

文章编号:1000-582X(2011)10-142-05

容器编程遗传算法的应用

任 薇, 张晓东, 邱玉辉

(西南大学 计算机信息与科学学院, 重庆 400715)

摘 要:随着高校课程与学生的高速增长,基于计算算法的高校自动排课方法成为目前算法设计研究的重要问题。基于遗传算法机制,提出了一种新的二进制编码机制和教室位置填充方法,用于改善排课的系统负载与效率。该算法摒弃了完全随机搜索的做法,依据适应度函数中各项权重比例的多寡为导向,定向随机生成染色体中的基因,精简了染色体信息量,大大缩短了系统的运行时间。在管理机制上,引入用户干预,高效处理初始化种群和排除冲突。仿真实验数据证明了该算法的收敛性与高效率。此外,在整个系统设计中,运用了基于容器的容器编程技术,实现了不定数据量的便捷处理。

关键词:遗传算法;容器;课程处理

中图分类号:TN929

文献标志码:A

An application based on container programming and genetic algorithm

REN Wei, ZHANG Xiao-dong, QIU Yu-hui

(College of Computer and Formation Science, Southwest University, Chongqing 400715, P. R. China)

Abstract: With the rapid growth of curriculums in colleges and universities and the number of students, automatic course arrangement method based on calculation algorithm has become an important issue of algorithm design and research. Based on Genetic Algorithm, a new binary encoding mechanism and classroom position filling method is put forward, which is applied to improve the load and efficiency of course scheduling system. In accordance with the ratio of weights in Fitness Function, this method abandons the way of random search, directionally and randomly generates chromosomes, simplifies the quantity of chromosome information, greatly shortens the running time of the system. In the aspect of management mechanism, user intervention is introduced, dealing with initial population with high proficiency and eliminating conflicts. Data from simulation experiment proves the astringency and high efficiency of this method. Moreover, the whole design of the system applies programming technology based on container, realizing the convenient processing of uncertain data.

Key words: genetic algorithm; container; course scheduling

随着高校招生逐年扩张,大学课程向着广度和深度发展,高校的教师、教室等一些资源越发显得紧张,自动排课算法已成为目前计算算法应用研究领

域的重要课题之一。目前,在自动排课系统设计方面,国内学者存在一些已有的研究成果^[1-4],但算法设计上仍存在不足之处^[5-13]。根据高校开课的具体

收稿日期:2011-02-03

基金项目:重庆市自然科学基金资助项目(CSTC2009AC2174)

作者简介:任薇(1981-),女,西南大学博士研究生,主要从事人工智能与复杂网络方向研究,(Tel)13883538497;
(E-mail)oiq@swu.edu.cn。

情况,采用优化的编码方案。与文献[1]中提到编码时染色体上基因片存储信息大小为 11 字节的方法相比,编码方式中 1 个基因片只用了 6 字节,缩短了编码存储空间,提高了运算速度,经试验证明其速度优于已有算法。此外,文献[5]中介绍把教室与时间等信息共同处理的思路,这可能会产生空间上的冲突。针对排课过程中的一些硬性和软性约束因素,利用遗传算法对课表进行了优化,提出将教室分离出以单独处理,从而解决了空间上的冲突问题,获得了无冲突、更人性的排课方案^[14]。

1 问题描述

遗传算法(genetic algorithm)是一类借鉴生物界的进化规律(适者生存,优胜劣汰遗传机制)演化而来的随机化搜索方法。其主要特点包括:直接对结构对象进行操作,不存在求导和函数连续性的限定;具有内在的隐并行性和更好的全局寻优能力;采用概率化的寻优方法,能自动获取和指导优化的搜索空间,自适应地调整搜索方向,不需要确定的规则。

一个串行运算的遗传算法(sequential genetic algorithm, SGA)按如下过程进行

- 1) 对待解决问题进行编码;
- 2) 随机初始化群体 $X(0) = (x_1, x_2, \dots, x_n)$;
- 3) 对当前群体 $X(t)$ 中每个个体 x_i 计算其适应度 $F(x_i)$, 适应度表示了该个体的性能好坏;
- 4) 应用选择算子产生中间代 $X_r(t)$;
- 5) 对 $X_r(t)$ 应用其它的算子, 产生新一代群体 $X(t+1)$, 这些算子的目的在于扩展有限个体的覆盖面, 体现全局搜索的思想;
- 6) $t := t + 1$; 如果不满足终止条件继续 3)。

