ADVANCES IN MATHEMATICAL PROGRAMMING FOR THE SYNTHESIS OF PROCESS SYSTEMS

I. E. GROSSMANN\textsuperscript{1}, J. A. CABALLERO and H. YEOMANS

Department of Chemical Engineering, Carnegie Mellon University
Pittsburgh, PA 15213, USA

Keywords: Process synthesis, mathematical programming, reaction, separation, energy systems.

Abstract
This paper presents a review of advances that have taken place in the mathematical programming approach to process design and synthesis. A review is first presented on the algorithms that are available for solving MINLP problems, and its most recent variant, Generalized Disjunctive Programming models. The formulation of superstructures, models and solution strategies is also discussed for the effective solution of the corresponding optimization problems. The rest of the paper is devoted to reviewing recent mathematical programming models for the synthesis of reactor networks, distillation sequences, heat exchanger networks, mass exchanger networks, utility plants, and total flowsheets. As will be seen from this review, the progress that has been achieved in this area over the last decade is very significant.

1. Introduction
The mathematical programming approach to design, integration and operation problems, or more generally synthesis problems, consists of three major steps. The first is the development of a representation of alternatives from which the optimum solution is selected. The second is the formulation of a mathematical program that generally involves discrete and continuous variables for the selection of the configuration and operating levels, respectively. The third is the solution of the optimization model from which the optimal solution is determined. As will be shown in this paper, significant advances have taken place with this methodology, which offers the possibility of developing automated tools to support the exploration of alternatives and optimization of chemical processes by design engineers.

Over the last decade there have been considerable advances in mathematical programming techniques. For instance, the solution of mixed-integer nonlinear programming problems and the rigorous global optimization of nonlinear programs has become a reality. Furthermore, there have been great advances in the capability of solving very large problems, particularly for linear and mixed-integer linear programming techniques. There has also been recently a trend towards new logic-based formulations that can facilitate the modeling and solution of these problems. Finally, the availability of modeling systems that can facilitate the formulation of optimization problems has also made great progress, as well as the development of several solution strategies.

It is the objective of this paper to present an overview of the major advances in mathematical programming techniques and strategies for the modeling and solution of design and synthesis problems. The paper is organized as follows. We first present an overview of methods for mixed-integer linear and nonlinear problems, and their more recent formulation as generalized disjunctive programming problems. We also give a brief review of methods for global optimization.

2. Mathematical Programming
Design and synthesis problems give rise to discrete/continuous optimization problems, which when represented in algebraic form, correspond to mixed-integer optimization problems that have the following form:

\[
\begin{align*}
\min Z &= f(x, y) \\
\text{st.} & \quad h(x, y) = 0 \\
& \quad g(x, y) \leq 0 \\
& \quad x \in \mathcal{X}, \ y \in \{0,1\}
\end{align*}
\]

where \(f(x, y)\) is the objective function (e.g. cost), \(h(x, y) = 0\) are the equations that describe the performance of the system (mass and heat balances, design

\textsuperscript{1} Author to whom correspondence should be addressed
equations), and $g(x, y) \leq 0$ are inequalities that define the specifications or constraints for feasible choices. The variables $x$ are continuous and generally correspond to the state or design variables, while $y$ are the discrete variables, which generally are restricted to take 0-1 values to define the selection of an item or an action. Problem (MIP) corresponds to a mixed-integer nonlinear program (MINLP) when any of the functions involved are nonlinear. If all functions are linear it corresponds to a mixed-integer linear program (MILP). If there are no 0-1 variables, the problem (MIP) reduces to a nonlinear program (NLP) or linear program (LP) depending on whether or not the functions are linear.

The formulation and solution of major types of mathematical programming problems can be effectively performed with modeling systems such as GAMS (Brooke et al., 1992), and AMPL (Fourer et al., 1992). While these require that the model be expressed explicitly in algebraic form, they have the advantage that they automatically interface with codes for solving the various types of problems. They also perform automatic differentiation and allow the use of indexed equations, with which large scale models can be readily generated. It should also be noted that these modeling systems now run mostly on desktop and PC computers, making their use and application widely available.

