

遗传算法的工程应用

肖俊

(上海交通大学软件学院 上海 200135)

摘要 分析了排课问题的各种约束条件,并研究了用遗传算法解决排课问题,给出了一个基于该算法的排课模型,并对涉及的各种问题进行探讨。

关键词 遗传算法,组合最优化,排课

Application of Genetic Algorithms

XIAO Jun

(Department of Software Academy, Shanghai Jiao Tong University, Shanghai 200135)

Abstract In this thesis an optimized scheme for course-arranging has been set up on the base of applying Genetic algorithm, discusses almost all the problems involved in the scheme and gives some analysis on the performance of it.

Keywords Genetic algorithms, Combinatorial optimization problem, Course-arranging problem

1 引言

排课调课是高校教务管理的一项日常工作,因其费时费力,手工完成多易出错,所以利用计算机进行自动排课的想法自然而生。我国也对其进行研究,才用过模拟手工排课、矩阵法等等。但目前大部分高校的排课工作却仍以手工为主,主要是由于排课问题约束条件较多且复杂,要考虑教师、教室、课程分布、时间分配、分合班、单双周、教师要求等多方面约束。

课程表问题又称为时间表问题,是典型的组合优化和不确定性调度问题,是解决资源调配引起的冲突。

20世纪70年代课表问题已经被论证为NP完全问题,解决这类问题的近似算法分为启发式方法和随机算法。启发式方法依赖于对问题的构造和性质的认识和经验,只能解决一些比较简单的问题。使用随机算法则比较耗时,而且在大量的搜索空间容易遗漏最优解。遗传算法则处于两者之间,作为现代最优化手段,被证明应用于大规模、含离散变量等情况下的全局最优化问题是最合适的,因此考虑用遗传算法解决排课问题,寻求一种新的解决排课问题的方法。

2 数学描述

排课系统中将包括以下的资源集合:

课程集合(Lesson): $L = \{L_1, L_2, \dots, L_n\}$

班级集合(Class): $C = \{C_1, C_2, \dots, C_a\}$

教师集合(Professor): $P = \{P_1, P_2, \dots, P_\beta\}$

时间表集合(Time): $T = \{T_1, T_2, \dots, T_\gamma\}$

教室集合(Room): $R = \{R_1, R_2, \dots, R_\delta\}$

其它集合: $Q = \{Q_1, Q_2, \dots, Q_e\}$

其中课程是最关键的资源,它包括6个属性:①周学时,决定一周课程要安排几次;②班级,上该课程的所有班级,可以有多个班级组成,即合班课,排课时就要牵涉到其它班级的时间;③授课老师,会有多个老师教授一门课程,比如实验课,排课时要考虑到这几个老师的排课冲突问题;④其他条件,教师

和教室安排的时间限定,以及其他的优先级方面的条件;⑤安排的时间;⑥安排的教室。前面四个为已知条件。

即:课程 $L = (H, C, P, Q, L, R)$

排课要遵循很多约束条件,可以分为两大类,硬约束和软约束。其中,硬约束为必须满足的条件,而软约束是为了让排出的课表更合理、更人性化,只要尽量满足。

硬约束包括:

- 教师授课时间冲突,即一个老师在一个时间段只能授一门课程。
- 教室授课时间冲突,即一间教室在一个时间段只能开一门课程。
- 班级上课时间冲突,一个班级在一个时间段只能上一门课程。
- 教室容纳人数必须大于开设课程的所选人数。
- 由于其他条件产生的约束条件,某时间段规定某教室不给排课;某些课程必须排在早上等等,硬性规定的约束条件。

以上条件必须全部遵守,才能排出一个正确的课表。

软约束包括:

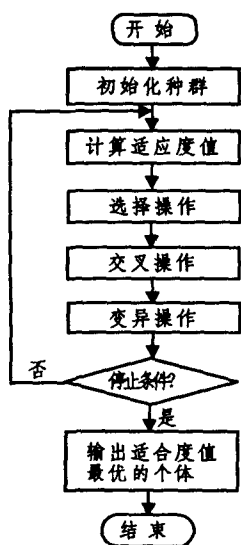
- 根据班级人数选择容量合适的教室,使最合理的利用教室资源。
- 一门课程如果一周上多次课,希望能隔开上。
- 满足教师教案的周期性。
- 等等。

由于学校的实际情况不同,软约束差别也很大。有些学校路途遥远,老师的课程要集中上,而有的学校老师希望分开来上课。软约束的种类繁多且很少有共同点,我们暂时列出以上几点。

3 算法模型

遗传算法类似于自然进化,通过作用于染色体上的基因寻找好的染色体来求解问题。与自然界相似,遗传算法对求解问题的本身一无所知,它所需要的仅是对算法所产生的每

个染色体进行评价,并基于适应值来选择染色体,使适应性好的染色体有更多的繁殖机会。在遗传算法中,通过随机方式产生若干个所求解问题的数字编码,即染色体,形成初始群体;通过适应度函数给每个个体一个数值评价,淘汰低适应度的个体,选择高适应度的个体参加遗传操作,经过遗传操作后的个体集合形成下一代新的种群。对这个新种群进行下一轮进化。