在排课算法设计中,对整个问题进行了数学描述。

算法对课程、班级、教师、时间及教室这 5 个要素定义如下^[2-4]

- 1) 课程集合: $L = \{l_1, l_2, \dots, l_A\}$;
- 2) 班级集合: $C = \{c_1, c_2, \dots, c_B\}$;
- 3) 教室集合: $R = \{r_1, r_2, \dots, r_Q\}$;
- 4) 教师集合: $S = \{s_1, s_2, \dots, s_K\}$;
- 5) 时间集合: $T = \{t_1, t_2, \dots, t_D\}$ 。

解决排课问题的基本算法流程如下

- 1) 初始化种群,即编码。编码时形成班级集合、课程集合、教师集合以及课程教师对 $L_S = \{(c_1, l_1, s_2), (c_2, l_3, s_8), \dots, (c_B, l_a, s_K)\}$;
- 2) 在时间集合上排序形成时间上无冲突的 1 个

班级-课程-时间-教师对集合 $A = \{(c_b, l_r, t_i, s_p), \dots, (c_B, l_R, t_I, s_K)\}$;

- 3) 根据限制条件,把班级-课程-时间-教师对关联到教室集合,形成完整的课表 $TB = \{(c_b, l_r, t_i, s_p, r_j), \dots, (c_B, l_R, t_I, s_K, r_J)\}$, 并解决空间冲突。

2 改进遗传算法的高校排课方法

根据一般解决方法,基于遗传算法的高校排课方法其流程可以分解为条件分析、初始化种群、冲突解决、个体冲突解决、个体适应度值计算、遗传操作,而遗传操作又分为选择算子、交叉算子、变异算子。在初始化种群时的编码与实现操作和教室位置处理两方面对传统遗传算法进行了进一步的改进,从而提高整个排课方法的效率,下面将对算法进行详细介绍。

2.1 染色体编码

染色体编码是按编码方案把排课信息对象抽象为由特定符号按一定顺序排成的串。编码把现实问题转化为遗传算法问题,所以编码是否高效直接影响后面的计算是否高效。

提出了一种新的二进制编码方案,即把排课信息中数据元素值转化成简便的二进制形式(如将十进制教师编号 00056 化为二进制编码 111000),缩短了编码的长度,而不像已有算法,将整个数据元素不经转化直接编码。其编码结构如图 1 所示

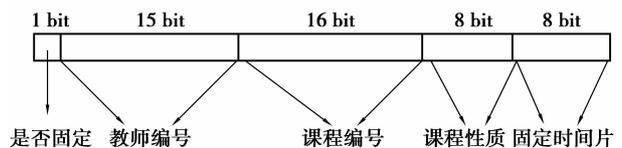


图 1 编码结构

编码结构中每个字段的定义如下

- 1) 是否固定:有些教师的课程是固定在某个时间段的,所以在排课过程中,判断第一位就可以得知可否移动此基因片。该字段针对某些教师存在的特殊的要求,如年龄、职务等的描述,反映了排课人性化;
- 2) 教师编号:提供 15 bit 描述教师代码,学校每个教师的编号都是唯一的,要求编码简单、高效;
- 3) 课程编号:提供 15 bit 课程描述代码,学校每门课程的编号都是唯一的;
- 4) 课程性质:此字段用于解决“特定资源”冲突问题。每一门课程都有其各自不同的特点,比如上机课需要在机房上课,英语口语需要在语音室上课,

体育课需要在操场上课,为此把 16 bits 平均分为 2 段,如图 2 所示

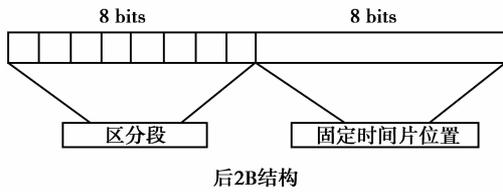


图 2 课程性质编码方式

根据高校一般课程设置方案,在此把课程分为专业必修课、专选科、公共课、上机实验课、专业实验课、体育课、英语语音课。前 8 位编码不同时表示不同的含义,例如:10000000(二进制)表示专业必修课。后 8 位编码表示当此教师固定在这个时间片上课时,所在的时间片值。此编码方案占用字节数少,仅占用 6B 就表示了要表达的信息,在大规模数据运算中速度快,效率更高。

2.2 种群初始化

在初始化种群时,基于容器的容器类的使用使其更加简捷。因为对于数据库需要对教学楼、教室、班级等一些数量未知的数据进行存储,运用数组显然不能很好地解决,会出现浪费内存、处理繁琐的情况,但容器就可以解决这些问题。

关联容器的使用方法如图 3 所示。

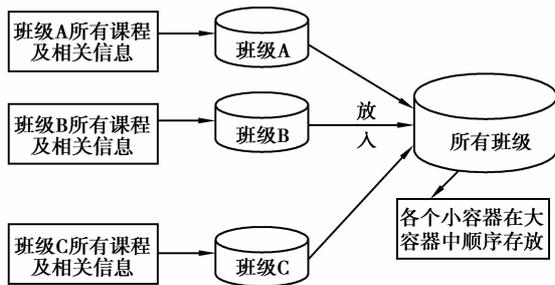


图 3 关联容器使用方法

2.3 教室位置填充

教室问题是单独处理的,这样在初始化种群和排除冲突时就可以有更快速地处理,更重要的是可以跟数据库结合(数据库中约束条件和范围),更加能够满足系统的要求,位置表由人去规划,计算机去实现,会使整个排课方案更优化。其功能实现如图 4。

对上面的处理迭代执行,直到所有时间片都处理结束,即可完成课程的安排和教室设置。对每一个时间片,对其进行操作,对之后的时间片将不会产

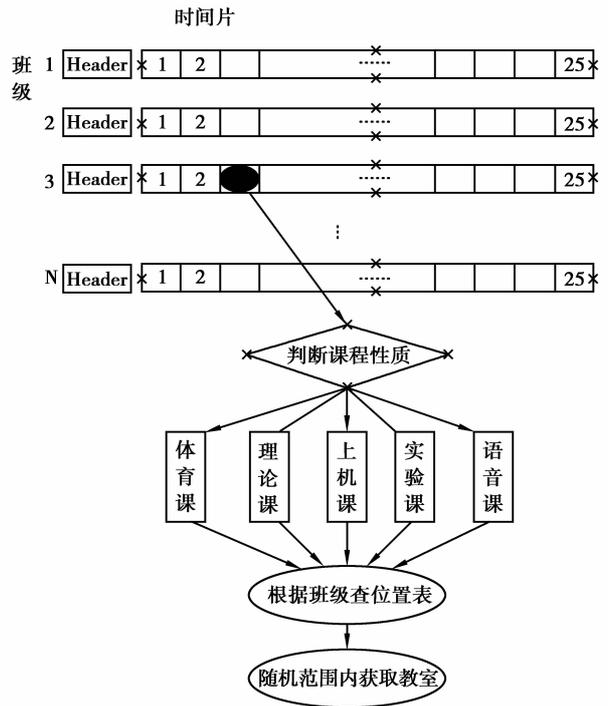


图 4 教室位置填充

生影响,因此,在本算法中,每列都使用所有的教室去填充,而不是上次填充剩下的教室。

3 实验结果及分析

为了验证算法的可行性和准确高效性,针对排课方法,使用了大量有效的数据进行了实验。实验平台是 Visual C++6.0 和 SQL server 2000。实验数据如表 1 所示。

表 1 排课数据

描述	值
班级数	102
教师数量	300
理论课教室数量	72
语音室数量	10
公共机房数量	12
所有课程数量	350
所有专业实验课教室数量	130
天数	5
每天节数	10
操场数量	2
教学任务记录	510

表 2 说明每天课时量的划分问题。

表 2 每天课时量划分

周一	周二	周三	周四	周五
T_1	T_6	T_{11}	T_{16}	T_{21}
T_2	T_7	T_{12}	T_{17}	T_{22}
T_3	T_8	T_{13}	T_{18}	T_{23}
T_4	T_9	T_{14}	T_{19}	T_{24}
T_5	T_{10}	T_{15}	T_{20}	T_{25}

每周 5 天涉及 25 个时间片。用 T_1, T_2, \dots, T_{25} 表示,其中 T_1, T_2, T_3, T_4, T_5 为星期一的 5 个时间片,依次类推。这样分主要是排课方便,时间块相等,也是根据大多数高校的教学方案。

