

文章编号: 1001-0920(2008)12-1427-05

基于粗糙规划的不确定加工时间的并行机调度

于艾清, 顾幸生

(华东理工大学 自动化研究所, 上海 200237)

摘要: 针对并行机调度中的不确定工件加工时间, 提出用粗糙变量表示不确定量, 并由此建立该问题的粗糙期望值规划模型. 提出一种应用于调度问题的进化规划算法, 改进了针对并行机问题的编码方式和变异方法. 采用粗糙模拟的方法计算个体的适应值, 即粗糙期望估计值, 并加以不同规模的算例进行仿真实验. 仿真结果表明, 改进进化规划算法得到的解优于遗传算法得到的解.

关键词: 调度; 并行机; 不确定性; 粗糙模拟; 进化规划

中图分类号: TP391.9 **文献标识码:** A

Parallel machines scheduling with uncertain processing time based on rough programming

YU Ai-qing, GU Xing-sheng

(Research Institute of Automation, East China University of Science and Technology, Shanghai 200237, China.

Correspondent: YU Ai-qing, E-mail: yuaiqing@mail.ecust.edu.cn)

Abstract: Considering the uncertain processing time in scheduling problems, rough variables are employed to address it and a rough expected value model for parallel machines scheduling is presented. In addition, an improved evolutionary programming algorithm for scheduling problems is proposed by redefining representation structure of individuals and mutation operator. To evaluate the fitness of individuals, rough simulation is adopted to calculate the expected values of objectives. Computational results of different scales of problems show that the rough simulation based evolutionary programming algorithm is efficient.

Key words: Scheduling; Parallel machines; Uncertainty; Rough simulation; Evolutionary programming

1 引言

目前针对并行机调度问题的研究大都建立在确定的生产环境的基础上, 然而在现实生产系统中存在机器因素、人力因素和环境因素的影响, 致使包括时间参数在内的相关信息无法准确预知, 存在不确定性. 因此, 根据调度环境的特点, 可分为确定性和不确定性调度^[1]. 处理这些不确定的信息一般有如下 3 种基本方法: 1) 随机概率论; 2) Zadeh^[2] 提出的模糊集理论; 3) Pawlak^[3] 提出的粗糙集理论.

目前, 很多学者采用随机概率论和模糊集理论来处理不确定信息, 它们分别根据统计学中的概率分布和模糊集理论中的隶属度或隶属函数对变量进行描述, 并建立随机规划和模糊规划模型^[4,5]. 但在实际生产过程中, 一些不确定信息的概率分布或隶

属度不易获得. 粗糙集(RS)理论是逐渐引起人们关注的一类处理不确定信息的新方法, 它是 20 世纪 80 年代初由波兰科学家 Pawlak^[3] 提出的, 从一个新的角度将知识定义为对论域的划分能力, 并将其引入数学中的等价关系进行讨论, 从而为数据分析, 特别是不精确、不完整数据分析提供了一套新的数学方法. 它的特点是不需预先给定默认特征或属性的数量描述, 而是直接从给定的问题的描述集合出发, 通过不可分辨关系和不可分辨类确定给定问题的近似域, 找出问题的内在规律. 由于它的这些特色, 粗糙集理论已成功地应用于决策分析、决策支持、数据挖掘、知识发现以及模式识别等诸多领域, 并取得了许多成果^[6,7]. 目前, 研究将粗糙集理论用于解决不确定的并行机调度问题的文献还较少. 作

收稿日期: 2007-10-05; 修回日期: 2007-12-14.

基金项目: 国家自然科学基金项目(60674075, 60774078).

作者简介: 于艾清(1981—), 女, 上海人, 博士生, 从事生产计划与调度、智能优化算法等研究; 顾幸生(1960—), 男, 江苏海门人, 教授, 博士生导师, 从事控制理论与应用、复杂工业过程建模等研究.

为一种新的处理不确定信息的方法,本文提出用粗糙集描述并行机调度问题中的不确定工件加工时间.

本文所考虑的并行机调度中的不确定参数是工件加工时间,用粗糙变量的上近似和下近似表示,由此建立并行机调度问题的粗糙期望值模型.针对模型的求解,改进了进化规划算法,提出了针对并行机问题的编码方式和变异方法,用粗糙模拟的方法计算个体的适应值,并加以不同规模的算例进行仿真实验.

