

文章编号: 1005—7854(2005)01—0087—04

遗传算法的原理与应用

李华昌¹, 谢淑兰¹, 易忠胜²

(1. 北京矿冶研究总院, 北京 100044; 2. 桂林工学院材料与化学工程系, 广西桂林 541004)

摘 要: 遗传算法是基于自然界生物进化基本法则而发展起来的一类新算法。本文在简要介绍遗传算法的起源与发展、算法原理的基础上, 对算法在优化、拟合与校正、结构分析与图谱解析、变量选择、与其他算法的联用等方面的应用进行了综述。该算法由于无需体系的先验知识, 是一种全局最优化方法, 能有效地处理复杂的非线性问题, 因此有着广阔的应用前景。

关键词: 遗传算法; 化学计量学; 优化

中图分类号: O242.1 **文献标识码:** A

THEORY AND APPLICATION OF GENETIC ALGORITHM

LI Hua-chang¹, XIE Shu-lan¹, YI Zhong-sheng²

(1. Beijing General Research Institute of Mining and Metallurgy, Beijing 100044, China;

2. Department of Material Engineering, Guilin Institute of Technology, Guilin 541004, Guangxi, China)

ABSTRACT: Genetic Algorithm (GA) is a kind of recursive computational procedure based on the simulation of principle principles of evaluation of living organisms in nature. Based on brief introduction of the principle, the beginning and development of the algorithms, the paper reviewed its applications in the fields of optimization, fitting and calibration, structure analysis and spectra interpretation, variable selection, and its usage in combination with others. The application of GA needs no initiating knowledge of the system, and therefore, is a comprehensive optimization method with extensive application in terms of processing complex non-linear problems.

KEY WORDS: Genetic Algorithm (GA); Chemometrics; Optimization

1 遗传算法的起源与发展

遗传算法 (Genetic Algorithm, GA) 是模拟自然界生物进化机制的一种算法, 即遵循适者生存、优胜劣汰的法则, 也就是寻优过程中有用的保留, 无用的则去除。在科学和生产实践中表现为, 在所有可能的解决方法中找出最符合该问题所要求的条件的解决方法, 即找出一个最优解。这种算法是 1960 年由 Holland 提出来的^[1-4], 其最初的目的是研究自然系统的自适应行为, 并设计具有自适应功能的软件系统。它的特点是对参数进行编码运算, 不需要有关

体系的任何先验知识, 沿多种路线进行平行搜索, 不会落入局部较优的陷阱, 能在许多局部较优中找到全局最优点, 是一种全局最优化方法。近年来, 遗传算法已经在国际上许多领域得到了应用。1985 年召开了第 1 届有关遗传算法的国际会议, 第 1 部关于这方面的专著在 1989 年问世。遗传算法是一种有广泛应用前景的算法, 但是它的研究和应用在国内尚处于起步阶段^[3]。

2 算法原理

在自然界, 由于组成生物群体中各个体之间的差异, 对所处环境有不同的适应和生存能力, 遵照自然界生物进化的基本原则, 适者生存、优胜劣汰, 将要淘汰那些最差个体, 通过交配将父本优秀的染色

收稿日期: 2004—08—20

作者简介: 李华昌, 测试研究所所长, 研究员、博士。

体和基因遗传给子代,通过染色体核基因的重新组合产生生命力更强的新的个体与由它们组成的新群体。在特定的条件下,基因会发生突变,产生新基因和生命力更强的新个体;但突变是非遗传的,随着个体不断更新,群体不断朝着最优方向进化,遗传算法是真实模拟自然界生物进化机制进行寻优的。在此算法中,被研究的体系的响应曲面看作为一个群体,相应曲面上的每一个点作为群体中的一个个体,个体用多维向量或矩阵来描述,组成矩阵和向量的参数相应于生物种组成染色体的基因,染色体用固定长度的二进制串表述,通过交换、突变等遗传操作,在参数的一定范围内进行随机搜索,不断改善数据结构,构造出不同的向量,相当于得到了被研究的不同解。目标函数值较优的点被保留,目标函数值较差的点被淘汰。由于遗传操作可以越过位垒,能跳出局部较优点,到达全局最优值。

