

遗传算法在甲醛生产过程优化中的应用

王 强 邵惠鹤

(上海交通大学自动化研究所·上海, 200030)

摘要: 遗传算法是一种模拟自然进化而提出的简单高效的组合优化算法. 本文研究了甲醛生产过程的优化问题, 该过程由于其反应动力学的固有复杂性而无法建模, 本文表明遗传算法可以有效地解决这一过程的寻优问题.

关键词: 遗传算法; 优化; 甲醛

1 前 言

遗传算法作为一种简单有效的组合优化算法正受到越来越广泛的重视, 它是一种基于自然选择和自然遗传的全局优化算法, 其基本原理最早由 Holland 及其学生提出, 采用从自然选择机理中抽象出来的几种算子对参数编码字符串进行操作, 由于这种操作是针对多个可行解构成的群体进行, 具有本质的并行计算特点, 故在其世代更替中可以并行地对参数空间的不同区域进行搜索, 并使得搜索朝着更有可能找到全局最优的方向进行, 且不至于陷入局部极小. 由于遗传算法具有这样的特点, 因而被广泛用于机器学习^[2]、函数优化^[3,4]、图像处理^[5]、系统辨识等多个领域, 并已召开了多次国际会议. Goldberg 的专著《Genetic Algorithms in Search, Optimization, and Machine Learning》对此作了较为全面的总结.

2 算法结构

对于一般的优化问题, 均可表示为:

$$\text{目标函数} \quad \min F = f(x, y, z), \quad F \in \mathbb{R},$$

$$\text{约束} \quad G(x, y, z) \geq 0, \quad (x, y, z) \in \Omega.$$

遗传算法的基本思想是将优化问题的一组基本可行解 (x_i, y_i, z_i) 编码为一组二进制的字符串, 每个字符串包含多个子字符串, 每个子字符串是目标函数空间的某一个参数的二进制编码. 子字符串的某一位或几位的组合称为一个基因 (gene). 这样的每个字符串代表一个寻优目标空间内的可行解, 根据优化问题的目标函数, 每个可行解对应了一个目标函数值, 称为适合度 (fitness). 多个基本可行解的集合被称为一个群体 (population), 该群体被用在每次寻优迭代通过几种简单算子的操作生成新的可行解群体, 称为一个世代 (generation). 最基本的操作有以下几种:

2.1 复制 (Reproduction)

复制是将亲代的个体信息传递到子代, 而这种传递是有所选择的, 每代中的每一个个体, 按照其适合度函数的大小决定它能够复制到下一代的概率, 这种概率是根据与平均适合度之比确定, 适合度较小的个体由于复制概率低而逐渐被淘汰, 而适合度大的个体由于复制概率大, 不仅可以复制, 而且有权力拥有不止一个后代. 通过复制, 使得群体中的优秀个体数

目不断增加,整个进化过程朝着更优解的方向进行,反映了优胜劣汰的原则。

2.2 交叉(Crossover)

复制虽然可以使可行解群体朝着最优解方向移动,但只能在现有群体内寻优,它不能产生与亲代不同的个体。在自然生态系统中的个体很少是单亲繁殖,一般均为双亲繁殖,子代从两个亲代获得遗传信息,也就可能从双亲处获得各自的有利的遗传特性,产生适应能力更强的个体。遗传算法中模仿这一特点引入了交叉算子,每一代的各个个体之间按一定的概率交换其部分基因,产生新的基因组合,使各个解有机会交流其优秀基因,可望获得比亲代更好的解的结构。图1所示为两个父代个体在第三位后进行交叉产生两个新个体的过程。图示是最简单的也是应用最广的单点交叉方式。一些较为复杂的交叉方式曾被采用过^[6]。

$$\begin{array}{ll} X = a_1 a_2 a_3 a_4 a_5 a_6 & [000000] \\ Y = b_1 b_2 b_3 b_4 b_5 b_6 & [111111] \\ & \downarrow \\ X = a_1 a_2 a_3 b_4 b_5 b_6 & [000111] \\ Y = b_1 b_2 b_3 a_4 a_5 a_6 & [111000] \end{array}$$