3.1 编码

遗传算法中的编码即把问题的可行解从解空间转换到算法能处理的搜索空间的转换方法。编码是应用遗传算法要解决的首要问题,是设计中的关键步骤,它将影响到后面交叉、变异算子等运算方法,很大程度上决定了如何进行群体的遗传进化运算以及遗传进化运算的效率。

二进制编码方式是遗传算法最常用的一种编码方式,但是,由于它不能直接反映问题的固有结构,加之精度不高,不适于排课中,因此在该算法中,我们采用多维实数编码方式,它可以直接在解的表现型上进行遗传操作,减低无效交叉发生的可能性。

3.2 适应度函数

遗传算法使用适应度在度量群体中个体的优良程度,适应度较高的个体遗传到下一代的概率较大。而度量个体适应度的函数即适应度函数。因此,适应度函数是用来区分群体中个体好坏的标准,是自然选择的唯一标准,选择的好坏直接影响算法的优劣。

通过上面算法描述中的软约束条件,我们可以得到下面的适应度函数:

$$F = 1 / \sum_{j=1}^n (\tilde{w}_j Cost_1(j) + \tilde{w}_j Cost_2(j) + \tilde{w}_m Cost_3(j) + \dots)$$

其中, W_i 为各罚值的权值,可以根据各个学校的情况调整加权系数, $Cost(j)$ 表示各个软约束的费用。

3.3 控制参数

3.3.1 选择 选择算子体现了个体的优胜劣汰操作,适应度的高低影响遗传到下一代的概率。遗传算法中的选择操作用来确定如何从父代群体中按某种方法选取哪些个体遗传到下一代群体中。

选择操作建立在对个体的适应度进行评价的基础上。选择操作的主要目的是为了避免基因缺失、提高全局收敛性和计算效率。

我们采用无回放余数随机选择,按照一种确定的方式进行选择操作。

1. 计算群体中各个个体在下一代群体中的期望生存数目 N_i :

$$N_i = M \cdot F_i / \sum_{i=1}^M F_i \quad (i=1, 2, \dots, M)$$

2. 用 N_i 的整数部分确定各个对应个体在下一代群体中的生存数目。这样可确定出下一代 M 个群体中的 $\sum_{i=1}^M \lfloor N_i \rfloor$ 个个体。

3. 以 $F_i - \lfloor N_i \rfloor \cdot \sum F_i / M$ 为各个个体的新适应度,用比例选择方法来随机确定下一代群体中还未确定的 $M - \sum_{i=1}^M \lfloor N_i \rfloor$ 个个体。

这种方法可保证适应度比平均适应度大的一些个体一定能够被保留在下一代群体中,选择误差比较小。

3.3.2 交叉 选择的效果提高了群体的平均适应值,但是损失群体的多样性,最好个体的适应值将无法得到改进。因此,遗传算法效仿生物的自然进化,两个同源染色体通过交配而重组,产生新的染色体的方法,形成交叉算法。交叉运算是产生新个体的主要方法,它决定了遗传算法的全局搜索能力。

排课系统中,由于牵涉到课表的合法性,无法使用单点交叉和双点交叉,可以采用交叉时间段的方法,同时,由于是一个班级内部进行交叉,必定会影响到其它班级的课程。因此,交叉运算同时要要进行冲突判断,以生成一个合法的课表。

定义交换概率 Γ , 交换时产生一个随机数 $r(0 < r < 1)$, 当 $r < \Gamma$ 的时候, 执行交换操作, 否则不执行。

3.3.3 变异 变异算法, 指将个体染色体编码串中的某些基因座上的基因值用该基因座的其他等位基因来替换, 从而形成一个新的个体。变异算法只是产生新个体的辅助方法, 但是它决定了算法的局部搜索能力。因此, 变异算子和交叉算子需要相互配合, 共同完成对搜索空间的搜索, 从而提高算法的搜索性能。

排课系统中由于课表合法性的问题, 采用简单的随机变异, 会出现大量的冲突, 这样耗时太多。因此, 我们考虑使用列交换, 这样交换后的课表不用纠错。定义变异概率 Ψ , 在交换时产生一个随机数 $r(0 < r < 1)$, 当 $r < \Psi$ 时, 执行变异操作, 否则不执行。

3.3.4 算法终止条件 采用的终止条件为: 用户满足已搜索到的最优解或系统已达到稳定。

结束语 本文提出的利用遗传算法解决优化排课问题的方案, 很好地解决了多约束下的排课问题。该算法的实现将极大地减轻教务管理人员的劳动强度, 实现教务管理工作的自动化。

参考文献

- 1 李明杰, 曾晋文. 课表编排系统的算法分析与设计. 计算机工程与设计, 2004, 25(10)
- 2 朱冠宇, 等. 利用遗传算法求解中学课表安排问题. 计算机工程与应用, 2004, 40
- 3 唐勇, 唐雪飞, 王玲. 基于遗传算法的排课系统. 计算机应用, 2002, 22(10)
- 4 周明, 孙树栋. 遗传算法原理及应用. 北京: 国防工业出版社, 1999. 6