The solution of LP problems relies largely on the simplex algorithm (Chvatal, 1983; Saigal, 1995), although lately interior-point methods (Marsten et al., 1990; Larsen et al., 1994) have received increased attention for solving very large problems because of their polynomial complexity. MILP methods rely largely on simplex LP-based branch and bound methods (Nemhauser and Wolsey, 1988) that consists of a tree enumeration in which LP subproblems are solved at each node, and eliminated based on bounding properties. These methods are being improved through cutting plane techniques (Balas et al., 1993), which produce tighter lower bounds for the optimum. LP and MILP codes are widely available. The best known include CPLEX, OSL and XPRESS, all which have achieved impressive improvements in their capabilities for solving problems. It is worth noting that since MILP problems are NP-complete it is always possible to run into time limitations when solving problems with large number of 0-1 variables, especially if the integrality gap is large.

The solution of NLP problems (Fletcher, 1987; Bazaara et al. 1994), relies either on the successive quadratic programming algorithm (SQP) (Han, 1976; Powell, 1978; Schittkowski, 1981), or on the reduced gradient method (Murtagh and Saunders, 1978, 1982). Major codes include MINOS and CONOPT for the reduced gradient method, and OPT (Vasantharajan et al., 1990) for the SQP algorithm. These NLP methods are guaranteed to find the global optimum if the problem is convex (i.e., convex objective function and constraints). When the NLP is nonconvex a global optimum cannot be guaranteed. One option is to try to convexify the problem, usually through exponential transformations, although the number of cases where this is possible is rather small. Alternatively, one could use rigorous global optimization methods, which over the last few years have made significant advances. These methods assume that special structures are present in the problem, such as bilinear, linear fractional and concave separable functions. Although this may appear to be quite restrictive, Smith and Pantelides (1996) have shown that algebraic models are always reducible to these structures, provided they do not involve trigonometric functions. For a general review on global optimization see Horst and Tuy (1993), Horst and Pardalos (1995); recent developments in chemical engineering can be found in Grossmann (1996). Computer codes for global optimization still remain in the academic domain, and the best known are BARON by Sahinidis and Ryoo (1995), and $\alpha$-BB by Floudas et al. (1996). It should also be noted that non-rigorous techniques such as simulated annealing (Kirkpatrick et al., 1983) and genetic algorithms (Goldberg, 1989), which have also become popular, do not make any assumptions on the functions, but then they cannot guarantee rigorous solutions, at least in finite amount of time. Also, these methods do not formulate the problem as a mathematical program since they involve procedural search techniques that in turn require some type of discretization. Furthermore, violation of constraints is handled through ad-hoc penalty functions.

Major methods for MINLP problems include first Branch and Bound (BB) (Gupta and Ravindran, 1985; Natar and Schrage, 1991; Borchers and Mitchell, 1994; Stubbs and Mehrotra, 1996), which is a direct extension of the linear case, except that NLP subproblems are solved at each node. Generalized Benders Decomposition (GBD) (Geoffrion, 1972), and Outer-Approximation (OA) (Duran and Grossmann, 1986; Yuan et al., 1988; Fletcher and Leyffer, 1994; Ding-Mai and Sargent, 1992), are iterative methods that solve a sequence of alternate NLP subproblems with all the 0-1 variables fixed, and MILP master problems that predict lower bounds and new values for the 0-1 variables. The difference between the GBD and OA methods lies in the definition of the MILP master problem; the OA method uses accumulated linearizations of the functions, while GBD uses accumulated Lagrangian functions parametric in the 0-1 variables. The LP/NLP based branch and bound by Quesada and Grossmann (1992) essentially integrates both subproblems within one tree search, while the Extended Cutting Plane Method (ECP) (Westerlund and Pettersson, 1995) does not solve the NLP subproblems, and relies exclusively on successive linearizations. All these methods assume convexity to guarantee convergence to the global optimum. Nonrigorous methods for handling nonconvexities include the equality relaxation algorithm by Kocis and Grossmann (1987) and the augmented penalty version.
of it (Viswanathan and Grossmann, 1990). A review on these methods and how they relate to each other can be found in Grossmann and Kravanja (1997). The only commercial code for MINLP is DICOPT (OAGAMS), although there are a number of academic versions (MINOPT by Floudas and co-workers, α-EC by Westerlund and co-workers).