图1 交叉因算子示意图

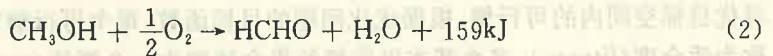
2.3 变异(Mutation)

复制和交叉只能在现有的基因型的排列组合内寻找最优,而不能产生新的基因型。如果所有字符串的某一位都是0,则无论经过多少次的复制与交叉该位置也都将是0。变异算子是对每个字符串的每一位按一定的概率由0变1或由1变0,产生新的基因型,扩大寻优范围。

遗传算法由于是采用群体进化的方式,有别于以往的单线索对目标函数空间搜索的优化算法,是对目标函数空间进行多线索的并行搜索,同时对多个可行解进行检查,并通过交叉算子在各可行解之间交换信息,不会陷入局部极小,可以有效地处理优化问题。其次它在使用中,需要的信息较少,只需要各个可行解的目标函数值,使用方便,且不需对目标函数作可微或连续等的假设。此外它的解的选择采用和模拟退火方法相似的概率选择,适应能力强。

3 遗传优化在甲醛生产过程中的应用

甲醛生产是以甲醇、空气、水蒸汽为原料,采用电解银作催化剂在绝热反应器内进行氧化脱氢反应生成甲醛,主要反应有



目前国内的厂家在生产中通过废热锅炉等节能技术,能耗等指标已大幅度降低,但是原料单耗与国际先进水平相比还有较大的差距,采用计算机在线优化控制等方法,可以在保证产品质量的基础上降低单耗,提高收率,延长催化剂使用寿命,具有很大的经济效益。

对于甲醛的生产装置来说,直接的优化目标函数是总的生产成本,由于整个过程中原料甲醇的成本占总成本的93%以上,因此也可以采用原料单耗作为优化的目标函数,两者是基本等价的。

甲醇脱氢生成甲醛的反应是在绝热固定床式反应器里进行的。针对该反应过程的动力学研究早在六十年代初就有研究,由于各研究者在实验方法、实验条件、催化剂制备及数据处理方法上的不同,所得到的动力学方程的表达也有所不同,有的过于繁琐,难以实时应用,

有的因反应温度不同而不适用于工厂生产装置的反应过程.事实上,甲醇氧化、脱氢制甲醛的气固相催化反应过程中,反应速率不仅与反应动力学有关,而且还与热质传递过程有关.对于这样一个复杂过程,采用建立数学模型然后进行优化是相当困难的,而采用遗传算法结合神经网络方法可以有效地解决这个问题.

甲醇脱氢反应器的反应结果由入口条件唯一确定,在催化剂和装填状况一定时,生产状况主要由以下四个指标反映:反应温度 T 、氧醇分子比 E 、蒸汽量 G_s 和空速 V .在实际生产中空速一般是固定不变的,此时 T, E, G_s 中只有两个是独立变量,只要确定其中任意两个,另一个就唯一确定.装置的控制采用反应器温度与甲醇蒸发器温度过程串级控制,为控制的方便起见,可以先确定 T, E , 然后决定蒸汽流量 G_s , 因此可以将三者的关系表示:

$$G_s = g(E, T). \quad (3)$$

在反应过程中,一般采用水醇比 H 而不是直接采用蒸汽流量 G_s 来描述,因此上式可写作:

$$H = h(E, T). \quad (4)$$

由于反应器中 (H, E, T) 三者关系无法得到解析表达,只能根据实际生产时的数据记录绘出的三者大致关系曲面如图 2. 为从得到的最优 E 和 T 得到 H , 我们采用神经网络方法得到圆满解决,具体方法这里不再详述.

