Available online at www.sciencedirect.com
ScienceDirect
Computers and Chemical Engineering 32 (2008) 1707–1713
Computers & Chemical Engineering
www.elsevier.com/locate/compchemeng

Multi-objective steady state optimization of biochemical reaction networks using a constrained genetic algorithm

Hannes Link [a,*], Julio Vera [b,1], Dirk Weuster-Botz [a,2], Néstor Torres Darias [c,d,3], Ezequiel Franco-Lara [a,2,4]

[a] *Lehrstuhl für Bioverfahrenstechnik, Technische Universität München, Boltzmannstr. 15, 85748 Garching, Germany*
[b] *Lehrstuhl für Systembiologie und Bioinformatik, Institut für Informatik, Universität Rostock, Albert Einstein Str. 21, 18051 Rostock, Germany*
[c] *Departamento de Bioquímica y Biología Molecular, Grupo de Tecnología Bioquímica, Dpto. Bioquímica y Biología Molecular, Facultad de Biología, Universidad de La Laguna, 38206 La Laguna, Tenerife, Islas Canarias, Spain*
[d] *Instituto Universitario de Bioorgánica Antonio González, Avda. Astfco. Fco. Sánchez, s/n, 38206 La Laguna, Tenerife, Islas Canarias, Spain*

Received 14 December 2005; received in revised form 19 February 2007; accepted 20 August 2007
Available online 23 August 2007

Abstract

A hybrid genetic algorithm-based method to solve constrained multi-objective optimization problems is proposed. Considering operation around a steady state of a dynamical system, the task of the algorithm consists on finding a set of optimal, but constrained solutions. The method is exemplified on a (bio)chemical reaction network in *Saccharomyces cerevisiae*. In the steady state the model reduces to a system of non-linear equations which must be solved by a search method. This iterative search was integrated into a genetic algorithm in order to look up for optimal steady states. The basic idea is to use individuals of the genetic algorithm as starting points for the search algorithm. The optimization goal was to simultaneously maximize ethanol production and reduce metabolic burden. Two alternative kinetic approaches are compared to Michaelis Menten-type kinetics: a S-System and a generalized mass action model, both based on Power-Law kinetics.

Keywords: Multi-objective optimization; Genetic algorithm; Metabolic engineering; Power-law; Trust region dogleg method

1. Introduction

Many technological researches demand an improvement of a process, a product or an integrated system. The search for the best feasible enhancement can be included under the topic optimization, which will be considered in several biotechnological research fields in the next years. Usually the optimization target is given intrinsically by the model of the system considered. Based on this model, a mathematical formulation of the objective function is derived, which has to be improved via an optimization process. Models of (bio)chemical systems are described in terms of coupled ordinary differential equations (ODEs) accounting for the rates or fluxes of each of the n involved variables (species), $\mathbf{x} = (x_1, x_2, \ldots, x_n)$:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = f_i(\mathbf{x}, \mathbf{p}) \quad \text{for } i = 1, 2, \ldots n \tag{1}$$

Besides the variables, most models depend on parameters (collected in a vector $\mathbf{p}$), which usually have constant values. Since most complex systems exhibit non-linear behavior, their models consist of a set of non-linear functions for which is, in general, not possible to obtain the exact optimal solution. Instead of an analytical approach, numerical methods are used, which reach from linear programming (LP) to more sophisticated algorithms based on stochastic principles.

* Corresponding author. Tel.: +49 89 28915713; fax: +49 89 28915714.
E-mail addresses: H.Link@lrz.tum.de (H. Link), jv030@informatik.uni-rostock.de (J. Vera), D.Weuster-Botz@lrz.tum.de (D. Weuster-Botz), ntorres@ull.es (N. Torres Darias), E.Franco-Lara@lrz.tum.de (E. Franco-Lara).
[1] Tel.: +49 381 4987575; fax: +49 381 4987572.
[2] Tel.: +49 89 28915713; fax: +49 89 28915714.
[3] Tel.: +34 922 318334; fax: +34 922 318334.
[4] Present address: Institut für Bioverfahrenstechnik, Technische Universität Braunschweig, Gaußstraße 17, 38106 Braunschweig, Germany. Tel.: +49 531 391 7650; fax: +49 531 391 7652.

0098-1354/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2007.08.009

Optimization in mathematical terms is defined as the search of a certain decision vector $\mathbf{x} \in S_X$ such that the objective function ϕ takes an extreme objective value, $y = \phi(\mathbf{x})$, where $y \in S_Y$.

The optimization gets more complicated when the interest lies on problems considering multiple and even conflicting objectives. This leads to a multi-objective optimization problem (MOP) where, in general, it is not possible to obtain a single optimal solution but an efficient set of best alternatives. The efficient set of a MOP consists of decision vectors, which cannot be improved in any objective without degradation in other objectives. These vectors are known as Pareto optimal.