实验中各操作数的参数值如表 3 所示。

表 3 操作数参数值

操作数	参数值
杂交概率	0.5
突变概率	0.02
进化代数	100
进化代数	200
种群总数	50 个体
种群总数	100 个体
种群总数	200 个体

经过在不同情况下实验得出进化结果,每次所排出的课表经检测无任何冲突。

当种群数是 100 时,种群中最优个体随着代数的增加而变化,如图 5 所示。

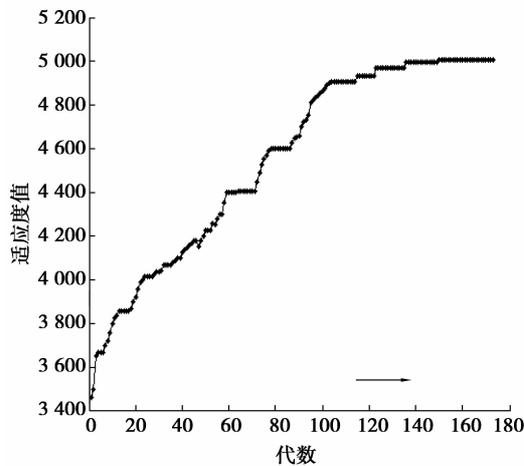


图 5 种群 100 时的变化曲线

当种群数是 50 时,种群中最优个体随着代数的增加而变化的曲线如图 6。

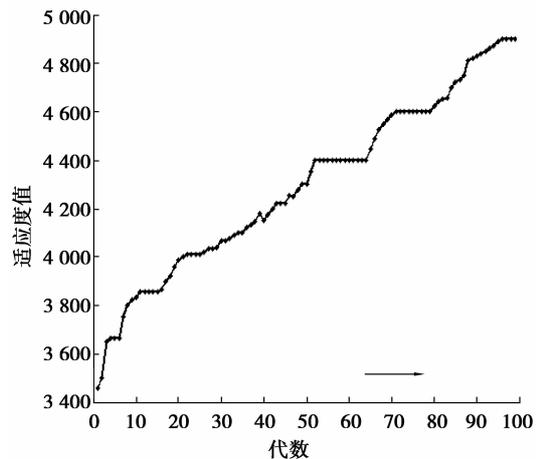


图 6 种群 50 的变化曲线

当种群数是 200 时,种群中最优个体随着代数的增加而变化的曲线如图 7。

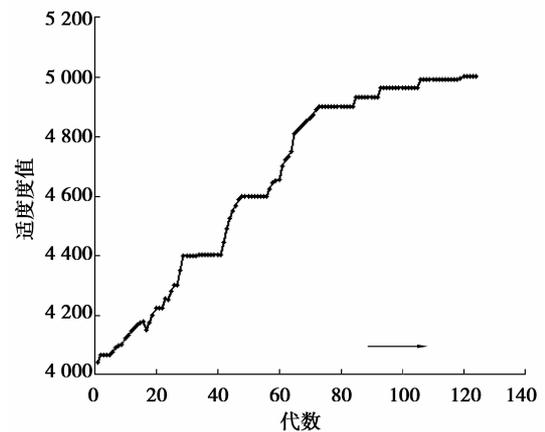


图 7 种群 200 时的变化曲线

当进化代数为 100 时,不同种群总数的 3 种情况的种群中每代最优个体走势做比较,如图 8。

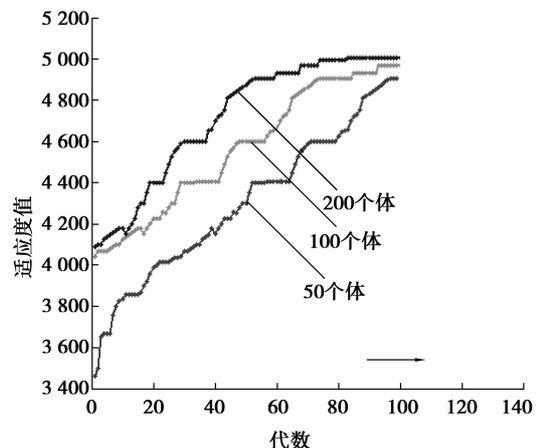


图 8 3 个种群最优个体走势比较

从上述图中可以看出,随着代数的增加,个体的适应度值是逐渐增加,而到了一定代数后就不再增加。个体数较多的种群中最优个体要比个体数较少的具有更高适应度值。新的教室填充的方式使得初始化种群和排除冲突更加快捷,更重要的是,可以跟数据库结合更加精确地满足系统的要求。