In recent years a new trend that has emerged in the formulation and solution of discrete/continuous optimization problems through a model that is known as Generalized Disjunctive Programming (GDP) (Raman and Grossmann, 1994). The basic idea in GDP models is to use boolean and continuous variables, and formulate the problem with an objective function, and subject to three types of constraints: (a) global inequalities that are independent of discrete decisions; (b) disjunctions that are conditional constraints involving an OR operator; (c) pure logic constraints that involve only the boolean variables. More specifically, the problem is given as follows:

\[
\begin{align*}
\min \, Z & = \sum_{x} c_x + f(x) \\
\text{s.t.} \quad & g(x) \leq 0 \\
& \forall \alpha, \exists A \, h_{\alpha}(x) = 0 \quad k \in K \\
& c_i = \gamma_{i} \\
& \Omega(y) = \text{True} \\
& x \in X, \, y_{\alpha} \in \{\text{True, False}\}
\end{align*}
\]

where \( x \) are continuous variables and \( y \) are the boolean variables. The objective function involves the term \( f(x) \) for the continuous variables (e.g. operating cost) and the charges \( c_x \) that depend on the discrete choices. The equalities/inequalities \( g(x) \leq 0 \) must hold regardless of the discrete conditions, and \( h_{\alpha}(x) = 0 \) are conditional equations that must be satisfied when the corresponding boolean variable \( y_{\alpha} \) is True for the \( \alpha \)th term of the \( k \)th disjunction. The set \( I_k \) represents the number of choices for each disjunction defined in the set \( K \). Also, the fixed charge \( c_k \) is assigned the value \( j_k \) for that same variable. Finally, the constraints \( \Omega(y) \) involve logic propositions in terms of boolean variables.

Problem (GDP) represents an extension of disjunctive programming (Balas, 1985), which in the past has been used as a framework for deriving cutting planes for the algebraic problem (MIP). It is interesting to note that any GDP problem can be reformulated as a MIP problem, and vice-versa. It is more natural, however, to start with a GDP model, and reformulate it as a MIP problem. This is accomplished by reformulating the disjunctions using the convex hull transformation (Turkay and Grossmann, 1996b) or with "big-M" constraints. The propositional logic statements are reformulated as linear inequalities (Raman and Grossmann, 1991; 1994). For the linear case of problem GDP, and when no logic constraints are involved, Beaumont (1991) proposed a branch and bound method that does not rely on 0-1 variables and branches directly on the equations of the disjunctions. This method was shown to outperform the solution of the alternative algebraic MILP models. Raman and Grossmann (1994) developed a branch and bound method for solving problem GDP in hybrid form; i.e. with disjunctions and mixed-integer constraints. For this they introduced the notion of "w-MIP representability" to denote those disjunctive constraints that can be transformed into mixed-integer form without loss in the quality of the relaxation. Hooker and Osorio (1999) developed a different branch and bound method which in a way is a generalization of Beaumont's method in that it does not introduce 0-1 variables, and addresses problems directly in the form of the GDP problem.

For the nonlinear case of problem (GDP), and for the case of process networks, Türkay and Grossmann (1996) proposed a logic-based Outer-Approximation algorithm. This algorithm is based on the idea of extending the Outer-Approximation algorithm by solving NLP subproblems in reduced space, in which constraints that do not apply in the disjunctions are disregarded, with which both the efficiency and robustness can be improved. In this method the MILP master problems correspond to the convex hull of the linearization of the nonlinear inequalities. Also, several NLP subproblems must be solved to initialize the master problem in order to cover all the terms in the disjunctions. Penalties can also be added to handle the effect of nonconvexities as in the method by Viswanathan and Grossmann (1990). This method has been implemented in the computer prototype LOGMIP, a GAMS-based computer code developed by Vecchiotti and Grossmann (1997). Finally, it should be noted that a new method for solving GDP problems has recently been reported by Lee and Grossmann (1999). These authors have developed reformulations and algorithms that rely on the convex hull of nonlinear convex inequalities. Also, their method is not restricted to process networks.

From the above review, it should be clear that LP and MILP codes have become quite powerful. NLP methods are being advanced by rigorous global optimization algorithms, which, however, can still be relatively expensive to apply. Finally, as for MINLP methods the new exciting direction is logic based optimization methods, such as Generalized Disjunctive Programming, which promise to facilitate problem formulation and improve the solution efficiency and robustness.

3. Superstructures

In the application of mathematical programming techniques to design and synthesis problems it is always necessary to postulate a superstructure of alternatives. This is true whether one uses a high level aggregated model, or a fairly detailed model. Most of the previous work has relied on representing the superstructure for each particular problem at hand, but without following some general principles. There are
two major issues that arise in postulating a superstructure. The first is, given a set of alternatives that are to be analyzed, what are the major types of representations that can be used, and what are the implications for the modeling. The second is, for a given representation that is selected, what are all the feasible alternatives that must be included to guarantee that the global optimum is not overlooked.