遗传算法是一种迭代算法,它在每一次迭代时都拥有一组解,这组解最初是随机生成的,在每次迭代时又有一组新的解由模拟进化和继承的遗传操作生成,每个解都有一目标函数给与评判,一次迭代成为一代。典型的算法步骤有:

- (1)初始化,即随机生成一个符号串群体;
- (2)基于适度函数对符号串进行评价;
- (3)应用一组遗传操作生成一个新的符号串群体;
- (4)重复步骤(2)和(3)直至结果收敛。

基本的遗传算法框图如图1所示,其中变量GEN是当前代数。它按不依赖于问题本身的方式作用在特征串群体上,算法仅用到与搜索空间中检查过的点相联系的适应值。不管求解问题本身,通过执行同样的、惊人简单的复制、杂交和偶尔的变异操作来完成它的搜索。在实际应用中,遗传算法能够快速有效地搜索复杂、高度非线性和多维空间,出人意外的是它并不知道问题本身的任何信息,也不了解适应值度量。遗传算法成功的关键在于符号串表示和遗传操作的设计。

3 遗传算法的应用

遗传算法可处理连续变量参数的优化问题,特别是适用于复杂非线性问题的处理。可用于NMR脉冲形状分析、RNA核苷酸测定、DNA构象分析、分子识别和设计、变量选择等,在分析化学、环境科学、机械设计中的应用也非常广泛。

3.1 优化

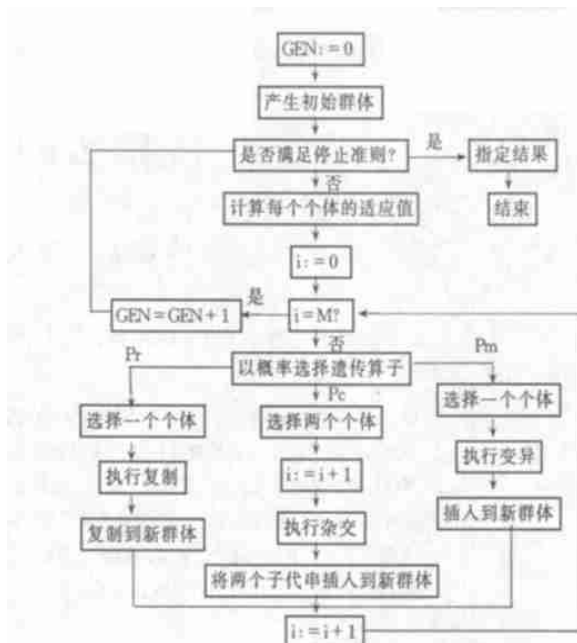


图1 遗传算法框图

Fig. 1 Sketch map for Genetic Algorithm

遗传算法用于优化的研究成果目前已经发表了较多的论文。这些成果有的是对此算法进行改进,如李通化等^[5]给出了一种动态的遗传算法,详细讨论了变异函数对优化进程的影响,并用红外光谱数据和二极管阵列检测液相色谱数据进行了验证。杨立民等^[6]运用双极性压缩函数适应度定标和基于排挤方法的选择算子改进标准遗传算法(SGA),使其成为简单通用、快速收敛的并行全局搜索算法;利用该算法优化误差反向传播网络(BPN),克服了BPN收敛慢和不具有全局收敛性的缺陷,并在此基础上,建立大气环境质量评价模型。张昆等^[7]为了使金属切削加工中切削参数能实现实时优化以保证产品质量和设备效率,提出采用遗传算法,结合现场实际工况的反馈信息实现了实时优化,在任一不同的生产条件下均能达到最优值。也有应用遗传算法去解决某一实际问题,如宋仁国等^[8]将人工神经网络用于建立7175铝合金的性能预测模型,在此基础上,采用遗传算法对其工艺进行优化;王举等^[9]将间歇化工过程的最优设计问题,分解为只包含离散变量的主导问题和只含连续变量的子问题,把遗传算法和线性规划法结合起来对其进行求解,并在算法中引入了一类新的算子,显著地提高了收敛概率。

