文章编号:1000-582X(2005)11-0058-04

遗传算法在排课问题中的运用*

江　齐[1],兰　竞[2]
(1.重庆大学 成人教育学院,重庆　400030;2.四川理工学院 招生办公室,四川 自贡　643000)

摘　要:遗传算法借鉴生物界自然选择和遗传机制,使用群体搜索技术,处理传统搜索方法难以解决的复杂的非线性问题.排课问题是一个多因素的优化决策问题,是组合规划中的典型问题,属于NP完全类问题.根据大学课表的特点,采用遗传算法,给出染色体编码和适应度函数,并采用了自适应的调整概率进行排课,数值试验证明了方法的有效性和可行性.

关键词:遗传算法;时间表问题;编码;交叉算子;自适应
中图分类号:TP301.6; G423.06　　**文献标识码**:A

遗传算法(Genetic algorithms, GA)是一类以达尔文自然进化论与遗传变异理论为基础的求解复杂全局优化问题的仿生型算法[1].它模拟生物进化过程,通过向自然学习来求解问题.它是一种借鉴生物界自然选择和自然遗传机制的随机搜索算法,其主要特点是群体搜索策略和种群中个体之间的信息交换.它尤其适用于处理传统搜索方法难以解决的复杂的和非线性的问题.如著名的巡回旅行商问题(Traveling Salesman Problem, TSP)、作业调度问题(Job Shop Scheduling Problem, JSP)、背包问题(Knapsack Problem, KP)、排课问题等.由于它具有这些特点,自20世纪80年代以来关于它的理论和应用研究都成了十分热门的课题,尤其是应用研究显得格外活跃,不仅应用领域扩大了,而且利用遗传算法进行优化和规则学习的能力也显著提高,目前它已被广泛应用于组合优化、机器学习、自适应控制、规则设计和人工生命等领域,被认为是21世纪有关智能计算中的关键技术之一.课表编排是一个多因素的优化决策问题,是组合规划中的典型的NP完全类问题,由于涉及信息较多且求解复杂性是课表规模的指数量级,对有一定规模的问题,一般采用启发式算法,但效果不好.这里研究遗传算法在课表编排中的应用[2].

1　排课问题的现实分析和数学模型

所谓课表,就是课号、课名、教室、学时、教师以及上课班级等信息构成的整体.课表的编排,简单的说就是给课程分配时间和教室,在排课过程中必须严格遵守以下基本原则[3]:

1) 每位教师在同一时间段内只能安排一门课程;

2) 每个自然班在一个时间段内只能安排一门必修课程;

3) 每门课程所用教室的类型和容量要满足该门课程的需求;

4) 每个自然班的每门课程只能安排一次,不能被重复安排.

一门课程的一次教学工作称为一项教学任务.所有教学任务得以顺利进行所必须的教室、班级、教师、实验室、体育场以及教学时间等称为教学资源.任何违反以上4条基本原则的排课都会导致2项(或以上)教学任务争夺某一教学资源,使教学工作不能正常进行,这一现象称为冲突.2门(或以上)课程若关联着相同的教师或上课班级,称为冲突课程.

为了使排出的课表更优化、合理,排课还应考虑以下因素:

* 收稿日期:2005-06-02
作者简介:江齐(1969-),四川成都人,重庆大学助理研究员,主要从事教学管理及教学研究.

1) 尽可能使每门课程在一周内的上课时间分布合理;

2) 尽可能使学生在连续的两门课之间更换教室的机率小;

3) 尽可能把同一课程的不同讲次安排在同一教室;

4) 尽可能使学生每天的必修课趋于平衡;

5) 尽可能满足每个课程教学的客观要求;

6) 满足个别教师的特殊上课时间要求.

大学课程一类为学院根据专业教学计划和教学大纲要求,针对各个专业年级安排的必修课、专业选修课和限选课;另一类为全校性选修、辅修课,这种课打乱专业及年级限定,根据选课人数安排教学. 为了叙述方便,笔者把前一类课称为C类课,后一类叫E类课.