As for types of superstructures, Yeomans and Grossmann (1998) have characterized two major types of representations. The first is the State-Task Network (STN) which is motivated by the work in scheduling by Kondili et al. (1993). The basic idea here is that the representation makes use of two types of nodes: states and tasks (see Fig. 1). The assignment of equipment is dealt implicitly through the model. Both the cases of one-task one-equipment [OTOE] or variable task equipment assignment [VTE] can be considered.

![State Task Network superstructure representation for distillation sequence](image)

Fig. 1. State Task Network superstructure representation for distillation sequence

The second representation is State Equipment Network (SEN) which is motivated by recent work of Smith (1995), and where the basic idea is to work with two types of nodes: states and equipment (see Fig. 2). The tasks in this case are treated implicitly through the model. This representation considers the case of variable task equipment assignment [VTE]. Yeomans and Grossmann (1998) have developed GDP models for each of the two different types of representations. These can then be used for solution with a GDP algorithm, or they can be used for reformulation as MILP or MINLP problems.

![State Equipment Network superstructure representation for distillation sequences](image)

Fig. 2. State Equipment Network superstructure representation for distillation sequences.

As for the issue on how to systematically generate the superstructure that includes all the alternatives of interest, Friedler et al. (1993) have proposed a novel graph theoretic approach that has polynomial complexity to find all the interconnections in process networks, given that nodes for processes and chemicals are specified. This procedure has been successfully applied for synthesizing process networks for waste minimization (Friedler et al., 1995). These authors have also used these ideas to perform more efficiently the search in the optimization (Friedler et al., 1996).

4. Models

Closely related to the selection of the superstructure, is the selection of level of detail of the optimization model. A common misconception about the mathematical programming approach is that models are always detailed and require a great deal of information. This, however, is not necessarily true. In general mathematical programming models can be classified into three main classes:

a) Aggregated models. These refer to high level representations in which the design or synthesis problem is greatly simplified by an aspect or objective that tends to dominate the problem at hand. Examples of aggregated models include the transshipment model for predicting minimum utility and minimum number of units in heat exchanger networks (Papoulias and Grossmann, 1983) and mass exchanger networks (El-Halwagi and Manousiouthakis, 1989a), the set of heat integration constraints based on the pinch location method (Duran and Grossmann, 1986; Grossmann et al., 1998), distillation models for minimizing cost of
utilities (Caballero and Grossmann, 1998), reactor network models for maximizing yield (Balakrishna and Biegler, 1992a). All these models are specific to the corresponding problem at hand. Daichendt and Grossmann (1994) have outlined a theoretical framework for deriving aggregated models, which however, must be adapted to each particular application.

b) Short cut models. These refer to fairly detailed superstructures that involve cost optimization (investment and operating costs), but in which the performance of the units is predicted with relatively simple nonlinear models in order to reduce the computational cost, and/or for exploiting the algebraic structure of the equations, especially for global optimization. Examples of such models include synthesis models for heat exchanger networks (Yce et al., 1990; Ciri and Floudas, 1991), distillation sequences (Aggarwal, and Floudas, 1990; Yeomans and Grossmann, 1998b), and process flowsheets (Kocis and Grossmann, 1989; Türkay and Grossmann, 1995a).

c) Rigorous models. These also rely on detailed superstructures, but involve rigorous and complex models for predicting the performance of the units. The area of synthesis of distillation sequences (ideal and non-ideal) is perhaps the one that has received the most attention for developing rigorous models. Examples are the work by Bauer and Stichlmair (1996,1998) and Smith and Pantelides (1995).

It should be noted that aggregated models give rise to simpler types of optimization models. They are often LP, NLP or MILP models of modest size that are simpler to solve than larger MINLP models. In contrast, both short cut and detailed models give rise almost exclusively to MINLP problems, which as mentioned above, can also be formulated as GDP problems. The important point to realize here is that mathematical programming can accommodate models of various degree of complexity.

5. Synthesis Strategies

There are several solution strategies that can be used in the optimization of mathematical programming models for design and synthesis. The two major strategies are simultaneous optimization, and the sequential optimization. In the simultaneous strategy a single model is optimized at once. The optimization is rigorous because all the trade-offs are taken simultaneously into account. Also, the simultaneous models are commonly of one type, but hybrids are possible. For example one can perform simultaneous optimization of a flowsheet in which the reaction, separation and heat integration are each represented by aggregated models. Alternatively, simultaneous optimization can be applied to the synthesis of subsystems, for example heat exchanger networks, or heat integrated distillation units. In the latter example, one might use detailed models for the distillation and heat integration, or detailed for distillation and aggregated for heat integration.