2 粗糙集信赖性理论

作为一种公理化方法,信赖性理论是研究粗糙现象的数学分支.刘宝碇^[8]建立了信赖性理论的公理基础,给出了4条公理和三元组的测度空间.为了建立粗糙决策系统的优化理论,刘宝碇^[9]引进了一系列概念,如粗糙空间、粗糙变量、期望值算子等.

定义 1^[9] 设 \mathcal{A} 为非空集合, \mathcal{A} 是由 \mathcal{A} 的子集构成的代数, α 为 \mathcal{A} 中的一个元素, f 为定义在 \mathcal{A} 上的满足4条公理的集函数,则称四元组 $(\mathcal{A}, \alpha, f, \mathcal{A})$ 为一个粗糙空间.

定义 2^[9] 一个粗糙变量 f 是从粗糙空间 $(\mathcal{A}, \alpha, f, \mathcal{A})$ 到实数集的一个可测函数,即设 $(\mathcal{A}, \alpha, f, \mathcal{A})$ 为一个粗糙空间, f 是从 \mathcal{A} 到实数集 R 的函数.若对 R 的每个 Borel 集 B , 有

$$\{ \alpha(A) \mid A \in B \} \in \mathcal{A}, \quad (1)$$

则称 f 为粗糙空间 $(\mathcal{A}, \alpha, f, \mathcal{A})$ 的粗糙变量.更进一步,有

$$\underline{f} = \{ \alpha(A) \mid A \in \mathcal{A} \}, \quad \bar{f} = \{ \alpha(A) \mid A \in \mathcal{A} \}, \quad (2)$$

分别称为粗糙变量 f 的下近似和上近似.

为了估计信赖性测度,计算期望值,刘宝碇还提出了粗糙模拟的概念.

定义 3^[9] 设 f 为定义在粗糙空间 $(\mathcal{A}, \alpha, f, \mathcal{A})$ 上的 n 维粗糙变量, $f: \mathcal{A} \rightarrow R$ 为可测函数.为了计算期望值 $E[f(\cdot)]$,按测度 α 分别从 \mathcal{A} 中抽样产生 $\alpha_1, \alpha_2, \dots, \alpha_N$, 从 \mathcal{A} 中抽样产生 $\alpha_{-1}, \alpha_{-2}, \dots, \alpha_{-N}$, 则当 N 充分大时,期望值 $E[f(\cdot)]$ 的估计值为

$$E[f(\cdot)] = \frac{\sum_{i=1}^N (f(\alpha_i) + f(\alpha_{-i}))}{2N}. \quad (3)$$

3 问题描述

3.1 并行机调度问题的假设

并行机调度问题一般具有以下形式:给定 m 台机器和 n 个工件 ($n > m$), 每个工件可在任一机器上完成加工.问题所追求的目标是使所有工件的总完工时间 (makespan) 最短,即怎样分配每台机器上所加工的工件及加工顺序,使所有工件的总完工时

间最小.

首先给出如下假设:

假设 1 每个工件只有一个加工工序,可在任何一台机器上加工;

假设 2 每台机器在同一时刻只能加工一个工件;

假设 3 所有工件在 0 时刻可以启动;

假设 4 工件加工时间为粗糙变量,用上、下近似区间表示.

3.2 粗糙并行机调度问题的数学描述

本文研究的是不确定条件下的并行机调度问题,主要指工件处理时间上的不确定性.用粗糙变量 $r = (x, \bar{r})$ 表示工件处理时间,在实际生产中所表达的含义是:工件处理时间最小范围在下近似区间 x 内,最大范围在上近似区间 \bar{r} 内.

符号表示如下: $j = 1, 2, \dots, m$ 为机器编号; $t_{x,j}(x, y, r)$ 为工件 x_i 在机器 j 上的加工时间,是粗糙变量; $x = (x_1, x_2, \dots, x_m)$ 为 n 个工件的排序,是整数变量; $y = (y_1, y_2, \dots, y_m)$ 为每台机器上的加工工件数,总和为 n ,是整数变量.