3.2 拟合与校正

分析仪器是通过分析信号的检测从而得到物质的浓度,因此对仪器信号的分辨和校正就显得尤为重要。蔡煜东等^[10,11]将遗传算法用于化学校正,运用改进的遗传算法拟合离子选择电极工作曲线,并

以半对数线性函数模型和二阶幂函数模型尝试了该算法的效果; 结果表明, 改进的遗传算法性能较好, 优于一般遗传算法和直线回归法, 可望成为各类传感器非线性曲线拟合的有效手段。蔡煜东还提出非线性多元函数拟合的遗传算法, 为分析化学中非线性函数拟合、曲线校正提供了一种性能较好的方法。孙雪峰等^[12]研究了遗传算法求解复杂穆斯堡尔谱的方法, 得到了满意的结果。姚萌等^[13]从一系列雷达信号图像获得必要的物体的相关信息, 以遗传算法对其进行处理, 将探测目标定位转换为曲线拟合, 在地下遗址探测中取得了较好结果。曲铁军等^[14]设计了使用遗传算法的自适应滤波器, 给出了离线和在线两种实现方案。离线方案以识别滤波参数为主要目的, 进而可以对状态进行较准确的事后估计; 在线方案以实时地对状态进行估计为目的, 对滤波器参数寻优采用具有良好性能的浮点编码遗传算法。

3.3 结构分析与图谱解析

结构分析在有机化学和药物化学领域中的应用是非常广泛的, 遗传算法主要应用在需要分子三维构像信息的时候, 这在合成新的有机化合物和药物、充分地利用天然资源具有重要的意义, 邓勃^[3]对这方面的应用作了简要的评述。侯廷军等^[15]对遗传算法在计算机辅助药物分子设计中的应用做了系统阐述。蔡文生等^[16]提出一种采用整数编码和基于节点基因交换方式的遗传算法, 并应用于化学结构图的同态研究。遗传算法在一组随机生成的表示目标结构与查询结构节点间映射关系的整数串中进行逐步优化, 直到找出与查询结构匹配的映射, 从而实现化学结构图的同态匹配。卢佩章等^[17]将遗传算法引入色谱峰解析程序, 结合色谱峰模型及峰形经验公式实现了恒温色谱图的全自动解析。陈闽军等^[18]提出了一种基于遗传算法的色谱指纹峰配对识别方法, 该方法根据色谱指纹图谱峰分布特性初选出若干标定峰, 将其存入一个候选标定峰库, 同时根据这些候选标定峰从待测指纹图谱中选出相应的候选标定峰, 以遗传算法对指纹峰进行识别。

3.4 变量选择

在各领域中, 都存在目标函数与影响因素的某种关系, 而这些影响因素是千变万化的, 它们的数量巨大, 如何从大量的影响因素中寻找最有利于建模的因素就成了研究的重要课题, 遗传算法在这方面显示了它的优势。Leardi R^[19], Barros A S 等^[20]将遗传算法用于主成分的选择。郑延斌等^[21]将遗

传算法应用于过饱和溶液 $\text{Li}_2\text{O} \cdot 3\text{B}_2\text{O}_3 - \text{H}_2\text{O}$ 体系结晶过程动力学参数辨识, 确定了结晶反应速率常数、热力学平衡浓度和表观反应级数。另外, 遗传算法也可以用于光度分析中波长的选择, 进而用回归算法求解。夏之宁等^[22]将遗传算法用于紫外光谱数据处理, 实现了多元分辨与校正, 并对多组分分析体系进行同时定量测量。杨杰等^[23]介绍了激光切割、拼焊的一种基于遗传算法的神经网络工艺参数选择方法的原理及实现方法, 通过遗传和变异的方法来优化和选择权值进行评价, 从而达到了智能化选择参数的目的。章元等^[24]在传统的遗传算法中引入 Metropolis 接受准则, 结合有序 Gram - Schmidt 正交化, 得到了预测能力较强的模型, 该法用于处理钢中微量元素及热处理工艺条件与钢的力学性能关系问题, 是一种较实用的变量筛选方法。