The sequential optimization strategy consists of solving a sequence of subproblems, normally at an increasing level of detail. The major motivation is to solve simpler problems to avoid solving a large single problem (normally MINLP). A good example is the procedure implemented in MAGNETS (Floudas et al., 1986) in which an LP is solved first to target the utility cost, next an MILP to determine the identity of the fewest number of matches, and finally an NLP superstructure in which interconnection of exchanger with the predicted matches are determined. Another example is the hierarchical decomposition procedure by Douglas (1988) in which the flowsheet is sequentially optimized through various levels, from a simple input-output model to the detailed structure.

While the simultaneous and sequential strategies have been used for a long time, there are several new variants that have been proposed. These include combining mathematical programming with physical insights (Gundersen and Grossmann, 1990), the state-space approach by Bagajewicz and Maniousiuothakis (1992), and the combined Hierarchical Decomposition and MINLP optimization by Daichendt and Grossmann (1998). These are discussed later in the paper. What is important to realize is that mathematical programming models can be applied through a variety of strategies, that in turn make use of models at various levels of complexity. For instance, in sequential decomposition one can use aggregated, short-cut and detailed models. In simultaneous optimization one can use a model of a single type, or a mix of several types. Finally, it is possible to use mathematical programming in combination with other approaches, most notably, physical insights.

6. Review of Synthesis Models Based on Mathematical Programming

After providing a general overview of general developments in mathematical programming techniques for process synthesis, we present in the remainder of this paper a review of models that have been proposed for subsystems (reactor networks, distillation sequences, heat exchanger networks, mass exchange networks, utility systems) and of process flowsheets. As will be seen, the amount of published papers in this area has been very substantial over the last decade or so.

6.1 Reactor Networks

Synthesis of reactor networks poses a difficult modeling problem as these are usually described by differential-algebraic equations. Compared, however, to heat exchanger networks, distillation systems or utility systems, the combinatorial part in reactor networks tends to be smaller.

Two significant mathematical programming strategies for synthesis of reactor networks are superstructure optimization and targeting. The
superstructure approach can be sub-optimal since the solution obtained is only rich as the initial superstructure chosen and it is difficult to ensure that all the possible networks are included in the initial superstructure. In targeting, the objective is to find an achievable bound to a performance index of the system independently of the actual reactor configuration. Although these two approaches appear independently, the concepts developed in the targeting approach are being used to generate superstructures that can ensure that the optimal solution is included.

In superstructure optimization Chitra and Govind (1981, 1985) studied PFR systems with a recycle stream from an intermediate point along the reactor, and optimized the recycle ratio as well as the point of recycle. The objective function was based on the yield of the reactions. Achenie and Biegler (1986, 1988) postulated a series parallel combination of axial dispersion reactors (ADR) and the advantage of the ADR is that more general reactor networks can be generated. Kokossis and Floudas (1989, 1990, 1991) postulated a large superstructure of isothermal networks of PFR's and CSTR's. They modeled the PFRs by series of CSTRs of equal size. Thus their MINLP formulations had not differential equations. Kokossis and Floudas (1994) extended their previous formulation to handle stability of reactor networks and integration with recycle streams, and the nonisothermal case (Kokossis and Floudas, 1994b), that includes options for pure adiabatic operation, options for perfectly controlled units and directly or indirectly intercooled or interheated reactors. Markoulaki and Kokossis (1996), used stochastic optimization (simulated annealing) to solve the complex MINLP associated with the formulation of the reactor network in an attempt of minimizing the effect of the nonconvexities. Smith and Pantelides (1995) proposed a synthesis technique for reaction and separation networks using detailed unit operations models. Complete connectivity among the units, both forward and recycle, was assumed in the superstructure chosen. In all these methods the optimal solution is as good as the superstructure, but there is no guarantee that the best solution is included in the formulation of the superstructure.

