Evolutionary Optimization of Catalysts Assisted by Neural-Network Learning

Martin Holeňa[1], David Linke[2], and Uwe Rodemerck[2]

[1] ICS AS, Pod vodárenskou věží 2 | FIT CTU, Kolejní 2
Prague, Czech Republic
martin@cs.cas.cz
[2] Leibniz Institute for Catalysis, Albert-Einstein-Str. 29a, 18059 Rostock, Germany

Abstract. This paper presents an important real-world application of both evolutionary computation and learning, an application to the search for optimal catalytic materials. In this area, evolutionary and especially genetic algorithms are encountered most frequently. However, their application is far from any standard methodology, due to problems with mixed optimization and constraints. The paper describes how these difficulties are dealt with in the evolutionary optimization system GENACAT, recently developed for searching optimal catalysts. It also recalls that the costly evaluation of objective functions in this application area can be tackled through learning suitable regression models of those functions, called surrogate models. Ongoing integration of neural-networks-based surrogate modelling with GENACAT is illustrated on two brief examples.

Keywords: evolutionary optimization, mixed optimization, constrained optimization, neural network learning, surrogate modelling, evolutionary algorithms in catalysis.

1 Introduction

In chemical engineering, much effort is devoted to increasing the performance of industrially important reactions, i.e., to achieving a higher yield of the desired reaction products without higher material or energy costs. Over 90% of chemical processes use a catalyst to this end. Catalysts are materials that decrease the energy needed to activate a chemical reaction without being themselves consumed in it. They typically consist of several components with different purpose, which can be selected from among many substances. Chemical properties of those substances constrain the possible ratios of their proportions, but they still allow for an infinite number of catalyst compositions. Moreover, the catalyst can usually be prepared from the individual components in a number of ways, and the preparation method also influences its performance in the chemical process. Consequently, the search for new catalytic materials leading to optimal performance of a chemical reaction entails high-dimensional constrained optimization tasks. Their *objective functions* cannot be analytically described, their values must be *obtained empirically*. Commonly used smooth optimization methods are not convenient to this end. Indeed, to obtain sufficiently precise numerical

K. Deb et al. (Eds.): SEAL 2010, LNCS 6457, pp. 220–229, 2010.
© Springer-Verlag Berlin Heidelberg 2010

estimates of gradients or second order derivatives of the empirical objective function, those methods need to evaluate the function in points some of which would have a smaller distance than is the measurement error. That is why *methods not requiring derivatives* have been employed to solve those optimization tasks - both deterministic ones, in particular the *simplex method* and *holographic strategy*, and stochastic ones, such as *simulated annealing*, or *evolutionary algorithms* [1]. Evolutionary, especially *genetic algorithms* (GA) are encountered most frequently, but their application to this area is far from any standard methodology. Main obstacles on a way to such a methodology are *mixed optimization* with respect to continuous and discrete variables, and *constraints*.

This paper describes how those two obstacles are tackled in the evolutionary optimization system GENACAT, developed in recent years at the Leibniz Institute for Catalysis in Rostock, in collaboration with the Institute for Computer Science in Prague. The overall functionality of the system has been outlined, from the point of view of the application domain, in the Journal of Chemical Information and Modeling [2]. The present paper, on the other hand, explains the principles of the underlying evolutionary approach to mixed constrained optimization, which have not been published yet.

The evaluation of the empirical objective functions encountered in the optimization of catalytic materials is *costly and time-consuming*. In particular, testing a generation of materials proposed by an evolutionary algorithm typically needs several to many days of time and costs thousands of euros. Therefore, evolutionary optimization usually proceeds only for 5–10 generations in this application area. A common approach to the optimization of such objective functions is to evaluate the original objective function only sometimes, and to replace it otherwise with a suitable regression model learned from the available data and called *surrogate model* of the objective function [3,4,5,6]. Several successful applications of this approach have been already reported also in catalysis [7,8,9], inciting us recently to incorporate surrogate models based on two kinds of artificial neural networks into GENACAT. An explanation of the integration of surrogate modelling with this evolutionary system would exceed the extent of the paper. Therefore, it has been presented in a separate more comprehensible companion paper [10]. Here, we mainly document the usefulness of surrogate modelling in the evolutionary optimization of catalytic materials on examples, not included in [10].

In the next section, the optimization task entailed by the search for new catalytic materials is formalized, and the solution adopted in the GENACAT system is explained and illustrated on an example. Section 3, on the other hand, shows two brief examples of neural-network based surrogate modelling.

2 Constrained Mixed Optimization in the Search for New Catalytic Materials

In the search for new catalysts leading to the optimal performance of a chemical reaction, the individual coordinates of points in the input space of the objective function typically convey some of the following meanings: