Computers and Chemical Engineering 72 (2015) 87–108

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Design of memetic algorithms for the efficient optimization of chemical process synthesis problems with structural restrictions

M. Urselmann*, S. Engell

Process Dynamics and Operations Group, Technische Universität Dortmund, Emil-Figge-Str. 70, 44227 Dortmund, Germany

ARTICLE INFO

Article history:
Received 27 January 2014
Received in revised form 2 August 2014
Accepted 7 August 2014
Available online 19 August 2014

Keywords:
Chemical process synthesis
Memetic algorithms
MINLP
Optimization

ABSTRACT

In Urselmann et al. (2011a,b) we presented a memetic algorithm (MA) for the design optimization of reactive distillation columns. The MA is a combination of a problem-specific evolutionary algorithm (EA) that optimizes the design variables and a mathematical programming (MP) method that solves the continuous sub-problems with fixed discrete decisions which are proposed by the EA to local optimality. In comparison to the usual superstructure formulation, the search space of the MA is significantly reduced without excluding feasible solutions. The algorithm computes many different local optima and can handle structural restrictions and discontinuous cost functions. In this contribution, a systematic procedure to modify the MA to solve more complex design problems is described and demonstrated using the example of a reactive distillation column with an optional side- or pre-reactor with structural restrictions on the number of streams. New concepts to handle connected and optional unit operations are proposed.

1. Introduction

Chemical process synthesis in practice is usually done by a team of chemists and engineers in a trial-and-error process based on heuristics, experiences, and expert knowledge over quite long periods of time. To reach the long-term profitability of the process, decisions that have to be made are the choice of the chemical production route, the choice of the pieces of equipment and of their connections by flows of materials and the selection of the operating conditions of the different units. The design alternatives proposed by the experts are then validated by extensive numerical simulation studies of detailed mathematical process models. The structure of the chemical production process, i.e. the pieces of equipment and their interconnections, is fixed before simulation is applied to determine promising operating conditions and the state variables of a few alternatives which are compared afterwards to define the base case that is further developed in detail engineering.

In order to improve and to shorten this process, systematic methods were developed to find the optimal design among all suitable alternatives (see e.g. Biegler et al., 1997; Barnicki and Siirola, 2004; Westerberg, 2004). The optimization-based methods (Grossmann, 1985; Yeomans and Grossmann, 1999) work on a so-called superstructure which is a representation of all structures of the production process that seem promising and that should be considered during the optimization. Usually a large monolithic mixed-integer nonlinear model (MINLP) that comprises the mathematical models of the superstructure, the unit operations, the restrictions, and the cost function is formulated in the following form:

$$\min\ F = f(x, y, z) \tag{1}$$

$$s.t.\ h(x, y, z) = 0 \tag{2}$$

$$g(x, y, z) \le 0 \tag{3}$$

$$x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{N}_0,$$

where F is the economic cost function, $h(x, y, z) = 0$ are the equations that describe the behavior of the processing system (mass and heat balances, geometric relationships, etc.), and $g(x, y, z) \le 0$ are the inequalities that define the specifications or constraints for feasible designs (Grossmann et al., 2000). The continuous variables x denote the operating conditions, while the discrete variables z represent structural design decisions. The state variables of the process are given by the continuous variables y. The optimization of the structure of the process and the operating conditions (z and x are summarized by the term 'design variables') and the determination of the corresponding state variables y are usually done simultaneously.

* Corresponding author.
E-mail addresses: maren.urselmann@bci.tu-dortmund.de (M. Urselmann), sebastian.engell@bci.tu-dortmund.de (S. Engell).

http://dx.doi.org/10.1016/j.compchemeng.2014.08.006

The resulting MINLP are non-convex and can exhibit many local optima. The MINLP-techniques which are usually applied for the solution decompose the problem into a mixed-integer or integer master problem (MIP/IP) to handle the structural decisions z and continuous nonlinear sub-problems (NLP) to optimize the continuous degrees of freedom (x and y) for fixed variables z.

Iterative methods as e.g. DICOPT (Jackson and Grossmann, 2001) which is an implementation of the outer approximation method (Duran and Grossmann, 1986; Fletcher and Leyffer, 1994) or GBD (Ciric and Gu, 1994) which is based on the general Bender's decomposition (Geoffrion, 1972) compute lower bounds and optimize the binary variables by solving a sequence of alternate NLP subproblems with fixed discrete variables. Extended Branch&Bound methods (Borchers and Mitchell, 1992) solve the MILP master problem by a tree search and solve an NLP sub-problem in each node of the tree with fixed and/or relaxed discrete variables. The solution of the NLP sub-problems is addressed by efficient mathematical programming methods as e.g. reduced gradient methods (Murtagh and Saunders, 1982), sequential quadratic programming (Schittkowski, 1981) or interior point methods. The Extended Cutting Plane Method (Westerlund and Pettersson, 1995) applies successive linearizations instead of solving the NLP subproblems. Due to the non-convexity of the problems, the NLP solvers cannot guarantee to find the global optimum (Grossmann et al., 2000). They find local optima and possibly exclude the global optimum from the search. The quality of the local solutions depends strongly on the initialization of the search.

An increase of the number of discrete degrees of freedom leads to an exponential increase in the size of the search space (Seferlis and Grievink, 2001). If realistic unit operation models are used, the resulting monolithic optimization model can often not be solved by the state-of-the-art methods in reasonable computation times. The complexity of the superstructure optimization model limits the number of structural degrees of freedom that can be efficiently optimized.

To overcome this problem, different approaches have been proposed in the literature. Some of these methods reduce the complexity of the optimization model by reducing the number of design variables or of the design alternatives by including special modeling or decomposition techniques or by the application of iterative procedures. A common approach is to avoid the explicit modeling of the structural decisions of the process by discrete variables and to use continuous variables that implicitly indicate the existence of certain structures, e.g. when the continuous flow rate of a feed stream that enters a distillation column on a certain stage is zero this indicates that there is no feed stream in the actual design. The scope of these techniques however is limited. If, e.g., the number of certain elements (e.g. the number of feed streams that enter a distillation column) is restricted, an explicit representation of the structures cannot be avoided.

Stein et al. (2004) eliminate the discrete decision variables by adding a set of continuous variables and constraints that represent the discrete decision space of the optimization problem. This procedure was extended by Harwardt et al. (2011) and Kraemer et al. (2009). The quality of the local solution found by the method depends strongly on the initial values, this influence can be reduced by the application of a sequential solution procedure that relaxes the constraints that are introduced to force the continuous variables to integer values in the first step, and then tightens them subsequently.

Lang and Biegler (2002) eliminate integer variables within the design optimization of distillation columns with trays by the definition of a differentiable distribution function (DDF) in which the placement of feed and side streams and the number of the trays are represented by continuous variables. All feed and side streams are mapped to all column trays by the DDF and then the distributed flow rate of entry or exit streams are directed to a specific tray based on the value of its DDF at that tray.

The solution methods described above finally lead to a local solution of the MINLP that depends on the initialization.

Different multi-level approaches have also been successfully applied. Kravanja and Grossmann (1997), Daichendt and Grossmann (1998) and Bedenik et al. (2004) decompose the original problem according to the hierarchical decomposition procedure proposed by Douglas (1988) by a multilevel tree search. For each node of the tree, a reduced superstructure is optimized by MINLP techniques.

Another approach is to use short-cut methods, heuristics or expert knowledge (Skiborowski et al., 2012, 2013; Caballero et al., 2009) or preliminary screening (Daichendt and Grossmann, 1994) to reduce the number of alternatives that are considered during the optimization and therefore the number of structural decisions a priori.

Two-stage procedures of superstructure optimization with reduced complexity were also successfully applied. In their work on a general superstructure and a global optimization approach for the design of integrated process water networks that is suitable for global NLP and MINLP, Ahmetović and Grossmann (2011) proposed a two-stage procedure for the solution of large-scale industrial models. They first solve an NLP without binary variables and without the corresponding constraints, where some of the structures are represented implicitly by continuous variables. The superstructure is then reduced by fixing variables that correspond to continuous variables which were set to zero by the optimization, and the reduced superstructure is optimized in the second step. Barkmann et al. (2008) also applied a two-stage procedure where some structures of the process are set to reasonable values and are kept unchanged, while other structural decision variables are optimized in the first step. The best values of the latter are then fixed and the remaining discrete variables are optimized in a second step.

In the approaches mentioned above, the reduction of the complexity of the design problem may lead to an exclusion of the globally optimal solution from the search.

Other methods reduce the complexity of the process models by the use of surrogate models (Henao and Maravelias, 2011; Shao et al., 2007; Xiong et al., 2007; Won and Ray, 2005).

An alternative approach is the formulation as a generalized disjunctive program (GDP) first introduced by Raman and Grossmann (1994). The GDP formulation involves binary and continuous variables that are specified in algebraic constraint, disjunctions and logic propositions. Logic-based methods can then be used to solve the GDP. By this procedure the effort needed to optimize structural decisions can be reduced and nonlinearities can be handled more efficiently (Grossmann and Ruiz, 2012).

Besides the large scale of the models that arise from superstructure optimization formulations with rigorous unit models, the non-convexity of the solution set often leads to the existence of many local optima that represent structurally different designs. When local optimization is employed, the convergence to one of these local optima can only be influenced by the initialization of the search. Rigorous global optimization methods (Sahinidis, 1996b; Floudas and Gounaris, 2009) can determine the global optimum, but at the moment are not yet computationally feasible for large problem sizes, although substantial progress was made in this area in the last years (see e.g. the literature review in Misener and Floudas (2013)).

Moreover, the knowledge of the set of all or at least of many possible local optima is useful for the designers who use optimization to explore the design space; e.g. to detect intervals of design parameters which lead to very similar economic results so that the exact choice in this interval is not crucial.