C类课由各学院管理,每个班某一学期开设的课是按要求指定的,并且给学生上课的教师基本上是固定的,排课的关键是把教师所讲课程制定到周课表的合适时间;E类课是教务科管理,根据各学院上报的开课计划汇总后下发到学院供全校学生登记选课. 两类课目标不同,后一种可在学校内部一般采用固定时间和教师分配固定教室再由学生自己根据C类课的上课情况以及个人的实际情况自由选择达到优化的目的,故只对C类课采用算法解决方案,这样做的好处是编码既能更好地贴近问题的描述,也符合当前大学授课的实际情况. 对C类课笔者提出的算法分为2个阶段[4]:第1阶段产生初始种群(可能包含冲突),使用遗传算法将其可行化;第2阶段从这个可行的初始种群开始,使用遗传算法进一步优化,以占用好的教学时间点且一周内同一门课分散作其适应度函数. 若第1阶段可行解未达到种群数,这说明求解空间内的可行解较少,这里不作重点讨论,但为了使问题讨论更加全面,笔者将冲突作为惩罚函数定义为适应度函数的方法给予解决.

2 排课问题的实现与优化

2.1 编码及其染色体表示

由于考虑的是周课表,并把每天用于上课的时间划分为5个时间片,上午8:00-10:00,10:00-12:00;下午2:00-4:00,4:00-6:00;晚上7:00-9:00. 假设周一到周五白天上C类课,剩下的时间自由安排E类课,则C类课的时间片为$5\times4=20$个,用$T_1,T_2,T_3,\cdots,T_{20}$表示,其中每4个为一天的时间片. 把班级和教室作为一个变量来对待. 如考虑的班有i个,则对应的教室编号为$R_1,R_2,\cdots,R_i$. 对于一个班级而言,由于这学期开的课程和教师是相对固定的,可以把课程和教师当作同一变量来对待. 如在课表上有5位老师,就给其编号1,2,3,4,5,同一编号教师可以在一个或多个班级出现,可以在一个班级上一门或多门课,因此一名教师一周可以有多次上课,他的编号出现次数为所有上课次数的总和. 通过以上把班级与教室等同、课程与教师等同的处理后,原课表五要素转化为三要素(班级、课程、时间). 把班级固定要上的课程看作是基因,用教师代码或0表示(0表示无课). 而染色体在这里就是N个班级和20个时间点组成的二维表. $N=6$时的教师代码如表1所示.

表1 班级与时间对应的教师代码

时间	教师代码					
	R_1	R_2	R_3	R_4	R_5	R_6
T_1	1	3	5	0	4	2
T_2	3	1	0	4	7	9
T_3	4	12	23	6	8	14
T_4	2	4	3	1	6	13
T_5	3	5	9	16	21	17
T_6	5	1	6	17	23	0
T_7	11	0	25	0	0	5
T_8	8	5	24	37	6	12
T_9	0	0	1	0	29	0
T_{10}	2	11	1	12	8	3
T_{11}	4	2	8	9	7	6
T_{12}	5	6	9	2	3	4
T_{13}	11	14	2	4	17	35
T_{14}	23	19	51	0	0	18
T_{15}	6	7	8	9	1	2
T_{16}	3	4	10	12	13	15
T_{17}	0	9	8	7	5	3
T_{18}	0	16	13	4	2	1
T_{19}	9	15	17	35	21	25
T_{20}	0	31	24	2	3	8

2.2 可行初始种群的产生

随机将各班级任课教师编码安排到20个时间点,从而产生一个染色体. 将随机产生的初始种群通过遗传操作,把适应度函数定义为冲突个数,即

$$f=\sum_{c\in C_j}v(c),$$

其中C_j为第j类约束集.

$$v(c)=\begin{cases}0, & 若 c 约束被满足;\\ 1, & 若 c 约束被违反.\end{cases}$$

保留其中适应度为0的个体到种群,此过程重复进行,直到可行解个数达到初始种群数,再进入第2阶

段,寻求近优解,如在指定最大迭代次数内未产生可行解则改变初始数据,若可行解未达种群数则定义惩罚函数

$$f = \frac{1}{1 + \sum_{约束类型 j} w_j \sum_{c \in C_j} v(c)},$$

其中 w_j 为违反第 j 类约束的惩罚值.

$$v(c) = \begin{cases} 0, & 若 c 约束被满足; \\ 1, & 若 c 约束被违反. \end{cases}$$

后一种情况不是考虑的重点,但按照笔者的方法亦可得出结论.