由此, n 个工件 m 台机器的并行机调度问题的完工时间为

$$C_{x_{l+i}j}(x, y, r) = t_{x_{l+i}j}(x, y, r), \quad l = \sum_{k=1}^j y_{k-1}, y_0 = 0; \quad (4)$$

$$C_{x_{l+i}j}(x, y, r) = C_{x_{l+i-1}j}(x, y, r) + t_{x_{l+i}j}(x, y, r), \quad 2 \leq i \leq y_j. \quad (5)$$

式(4)表示第 j 台机器上第 1 个加工工件的完工时间,其中 l 为工件 x_{l+i} 的下标,表示在第 i 个工件加工前,已经被加工的工件个数.式(5)表示第 j 台机器上第 i 个加工工件的完工时间. makespan 则为

$$C_{\max} = \max(C_{x_{l+y_j}j}(x, y, r)),$$

目标函数为 $\min(\text{makespan})$.

3.3 粗糙并行机调度问题的期望值规划模型

根据 3.2 节给出的粗糙并行机调度问题的数学描述,建立该问题的粗糙规划模型如下:

$$\begin{aligned} \min & E(C_{\max}). \\ \text{s.t.} & 1 \leq x_i \leq n, \quad i = 1, 2, \dots, n; \\ & y_j = n, \quad 1 \leq y_j < n, \\ & j = 1, 2, \dots, m; \\ & x_i, y_j \in Z. \end{aligned} \quad (6)$$

目标函数为 makespan 的粗糙期望值,满足约束条件: x_i 为区间 $[1, n]$ 内的任一整数变量; y_j 为区间 $[1, n]$ 内的任一整数变量,其和为 n .

4 改进进化规划算法及其在并行机调度中的应用

进化规划 (EP) 算法最初是由美国的 Fogel 等^[10] 在 20 世纪 60 年代首先提出的,当时并未引起足够的重视. 30 多年后,文献^[11] 改进了这种方法,从而使其作为进化算法的一个分支在实际中得到了广泛应用.

最近几年,进化计算方法在调度问题中的应用已成为调度理论研究的热点^[12]. 对于这类问题,大多采用遗传算法求解,但近年来进化规划已被广泛用于求解此类问题,并取得了比较满意的结果. EP 也是模拟自然界中生物的进化规律,在优胜劣汰的竞争环境中不断进化,逐渐逼近最优解. 与遗传算法着重于对染色体的算子研究不同,进化规划更注意个体和种群行为上的进化,即强调父代和子代行为上的联系,而不是基因方面的联系. EP 的显著特征是没有杂交过程,其适应值用于指导子代个体的选择.

进化规划通过变异、竞争、选择等算子的作用来模拟生物的进化过程. 本文针对所研究的并行机调度问题,提出一种改进的进化规划算法 (SEP),其具体执行过程如下:

Step1 初始种群的产生. SEP 的初始种群个体是由目标各分量组成的数字串. 为提高 SEP 的效率以及改进解的质量,结合并行机调度问题本身的一些表述知识,对于初始种群个体,采用两部分编码 X / Y 更易于理解. $X = (x_1, x_2, \dots, x_n)$ 为 n 个工件的排序,是字符编码,字符串长度为 n ; $Y = (y_1, y_2, \dots, y_m)$ 为每台机器上的加工工件数,是字符编码,满足约束条件 $\sum_{j=1}^m y_j = n$,字符串长度为 m . 确定编码方式后,随机产生初始种群. 对于 X 部分编码,随机产生 n 的排序即可. 对于 Y 部分编码,考虑到每台机器上至少要加工一个工件的约束,每个 y_j 是在一定区间范围内随机产生的,区间为 $[1, n - (m - j) - \sum_{k=1}^{j-1} y_{k-1}]$,最后一台机器上加工的工件数为 $n - \sum_{k=1}^j y_{k-1}$.

Step2 适应度计算. 作为评判标准,可根据适应度选择子代个体. 在一般确定环境下, makespan 的取值可直接计算,而本文的工件加工时间是不确定的. 鉴于此,基于式 (3) 采用粗糙模拟的方法计算不确定条件下的目标函数 C_{max} 的期望值

$$E[C_{max}] =$$

$$\frac{\sum_{i=1}^N (C_{max}(x, y, r_i) + C_{max}(x, y, \bar{r}))}{2N} \quad (7)$$