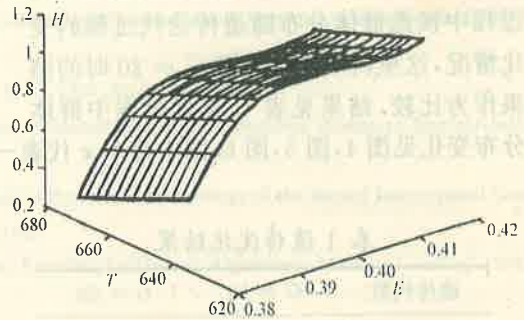


图 2 甲醛反应器中水醇比与氧醇比、反应温度三元关系图

以生产中单耗 S 最低为目标函数,操纵变量为 E 和 T ,所要解决的优化问题可描述如下:

$$\min S = f(E, T). \quad (5)$$

约束条件包括:

- 1) 反应温度范围,由催化剂的适用温度确定

$$635^{\circ}\text{C} \leq T \leq 672^{\circ}\text{C}. \quad (6)$$

- 2) 产品中酸度规定,可化为相应的 T 和 E 的约束

$$(15473.15 - 45.05908T + 3.36324 \times 10^{-2}T^2) \leq E \times 10^3 \leq (313.4556 + 0.1671598T). \quad (7)$$

对于这样一个非线性约束的优化问题,其目标函数值可以通过实测得到,由于无法得到(5)式的解析表达式,无法采用一般的数值优化算法.这里采用遗传算法进行寻优.首先通过采集该反应器的运行数据,得到反应单耗 S 与氧醇比 E 和反应温度 T 的大致函数图象如图 3.

采用遗传算法对该问题进行寻优的算法如下:

- 1) 随机地在可行区域内产生 50 个候选解,采用二进制方法对自变量 E 和 T 编码;
- 2) 按照预定的交叉率在候选解群体中随机抽取数对候选解进行交叉操作;
- 3) 按照预定的变异率抽取候选解进行变异操作;
- 4) 计算每个候选解的目标函数值,根据所选的淘汰率淘汰候选解中目标函数最差者.

所空缺位置从目标函数较好的候选解中产生;

5) 判断如得到最优或到了预定迭代次数则结束, 否则转 2).

针对本文问题, 为保证有足够精度, 同时由于在 PC 机上计算, 为方便编程, 参数 E 和 T 的编码二进制串长取 16 位, 所使用的遗传参数为: 群体大小 $P = 50$; 淘汰率 $D = 0.02$; 变异率 $m = 0.01$; 交叉率 $C = 0.06$. 整个寻优过程经过 10 次迭代即已得到稳定最优解, 可结束迭代. 为更好说明寻优过程中候选群体分布随遗传迭代过程的变化情况, 这里再取遗传代数 $G = 20$ 时的结果作为比较, 结果见表 1, 寻优过程中群体分布变化见图 4, 图 5, 图 6. 图中每个 x 代表一个候选.

表 1 遗传优化结果

| 遗传代数 | $G = 10$ | $G = 20$ |
|------------|----------|----------|
| 最优反应温度 T | 638.54 | 638.65 |
| 最优氧醇比 E | 0.4038 | 0.4035 |
| 最小单耗 | 437.59 | 437.59 |