2.3 遗传操作

采用"轮盘赌"选择方法,它由函数 select(…)实现.同时采用 PMX 交叉算子,这种算子较简单地直接交换扩大了搜索空间,还使交叉后的非法结合法化.其过程是随机选定 2 个配对个体,然后随机选出交叉点,该点即为染色体的一列,在这列上随机选择 2 点,交换其中间部分,然后修改前后两部分使其合法化.遗传算法中,交叉算子因其全局搜索能力而作为主要算子,变异算子因其局部搜索能力而作为辅助算子.遗传算法通过交叉和变异这一对相互配合又相互竞争的操作而使其具有兼顾全局和局部的均衡搜索能力.变异操作基本步骤如下:

1) 在所有个体的码串范围内随机确定基因座;

2) 以事先确定的变异概率 P_m 来对这些基因座的基因值进行变异,如表 2 和表 3 所示.

表 2 变异前班级与时间对应的教师代码

时间	教师代码				
	R_1	R_2	R_3	…	R_{10}
T_1	3	4	5	…	6
T_2	1	2	3	…	8
T_3	0	11	21	…	9
⋮	⋮	⋮	⋮	⋮	⋮
T_{20}	4	7	2	…	12

表 3 变异后班级与时间对应的教师代码

时间	教师代码				
	R_1	R_2	R_3	…	R_{10}
T_1	3	11	5	…	6
T_2	1	2	3	…	8
T_3	0	4	21	…	9
⋮	⋮	⋮	⋮	⋮	⋮
T_{20}	4	7	2	…	12

杂交概率 P_c 和变异概率 P_m 的选取很大程度上影响算法的收敛速度和近优解的质量,笔者希望每一当前代的最优个体不参加杂交和变异,较优个体多参加杂交和变异,以便加快算法的搜索效率和有效防止陷于局部最优解,从而采用自适应的调整概率[5]:

$$P_c = \begin{cases} k_1 \sin\left(\frac{\pi}{2} \times \frac{f_{max} - f_c}{f_{max} - f_{avg}}\right), & 当 f_c \geq f_{avg}; \\ k_2, & 当 f_c < f_{avg}. \end{cases}$$

$$0 < k_1, k_2, k_3, k_4 \leq 1.$$

$$P_m = \begin{cases} k_3 \sin\left(\frac{\pi}{2} \times \frac{f_{max} - f_m}{f_{max} - f_{avg}}\right), & 当 f_m \geq f_{avg}; \\ k_4, & 当 f_m < f_{avg}. \end{cases}$$

其中,f_{max} 是群体的最大适应度函数值,f_{avg} 是平均适应度函数值,f_c 是杂交的两染色体串中适应度较大者,f_m 是变异串的适应度函数值.当 $f_c = f_{max}$,$f_m = f_{max}$ 时概率为 0,当前代可以直接转入下一代.采用过滤算法排除冲突,即排除同一时间一个班上多门课,同一时间一个教室上多门课,教室容量不足等冲突现象,但它可能发生同一教师在同一时间为 2 个获益班上课的情况,表现在染色体中为同一行有相同的编码.虽然初始种群均为可行解,但这些可行解经过交叉变异操作后可能会产生不可行解.过滤操作就是对交叉变异后的个体及时检查并进行修复,从而保证每个个体的可行性.由于父代的无冲突性,交叉后的冲突检查只需对 2 个子代交叉所在列进行,修复操作也在该列进行.下面给出具体步骤:

step1 选定 2 个父代;

step2 jnum = 0(控制列的选择次数);