Step3 变异操作. 标准 EP 算法常采用高斯变异,即对当前群体中个体的每个分量添加一个符合正态分布的随机数. 针对并行机调度问题的两部分编码形式,SEP 的变异方式有所改进,不同的编码部分采用不同的变异方法. 对于 X 部分编码,借鉴 Fogel 在文献^[11] 中的变异方法,随机选取 X 部分编码中的两点,并反转两点间的编码来产生新的个体. 对于 Y 部分编码,根据所考虑问题的编码形式的特殊性,在机器编码部分,本文提出一种新的变异方法. 机器编码部分必须满足以下 2 点约束: 1) y_j 为正整数; 2) 加工工件数必须为 n . 这便促使了新的变异方法的产生. 随机选取 Y 部分编码的 2 点,将其中一点进行高斯变异,另一点则为满足约束而补差其变异前后的变化量. 这一点的高斯变异原则如下:

$$y_j = y_j + [\cdot N(0, 1)], \quad (8)$$

$$= \times \frac{FIT_{old}}{FIT_{opt}} \quad (9)$$

式中: y_j 和 y_j 分别代表子代和父代在 Y 部分中的变异点, $N(0, 1)$ 表示均值为 0 且标准差为 1 的高斯随机变量, $[\cdot]$ 表示变动范围; FIT_{old} 为父代的适应度, FIT_{opt} 为种群中的最优适应度, $[\cdot]$ 表示取整. 变异后的子代加上原来的父代共同形成混合群体,并计算其适应值.

Step4 竞争. SEP 采用随机 q 竞争选择方法使优良个体尽可能入选,但也允许极少数性能较差的个体入选,从而使少数性能较差的个体也可以保留在下一代群体中,这种选择方式也是动态、不确定的. 在混合群体中,通过 q 竞争选择法产生新一代群体.

Step5 若满足终止条件,则停止计算;否则,重复 Step2 ~ Step4,直至满足停止条件,获得满意解为止. 终止准则为选择最大迭代次数和最优个体与平均值的偏差作为判据.

5 仿真实验

为了验证 SEP 算法的有效性,考虑不同规模下随机生成的并行机调度算例进行仿真研究. 计算机配置为 Intel Pentium 4 CPU 3.06 GHz,内存为 512 M,编程环境为 VC++ 6.0. 本文研究的是工件加工时间的不确定性,用粗糙变量 $r = (x, \bar{r})$ 表示,在现实生产中可表示为某工件的粗糙加工时间为 $([8, 10], [7, 12])$. 实际上,工件处理时间的最小范围是区间 $[8, 10]$,最大范围是 $[7, 12]$. 对于粗糙数据的产生,本文采取如下方法: 1) 在区间 $[a, b]$ 内随机

生成粗糙加工时间最大范围的下区间数据 \underline{p}_1 , 其上区间数据 \bar{p}_1 在区间 $(\underline{p}_1, \bar{p}_1]$ 中随机产生, $\bar{p}_1 > 1; 2)$ 最小范围的加工时间 $[\underline{p}_2, \bar{p}_2]$ 在区间 $(\underline{p}_1, \bar{p}_1)$ 内随机产生. 为了计算方便, 设置参数 $a = 1, b = 20, \alpha = 1.5$.

在应用第4节提出的SEP求解并行机调度问题中提到了适应度计算的粗糙模拟方法, 其中对于 N 取值的确定, 本文采用了多次仿真模拟的方法. 考虑3机器10工件的一个工件和机器排序, 粗糙模拟计

算它的 makespan. 图1展示了 N 取不同值时的粗糙模拟过程及其相对应的 makespan 值. 可以看出, N 取较小值时, 粗糙模拟的值跳跃非常大; 随着 N 的增大, makespan 逐步趋于稳定; N 充分大时, 粗糙模拟值即为 makespan 的期望估计值. 综合权衡计算时间和计算结果, 对于本文的调度问题, 取 $N = 3500$.

改进进化规划算法 SEP 的参数设置为 $n = 20, q = 0.9n, \alpha = 0.1$. 考虑3种规模的并行机调度问题, 机器数为3, 加工工件数分别为10, 20和30. 每