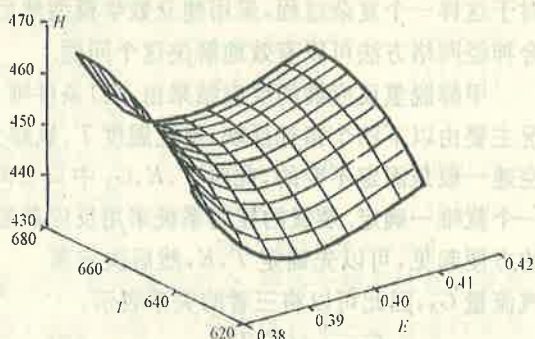


图 3 生产单耗与氧醇比及反应器三元关系图

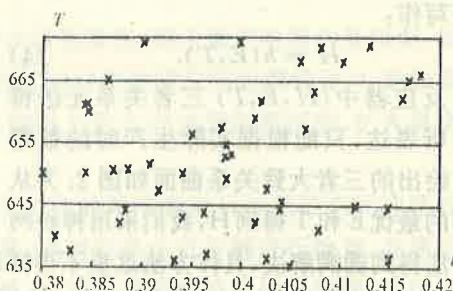


图 4 初始候选解分布

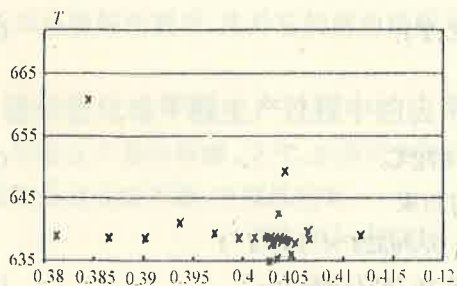


图 5 第 10 代遗传结果

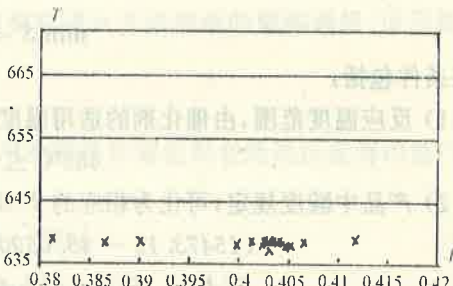


图 6 第 20 代遗传结果分布

经过遗传优化后得到最优操作点为 $(0.404, 638.7)$, 在此条件下单耗 $S = 437.6$.

4 结果分析

从以上的遗传寻优结果分布可见, 初始候选解分布基本随机的, 随着遗传迭代的进行, 大部分的可行解逐渐向较优区域集中, 表示遗传过程正在逐渐淘汰不良解, 同时使得较优区域内的搜索密度增强, 找到最优解的概率也加大. 同时在可行区域的其它部分仍有候选解存在, 可以有效地防止陷入局部极小. 经过 10 代遗传后大部分候选解均已集中 $(0.404, 639)$ 附近, 此时可以认为已经找到最优点. 从图 6 也可以看出, 再继续迭代 10 次后目标函数下降很少, 只是候选解分布更加密集, 可见遗传算法的寻优效率是比较高的.

前面通过遗传优化得到的最优操作条件下单耗 $S = 437.6$, 而根据厂方两年的生产记录显示年平均单耗 $S = 461$, 采用遗传算法得到的最优工艺参数指导生产具有明显的经济效益. 遗传算法进行优化计算过程中只用到目标函数值而不需要数学模型或目标函数导数等附加条件, 使用方便.

参 考 文 献

- [1] Kosmogorov. On the Representation of Continuous Functions of Many Variables by Superposition of Continuous Functions of One Variable and Condition. Dokl. Akad. Nauk, USSR, 114, 1957, 953—956
- [2] Marco Dorigo, Uwe Schnepf. Genetics-Based Machine Learning and Behavior-Based Robotics: A New Synthesis IEEE Trans. SMC, 1993, 23(1): 141—154
- [3] Larry Rendell. Induction as Optimization: IEEE Trans. SMC, 1990, 20(2): 326—338
- [4] Hollstien, R. B. Artificial Genetic Adaption in Computer Control Systems. Dissertation Abstracts International, 1971, 32(3): 1510B
- [5] Stadnyk, L. . Schema Recombination in Pattern Recognition Problems. Proceedings of the Second International Conference on Genetic Algorithms and Their Applications, 1987
- [6] Schraudolph, N. N. and Belew, R. K. . Dynamic Parameter Encoding for Genetic Algorithms. Machine Learning, 1992, 9(1): 9—21

Genetic Optimization in Formaldehyde Production Process

WANG Qiang and SHAO Huihe

(Research Institute of Automatic Control, Shanghai Jiao Tong University · Shanghai, 200030, PRC)

Abstract: Genetic algorithm is a simple but efficient combinatorial optimization algorithm based on the principle of nature evolution. This paper studied the optimization of formaldehyde production process. Due to the complexity experienced in the process reaction dynamics, it is difficult to set up precise model. It is showed in this paper that genetic algorithm is capable of handling such sort of optimization problems.

Key word: genetic algorithm; optimization; formaldehyde

本文作者简介

王 强 1968年生. 1993年毕业于华东化工学院, 现在上海交通大学自动化系攻读博士学位. 主要从事化工过程控制及稳态优化, 神经网络, 遗传算法等方面的研究.

邵惠鹤 1936年生. 1961年毕业于华东化工学院, 现为上海交通大学教授, 博士生导师. 目前研究兴趣是过程模型化及优化控制, 生化反应器控制, 智能控制理论研究及应用.