MOPs arising from (bio)chemical systems are often constrained by steady state conditions in order to guarantee a continuous optimal operation. At steady state, none of the variables change their value as a function of time, even though material is flowing through the system. Mathematically the situation requires that all ODEs in Eq. (1) are equal to zero, reducing the model to an algebraic system of non-linear equations (NLS):

$$F(\mathbf{x}, \mathbf{p}) = (f_1(\mathbf{x}, \mathbf{p}), \ldots, f_n(\mathbf{x}, \mathbf{p}))^{\mathrm{T}} = 0 \tag{2}$$

Any steady state of the model corresponds to a stationary operating point of a continuous process. A particular value $\mathbf{p} = \mathbf{p}'$ fixes design and operational parameters of the process and finding an appropriate value $\mathbf{p}'$ therefore amounts to designing the process (Marquardt & Mönnigmann, 2005).

With appropriate values for $\mathbf{p}'$, a solution of Eq. (2) can be given by a vector $\mathbf{x}^*$. To compute $\mathbf{x}^*$, Eq. (2) is solved iteratively by a search algorithm. Starting from a point $\mathbf{x}_0$, series of vectors $\mathbf{x}_1$, $\mathbf{x}_2$,... are generated taking into account a convergence criteria to approximate the solution asymptotically.

The constrained MOP consists then on the optimization (exemplified in Eq. (3) with the maximization) of the objective functions ϕ_i, which depend on the variables $\mathbf{x}$ and on the parameters $\mathbf{p}$:

$$\max \ \phi_i(\mathbf{x}, \mathbf{p}) \quad \text{for } i = 1, 2, ..., m \tag{3}$$

The objective functions might take maximal or minimal values under the condition that Eq. (2) holds true. In order to search for these efficient steady states, the variables and parameters are altered in a given predefined constrained search space for the systems variables and parameters, which are also considered as decision variables in the optimization process.

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \in S_x, \qquad \mathbf{p} = (p_1, p_2, \ldots, p_l) \in S_p \tag{4}$$

In general, there are few evolutionary algorithms developed for constrained multi-objective optimization problems (Sarker, Abbass, & Karim, 2001). In this work, optimization is carried out with a genetic algorithm (GA), which is based on a stochastic search considering only the steady state solutions of the system.

Biochemical reactions are enzyme catalyzed and commonly described in terms of Michaelis Menten (MM) kinetics. However, its structure has been criticized by some authors, who suggest kinetics that are based on the Power-Law formalism (Savageau, 1969a). Similar to chemical reactions the rate law depends on a rate constant α and the involved species x_i raised to a power, the kinetic order g_i.

$$v = \alpha \prod_{i=1}^{n} x_i^{g_i} \tag{5}$$

As the kinetic orders can take any value (even negative), these kinetics are also labelled fractal kinetics. Savageau (1969a,b) based the Biochemical Systems Theory (BST) on this approach. Within the BST are the S-Systems and generalized mass action (GMA) models the alternative representations of MM-type models.

S-Systems are non-linear models (as they are based on Eq. (5)) but suppose an effective aggregation of chemical fluxes into a net input and output fluxes in each differential equation. With this property, S-Systems models can be optimized in a linear domain by LP after a logarithmic transformation using the so called indirect optimization method (IOM) (Torres, Voit, González-Alcón, & Rodríguez, 1997). Accordingly, a performance comparison with the proposed non-linear method set against the efficient states of a S-System calculated with the IOM was possible. The evaluation was extended as well for the efficient (steady) states using a MM model, a GMA model and a S-System of a (bio)chemical reaction network for *Saccharomyces cerevisiae*. The present work shows how the non-linear method is able to compute solutions close to those of the true efficient set obtained with the IOM.

2. The multi-objective genetic algorithm

GAs were found useful for solving MOPs, as they have some advantages over traditional operational research techniques. For example, considerations for convexity, concavity, and/or continuity of functions are not necessary in GAs, whereas they form a real challenge in traditional optimization techniques. MOPs are considered as "difficult problems" in the specialized literature, being the constrained MOPs even more difficult (Sarker et al., 2001).

For this work, a modified version of the Strength Pareto Evolutionary Algorithm (SPEA) for multi-objective optimization is discussed. Comparative studies for a large number of case studies have shown that, among all major multi-objective GAs, the SPEA is clearly superior (Zitzler & Thiele, 1999). The SPEA uses two populations and is based on the principles of Pareto dominance. The original SPEA was modified in order to handle the steady state constraints. The simplified flow of the algorithm reads as follows:

- *Step 1*. Produce an initial population P of variables $\mathbf{x}_k \in S_x$ and of parameters $\mathbf{p}_k \in S_p$ for N individuals I

  $$I_k = (\mathbf{x}_k, \mathbf{p}_k) \quad \text{for } k = 1, \ldots, N$$

- *Step 2*. Solve the NLS $F(\mathbf{x},\mathbf{p}_k) = 0$ for $\mathbf{x}$ with a search algorithm starting at $\mathbf{x}_k$ and replace $\mathbf{x}_k$ by the solution $\mathbf{x}_k^*$ in the population P (see Section 3 for a detailed description).
- *Step 3*. Copy Pareto optimal members in P to an extern population P'.