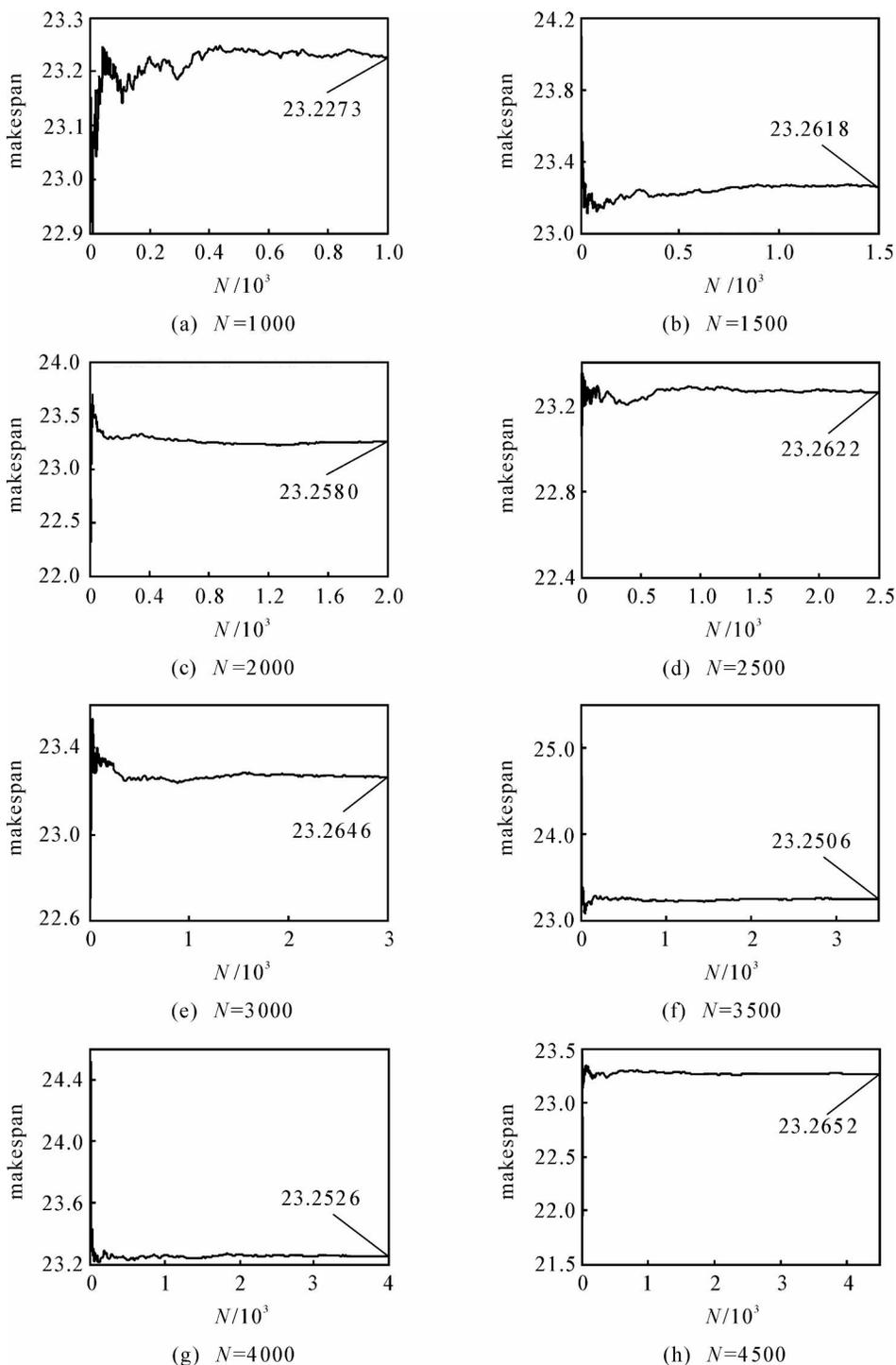


图1 粗糙模拟期望值演化曲线

表 1 不同规模算例下 SEP 和 GA 算法 10 次仿真的结果对比

问题规模	SEP			GA		
	最优值	平均值	CPU time/s	最优值	平均值	CPU time/s
3 × 10	17.0945	17.1473	26.156	17.1381	17.5576	30.225
3 × 20	65.0044	65.1284	101.797	66.1450	66.7171	115.206
3 × 30	115.2052	115.2859	130.203	116.4540	117.1329	165.406