3.5 与神经网络的联用

近年来神经网络研究成为国际前沿的科学研究领域之一, 将神经网络与遗传算法联用, 具有更强的建模、寻优等能力。孙卫国等^[25]用遗传算法训练神经网络, 为苯乙酸胺类化合物的 QSAR 建模, 效果良好。王雅琳等^[26]针对电力部门提出的分时计价原则, 以锌电解过程为研究对象, 用生产过程数据建立了不同酸度下电流密度与电流效率间关系, 由此建立了分时供电优化模型, 并用一种变形的多变量递推最小二乘自适应策略不断修正优化模型, 采用高效模拟退火优化算法, 引入惩罚算子, 求解带非线性约束和边界约束的优化问题, 获得了合理的分时供电方案。路文江等^[27]利用人工神经网络的自学习以及非线性逼近能力对材料元素与硬度的相关性进行拟合和预测, 并用遗传算法的强寻优能力对喷涂材料成分进行优化。张玉祥^[28]将遗传算法、小波分析、人工神经网络和模拟退火思想结合起来, 提出了用遗传算法来学习小波神经网络层间的权值、尺度参数和位置参数, 将其应用于矿压预报, 得到了比传统神经网络更优的效果。何耀华等^[29]针对化工过程系统中难于建立机理模型的过程或设备, 提出了集成运用神经网络和遗传算法的优化方法 (NN - GA 方法), 并介绍了对神经网络训练算法的有效改进以及 NN - GA 方法的实施过程, 给出了应用实例, 分析了 NN - GA 方法与统计回归分析相比所具有的独特优势。

4 结 语

遗传算法是一种全局最优化方法, 在优化过程

中,它无需体系的先验知识,能在许多局部较优中找到全局最优点,能有效地处理复杂的非线性问题。目前,尽管对于遗传算法及其在分析化学、环境科学、机械设计中的应用研究取得了部分可喜的成果,但对于算法本身来说,如何进一步提高收敛速率、遗传算子的研究、与其他算法的联用等方面都还可以做较多的工作;在应用研究方面,除上述领域外,还可以在其他诸如化工、冶金、材料、仪器仪表等诸多需要进行优化设计的领域应用,有广阔的应用前景。

参考文献:

- [1] Holland J H. Adaptation in Natural and artificial Systems [M]. Ann Arbor: The University of Michigan Press, 1975.
- [2] Cong Peisheng, Li Tonghua. Numeric genetic algorithm [J]. Anal., Chem. Acta., 1994, 293: 191—203.
- [3] 邓勃, 刘嘉. 遗传算法在分析化学中的应用[J]. 分析科学学报, 1997, 13(2): 160—168.
- [4] 李军, 边肇祺. 用于最优化的计算智能[M]. 北京: 清华大学出版社, 1999.
- [5] 李通化, 张奇志, 朱仲良, 等. 用数值遗传算法计算配合物的稳定常数[M]. 高等学校化学学报, 1995, 16(3): 354—358.
- [6] 杨立民, 许有鹏. 改进遗传神经网络方法在大气环境质量评价中的应用[J]. 环境科学研究, 1999, 12(2): 28—31.
- [7] 张昆, 韩至骏, 于国军. 用基因算法实现切削参数的现场实时优化[J]. 清华大学学报(自然科学版), 1999, 39(2): 27—29.
- [8] 宋仁国, 董敏, 张奇志, 等. 7175 铝合金工艺优化的遗传算法[J]. 材料科学与工程, 1998, 16(1): 28—31.
- [9] 王举, 陈中州, 袁希钢. 用于间歇化工过程最优设计的遗传算法[J]. 计算机与应用化学, 1999, 16(1): 24—28.
- [10] 蔡煜东. 分析化学中非线性多元函数的遗传算法[J]. 分析化学, 1995, 23(7): 790—792.
- [11] 蔡煜东. 运用改进的遗传算法拟合离子选择电极工作曲线[J]. 分析化学, 1995, 23(6): 640—643.
- [12] 孙雪峰, 林俊, 黄红波, 等. 用遗传算法拟合穆斯堡尔谱[J]. 计算物理, 2003, 20(3): 264—266.
- [13] 姚萌, 孟海燕. 基于遗传算法的地下遗址 GPR 数据分析[J]. 华东师范大学学报(自然科学版)[J]. 2003(1):

