 刘彬, 金涛, 陈大平. 遗传算法在试题组卷中的应用 [J]. 燕山大学学报, 2002, 26(3): 193- 196.
L I U Bin, J N Tao, CHEN Da ping. Application of a geneti algorithm to composing a test paper [J]. Journal of Yanshan University, 2002, 26(3): 193- 196.
- [5] Takahashi K. & Yamada T. Application of an immune feedback mechanism to control systems [J]. JSM E International Journal, Series C, 1998, 41(2): 184- 191.
- [6] Nikolaev N I, Iba H, Slavov V. Inductive Genetic Programming with Immune Network Dynamics [M]. In Advances in Genetic Programming 3, MIT Press, chapter 1999(15): 355- 376.
- [7] Kayawa M. Sugita Y. Morooka. Sensor Diagnosis System Combining Immune Network and Learning Vector Quantization [J]. Electrical Engineering in Japan, 1996(117), 5: 44- 55.
- [8] Mori K, Tsukiyama M, Fukuda T. Application of an Immune Algorithm to Multi- optimization Problems [J]. Electrical Engineering in Japan, 1998, 122(2): 30- 37.