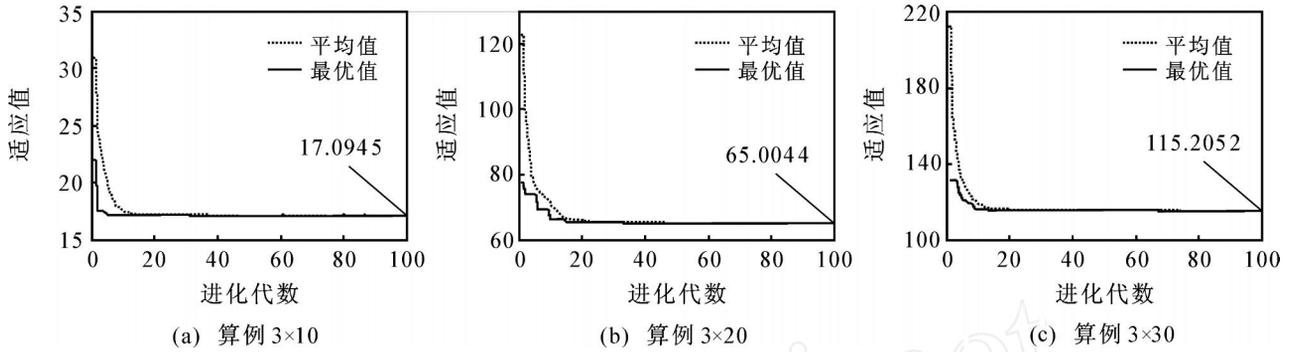


图 2 3 种规模算例的最优值演化曲线

种规模的算例仿真 10 次. 取最优值、平均值和计算时间(CPU time) 列于表 1, 并与遗传算法比较. 从表 1 可看出, 在同一算例的情况下, SEP 找到的解优于遗传算法. 不同规模下的最好结果的最优值演化曲线如图 2 所示. 其中: 横轴为进化代数; 纵轴为适应度, 即 makespan 的期望估计值; 点线为种群适应度平均值曲线, 对应着演化代数和当前代种群中所有个体适应值的平均值; 实线为种群最优值曲线, 对应着每代种群中的最优适应值. 图 2 验证了改进进化规划算法 SEP 的收敛性.

6 结 论

目前对于不确定条件下的并行机调度问题的研究主要集中在工件处理时间上, 研究方法和角度有多种. 本文研究的粗糙变量的表述, 从实际生产的角度考虑了工件处理时间的不确定性, 由此建立了该调度问题的粗糙期望值规划模型, 提出了适用于生产调度问题的改进进化规划算法. 不同规模的仿真实验验证了该算法的有效性和普适性. 今后可从模型入手, 加入新的约束或调整目标函数, 从另一角度更深入地探究粗糙集描述不确定信息的其他方法.

参考文献(References)

[1] 顾幸生. 不确定性条件下的生产调度[J]. 华东理工大学学报, 2000, 26(5): 441-446.
(Gu X S. Production scheduling under uncertainty[J]. J of East China University of Science and Technology, 2000, 26(5): 441-446.)

[2] Zadeh L A. Fuzzy sets[J]. Information and Control, 1965, (8): 338-356.

[3] Pawlak Z. Rough sets[J]. Int J of Information and Computer Sciences, 1982, 11(5): 341-356.

[4] Tavakkoli-Moghaddam R, Jolai F, Vaziri F, et al. A

hybrid method for solving stochastic job shop scheduling problems[J]. Applied Mathematics and Computation, 2005, 170: 185-206.

[5] 刘琦, 顾幸生. 基于模糊规划的处理时间不确定条件下的 Job shop 问题[J]. 华东理工大学学报, 2001, 27(5): 442-445, 450.
(Liu Q, Gu X S. Job shop problem under uncertain processing time based fuzzy programming[J]. J of East China University of Science and Technology, 2001, 27(5): 442-445, 450.)

[6] Pawlak Z. Rough sets and their applications [J]. Microcomputer Applications, 1994, 13(2): 71-75.

[7] 于艾清, 顾幸生. 基于广义粗糙集的不确定条件下的 Flow Shop 调度[J]. 系统仿真学报, 2006, 18(12): 369-372.
(Yu A Q, Gu X S. Flow shop scheduling problem under uncertainty based on generalized rough sets[J]. J of System Simulation, 2006, 18(12): 369-372.)

[8] Liu B D. Uncertainty theory: An introduction to its axiomatic foundations [M]. Berlin: Springer-Verlag, 2004.

[9] Liu B. Theory and practice of uncertain programming [M]. Heidelberg: Physica-Verlag, 2002.

[10] Fogel L J, Owens A J, Walsh M J. Artificial intelligence through simulated evolution [M]. New York: Wiley, 1966.

[11] Fogel D B. Applying evolutionary programming to selected traveling salesman problems[J]. Cybernetics and Systems, 1993, 24(1): 27-36.

[12] Min Liu, Cheng Wu. Scheduling algorithm based on evolutionary computing in identical parallel machine production line[J]. Robotics and Computer Integrated Manufacturing, 2003, (19): 401-407.