4 结 论

研究过程中,充分地分析了遗传算法在排课问题中的应用,对使用遗传算法研究及实现排课系统的编码模式进行了深入地讨论,以对课表进行最大优化为目标,对传统编码模式进行了修正和改进。提出了一套新的二进制编码和教室位置填充方法,精简了染色体信息量,大大缩短了系统的运行时间,最终构建了适合高校排课的遗传算法模型。另外,在系统设计中,运用了基于容器的容器编程技术,实现了不定数据量的便捷处理。模型适用于冲突任务的合理优化安排问题。此模型得出的结果无冲突并且令人满意,但是还有进一步完善的地方,如要充分考虑学生的意愿,选修课的排课等。

参考文献:

- [1] 徐艳斌. 基于遗传算法的高校排课系统设计与分析[D]. 广州:广东工业大学, 2007.
- [2] HIROTO M. Model specification search using genetic algorithm for factor analysis model [J]. Japanese Psychological Research, 2008, 49(3): 179-191.
- [3] SPYROS K, VASSILIOS P, PAVLINA F. Solving university timetabling problems using advanced genetic algorithms [EB/OL]. (2005-10-16) [2009-09-06]. <http://www.teiser.gr/arximidis/pdf/kazarlis/Dhmosieyseis/1.pdf>.
- [4] 陈兴刚, 孟祥婧, 李静. 多约束条件下高校排课系统的算法研究[J]. 电脑知识与技术, 2008, 3(27): 1958-1959.
CHEN XING-GANG, MENG XIANG-JING, LI JING, et al. Research on the algorithms of the course scheduling system for college under the multi-constrained conditions [J]. Computer Knowledge and Technology, 2008, 3(27):1958-1959.
- [5] 廖远, 黄勤, 曹培霞. 基于三维编码的自适应遗传算法在排课系统上的应用[J]. 计算机与现代化, 2008, 33(4):23-24.
LIAO YUAN, HUANG QIN, CAO PEI-XIA. Application of self-adaptive genetic algorithm based on three-dimension coding in curriculum scheduling system [J]. Computer and Modernization, 2008, 33(4):23-24.
- [6] 吴仕勇, 王天志. 基于数值计算方法的遗传算法的优化研究 [J]. 计算机工程与设计, 2009, 30(12): 2966-2967.
WU SHI-YONG, WANG TIAN-ZHI. Optimization research of genetic algorithm based on numerical calculation method [J]. Computer Engineering and Design, 2009, 30(12): 2966-2967.
- [7] SUSSNER P, VALLE M E. A general framework for fuzzy morphological associative memories [J]. Fuzzy Sets and Systems, 2008, 59(7):747-768.
- [8] 曹晓燕, 邵定宏. 基于混合遗传算法的测试数据自动生成研究 [J]. 计算机工程与设计, 2010, 31(2): 4670-4673.
CAO XIAO-YAN, SHAO DING-HONG. Study of automatic test data generation based on hybrid genetic algorithm [J]. Computer Engineering and Design, 2010, 31(2):4670-4673.
- [9] ZWICKL D J. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion [D]. Texas: the University of Texas, 2006.
- [10] 朱玉, 张虹, 孔令东. 基于免疫遗传算法的多维多层关联规则挖掘 [J]. 计算机工程, 2009, 35(3): 181-186.
ZHU YU, ZHANG HONG, KONG LING-DONG. Multi-dimension and multi-level association rule mining based on immune genetic algorithm [J]. Computer Engineering, 2009, 35(3): 181-186.
- [11] 傅平, 罗可. 基于信息熵的免疫遗传算法聚类分析 [J]. 计算机工程, 2008, 34(6): 227-229.
FU PING, LUO KE. Clustering analysis of immune-genetic algorithm based on information Entropy [J]. Computer Engineering, 2008, 34(6): 227-229.
- [12] DARRELL W. A genetic algorithm tutorial [J]. Statistics and Computing, 1994, 4(2): 65-85.
- [13] RAFAL K, TOMASZ A, KENNETH D J. Evolutionary computation and structural design: a survey of the state of the art [J]. Computers & Structures, 2005, 83(23-24):1943-1978.
- [14] PEZZELLA F, MORGANTI G, CIASCETTI G. A genetic algorithm for the flexible Job-Shop scheduling problem [J]. Computers & Operations Research, 2008, 35(10):3202-3212.
- [15] GONCALVES J F, MENDES J J M, RESENDE M G C. A genetic algorithm for the resource constrained multi-project scheduling problem [J]. European Journal of Operational Research, 2008, 189(3):1171-1190.