step3 1 ~ n 间产生随机数字 j, jnum = jnum + 1, 用 pmx 算法对第 j 列杂交生成子代;

step4 子代有冲突,执行步骤 5,排除冲突,否则返回该子代;

step5 inum = 0(控制行的选择次数);

step6 1 ~ m 间产生随机数字 i; inum = inum + 1;

step7 第 j 列,是冲突元素与第 i 行对应元素互换;

step8 若子代冲突 &&inum ≤ m,则执行步骤 6,否则执行步骤 9;

step9 若子代冲突 &&jnum ≤ n,则执行步骤 3,否则执行步骤 10;

step10 若子代冲突,执行步骤 1,否则返回该可行子代.

n 为班级总数, m 为时间片数.

对于一个给定的优化问题[6],如何设计构造适应度函数是一个很关键的问题,这需要对问题有深入的

了解. 在过滤算法之后，产生的肯定是一个可行解，一个好的课表应更多地满足模型中提出的经验常识，因此在第 2 阶段的适应度函数中应考虑下面 2 点：课占用更好的时间点；同一教师带课的离散化. 以 $g(t(s_i))$ 表示时间片的期望值，$h(s_i)$ 表示同一门课程的离散度，不妨设定：

$$f = \sum_i g(t(s_i)) + h(s_i).$$

时间点与两课时间差所对应的期望值见表 4、表 5.

表 4　不同时间点的期望值

时间点	$T_1,T_5,T_9,T_{13},T_{17}$	$T_2,T_6,T_{10},T_{14},T_{18}$	$T_3,T_7,T_{11},T_{15},T_{19}$	$T_4,T_8,T_{12},T_{16},T_{20}$
期望值	12	8	4	1

表 5　时间差与期望值的对照

两课时间差	1,19	2,3,17,18	4,5,15,16	6,7,13,14	8,9,10,11,12
期望值	0	2	4	9	16

2.4　数值结果

为了直观地验证排课结果，笔者取了真实的数据作出部分改动. 时间片仍是前面提到的 20 个，用 $T_1, T_2, \cdots, T_{20}$ 表示. 任意选取 7 个教学班($R_1, R_2, \cdots, R_7$)，教学班固定要上的课，用教师代码表示(即 1,2,…,33 分别代表 33 位教师所授的 33 门课)，7 个班上课教师分别为：

$R_1 = (1,2,2,3,3,4,4,5,5,6,6,8,8,9,10)$,

$R_2 = (11,17,18,19,19,19,20,21,22,22,23)$,

$R_3 = (11,24,24,25,25,27,27,28,29,29)$,

$R_4 = (9,11,24,24,25,25,27,27,32,32)$,

$R_5 = (9,11,26,26,29,29,31,31,33,33)$,

$R_6 = (11,26,26,28,29,29,30,30,33,33)$,

$R_7 = (2,2,4,4,7,7,8,8,10,11)$.

取种群数为 20，最大迭代数为 1 000 代，对 k_1、k_2、k_3、k_4 取不同的值，测试其给定代数内近优解的最大适应度值以及达到此适应度值的迭代次数，如表 6.

表 6　不同 k 值所对应的迭代次数

组数	k_1	k_2	k_3	k_4	近优解值	迭代次数
1	1.00	1.00	0.50	0.50	714	810
2	0.80	0.80	0.85	0.85	692	530
3	1.00	1.00	0.60	0.60	681	420
4	0.90	0.90	0.50	0.50	680	400
5	0.90	0.90	0.60	0.60	685	220
6	1.00	1.00	0.80	0.80	734	750
7	0.50	0.50	0.80	0.80	673	760

取第 6 组在 750 代时出现近优解值 734，得出最优排课表如表 7(其中"0"表示无课).