48—54.

- [14] 曲铁军, 杨旭东, 葛升民. 使用遗传算法的自适应滤波器[J]. 哈尔滨工业大学学报, 2003, 35(6): 655—659.
- [15] 侯廷军, 徐筱杰. 遗传算法在计算机辅助药物分子设计中的应用[J]. 化学进展, 2004, 16(1): 35—41.
- [16] 蔡文生, 邵学广, 潘忠孝. 遗传算法用于化学结构图的同态研究[J]. 高等学校化学学报, 1999, 20(1): 32—36.
- [17] 陈开, 李通化, 卢佩章. 遗传算法在色谱图解析中的应用[J]. 分析化学, 2003, 31(2): 58—162.
- [18] 陈闽军, 程翼宇. 基于遗传算法的色谱指纹峰配对识别方法[J]. 分析化学, 2003, 31(5): 513—517.
- [19] Leardi R, Lupiancz Gonzalez A. Genetic algorithms applied to feature selection in PLS regression: how and when to use them[J]. Chemometrics and intelligent laboratory systems, 1998, 41(2): 195—208.
- [20] Barros A S, Rutledge D N. Genetic algorithms applied to the selection of principal components[J]. Chemometrics and intelligent laboratory systems, 1998, 40(1): 65—82.
- [21] 郑延斌, 曹益林, 宋予民. 遗传算法用于结晶过程动力学参数辨识[J]. 计算机与应用化学, 1999, 16(2): 121—124.
- [22] 杨杰, 刘云, 汤漾平. 基于遗传算法的激光加工工艺参数选择[J]. 湖北工学院学报, 2003, 18(2): 14—15.
- [23] 章元, 朱尔一, 李静, 等. 模拟退火算法与遗传算法结合用于变量筛选[J]. 分析化学, 1999, 27(10): 1131—1135.
- [24] 夏之宁, 胡芳, 邱细敏, 等. 紫外光谱与遗传算法用于多组分氨基酸同时测定[J]. 重庆大学学报(自然科学版), 1998, 21(1): 107—112.
- [25] 孙卫国, 陈德创, 陈亚秋, 等. 基于遗传算法的神经网络为苯乙酸胺类农药构效关系建模的研究[J]. 高等学校化学学报, 1998, 19(6): 871—875.
- [26] 王雅琳, 黄泰松, 桂卫华, 等. 锌电解过程分时供电优化决策系统[J]. 矿冶工程, 1999, 19(1): 47—52.
- [27] 路文江, 汤富领, 刘树群, 等. 用神经网络—遗传算法优化镍基喷涂材料[J]. 甘肃工业大学学报, 1998, 24(1): 11—15.
- [28] 张玉祥. 小波神经网络遗传算法及其在矿山压力预报中的应用[J]. 中国有色金属学报, 1999, 9(2): 448—452.
- [29] 何耀华, 夏志忠, 华贲. 非机理模型优化问题的 NN—GA 方法研究[J]. 石油化工, 1999, 28(6): 372—380.