表 7　7 个班的最优课表

时间	教师代码						
	R_1	R_2	R_3	R_4	R_5	R_6	R_7
T_1	0	21	25	0	0	28	0
T_2	0	19	11	0	31	26	0
T_3	3	0	0	11	0	0	0
T_4	4	19	24	32	0	0	0
T_5	5	23	0	0	26	33	4
T_6	8	0	29	27	9	0	7
T_7	6	0	25	9	11	29	0
T_8	0	0	0	0	0	0	0
T_9	5	20	27	0	33	30	2
T_{10}	4	0	0	25	31	0	8
T_{11}	2	19	0	32	0	0	4
T_{12}	0	22	0	24	0	29	0
T_{13}	3	22	29	0	33	26	10
T_{14}	8	18	28	0	29	11	2
T_{15}	6	0	0	0	0	0	11
T_{16}	1	11	0	24	0	0	0
T_{17}	10	0	27	25	29	30	8
T_{18}	0	17	0	27	26	0	7
T_{19}	9	0	29	0	0	0	0
T_{20}	2	0	0	0	0	33	0

3　结　语

排课问题是一个多因素的优化决策问题，是组合规划中的典型问题. 笔者采用遗传算法给出了染色体编码和适应度函数，能较好地反映排课的需求. 同时采用自适应的调整概率，在循环计算过程中，根据排课过程应遵守的 4 个基本原则及应考虑的 6 个因素等优化原则得出最优排课表.

参考文献：

[1]　刘勇，康立山，陈毓屏. 非数值并行算法—遗传算法[M]. 北京：科学出版社，1998.

[2]　陈国良，王煦法，庄镇泉，等. 遗传算法及其应用[M]. 北京：人民邮电出版社，1996.

[3]　吴金荣. 关于大学课程表问题的研究[J]. 运筹与管理，2002，11(6)：66－71.

[4]　李增智，王云岚，陈靖. 课程表问题的一种混合型模拟退火算法[J]. 西安交大学报，2003，37(4)：343－350.

[5]　张春梅，行飞. 用自适应算法求解大学课表问题[J]. 内蒙古大学学报，2002，33(4)：459－464.

[6]　王凌. 智能优化算法及其应用[M]. 北京：清华大学出版社，2001.

(下转第 72 页)

Method of Computing Sinusoidal Steady-state Solutions of Lossy Uniform Transmission Lines

SUN Tao[1], *LIU Zong-hang*[2], *JIANG Ze-jia*[1]

(1. College of Electrical Engineering;

2. College of Communication Engineering, Chongqing University, Chongqing 400030, China)

Abstract: There is no universal method of finding the analytic solutions to transmission lines discribed by partial differential equations, so many researchers are studying and developing transmission line theories. Computing steady-state solutions of uniform transmission lines is one part of the study. The paper introduces another method of computing sinsoidal steady-state solutions of lossy uniform transmission lines. First, the complex expressions of voltage and current with zero initial state are obtained from the complex frequency-domain model of lossy uniform tansmission lines. The network functions, which are the ratios of voltage and current's image functions to the excitation's image function, can be found from the complex expressions. Sinusoidal steady-state solutions can be obtained by using the relation between network function and system's frequency characteristic. Finally, the method is demonstrated to be effective by an example.

Key words: lossy uniform transmission lines; network function; frequency characteristic; sinusoidal steady-state solutions

(编辑 李胜春)

(上接第61页)

Application of the Genetic Algorithm in Timetable Problem

JIANG Qi[1], *LAN Jing*[2]

(1. College of Further Education, Chongqing University, Chongqing 400030, China;

2. Recruiting Office, Sichuan University of Science&Engineering, Zigong, Sichuan 643000, China)

Abstract: Genetic Algorithm is based on the biological mechanism of natural selection and heredity, leveraging colony searching technology, and is particularly applicable for the resolution of complicated non-linear problems intractable with traditional searching methods. Timetable problem is a multi-factor optimized decision problem and is typical problem in constitution and planning. It has been proved as a kind of NP-complete problem. According to the character of courses assignment in an university, a kind of codes and fitness function are designed and solved by Genetic Algorithm. With adaptive crossover and mutation probability employed, the experiment verifies that this method is both efficient and effective for the problem.

Key words: Genetic Algorithms; timetable problem; codes crossover operator; adaptive

(编辑 张 苹)