The targeting reaction network synthesis is based in the concept of “attainable region” which was first suggested by Horn (1964). The attainable region is the convex hull of concentrations that can be achieved starting from the feed point by reaction and mixing. Glasser et al. (1987) and Hildebrandt et al. (1990) developed the geometrical concepts that allow to obtain the entire region in the concentration space that is attainable for a given feed concentration through reaction and mixing. Although the graphical representation of the attainable region is constrained to two or three dimensions, Feinberg and Hildebrandt (1992) show that the geometric insights gained from this representation can be useful in problems with higher dimensions. In this work some of the characteristics of the attainable region were pointed out and formally established in the recent work of Feinberg and Hildebrandt (1997). In particular they showed that the boundary of the attainable region is made up of PFR trajectories and straight line segments. As a result, all points of this boundary can be found by a combination of PFRs, CTRs, and differential side streams reactors (DSRs). Due to the convex nature of the attainable region it is possible to formulate superstructures formed only by combination of these three reactors that will ensure the inclusion of the optimal solution with respect to a target objective. However, constructive procedures for higher dimensional attainable regions that incorporate these properties still need to be developed. Omtveit and Lien (1993) extended the representation of the attainable region to account for recycles in flowsheets. Glasser et al. (1994) compared the conversion obtained with segregated and maximum mixed reactors with the bounds given by the attainable region. These authors showed that, in general, these reactors do not bound the attainable region.

Balakrishna and Biegler (1992a, b) and Lakshmanan and Biegler (1996a) formulated the geometric technique for targeting as a mathematical programming model. They proposed a general targeting procedure based on optimization flows between regions of segregation (PFR) and maximum mixedness as a mixed integer dynamic optimization problem. Since this is an optimization based procedure, it overcomes the dimensionality problem of the geometric technique and can be extended to nonisothermal systems where the temperature profile is an additional control profile. Lakshmanan and Biegler (1996b) extended the model to deal with the problem of simultaneous chemical reaction and mass integration. Balakrishna and Biegler (1993) also developed a targeting model for reaction separation and energy management.

Bikic and Glavic (1995) addressed the problem in which the reactor network has multiple feeds. The problem is solved in two stages. In the first, candidate reactors that satisfy at least the necessary conditions for the optimum reactor network are generated and then, in the second stage, the optimal flowsheet is extracted by optimizing the process that comprises all the candidate reactors. The method is restricted to 2 and 3 dimensional reaction schemes. Later, Bikic and Glavic (1996) extended the model to more complicated cases, when reactions take place in nonisothermal systems with external heat sources and sinks. Hopley et al. (1996) extended the attainable region for the case of reversible reactions with complex kinetics, and used the attainable region to obtain the optimal structure.

6.2 Distillation Sequencing

Distillation has been, after heat exchanger networks, the most studied of all the subsystems in process synthesis. Although distillation is an expensive operation in terms of capital and operating costs, it
continues to be the most important separation technique, even for nonideal and azeotropic mixtures. A general review of distillation synthesis can be found in Westerberg (1985), Floquet et al. (1988, 1994), Gert-Jan et al. (1994), Juergen et al. (1995), and Westerberg and Wahnjaß (1996). Several approaches have been proposed for the design of efficient separation systems: heuristic methods (Seader and Westerberg, 1977), evolutionary techniques (Stephanopoulos and Westerberg, 1976), mean-end analysis (Sirola and Rudd, 1971), hierarchical decomposition (Douglas, 1988), implicit enumeration (Johns and Romero, 1979), dynamic programming (Fraga and McKinnon, 1995; Fraga, 1996), stochastic methods (Fraga and Matias, 1996), phenomenon based design (Laroche et al., 1992; Rooks et al., 1998; Hauan and Lien, 1998) and superstructure optimization, which will be treated extensively in the following paragraphs.

One of the pioneering works in superstructure optimization is due to Sargent and Gaminbandara (1976). These authors proposed a superstructure of linked columns that includes not only simple sharp splits, but also complex columns such as the Petyuk configuration (Petyuk et al., 1965). A nonlinear programming model was used to solve this model that consists of tray by tray mass balance equations and vapor-liquid equilibrium assuming constant K values. Andreuovich and Westerberg (1985) presented a network superstructure for the separation of near ideal mixtures into pure components. The model considered only sharp splits and is based in short cut methods, but its major contribution was the formulation of a MILP problem associated with the superstructure representation. The model is suitable for the design of heat integrated distillation sequences, and uses a network representation instead of a tree representation (Hendry and Hughes, 1972).

The model of Kakhu and Flower (1988) is possibly the first model capable of describing complex column configurations explicitly, such as side strippers, side rectifiers and Petyuk configurations. It is worth noting that the model of Sargent and Gaminbandara (1976) can also be rearranged to tackle these configurations. Floudas and Paules (1988) proposed a MINLP model for the design of heat integrated distillation processes for the separation of ideal boiling multicomponent mixtures that extended the Andreuovich and Westerberg model, by allowing the use of nonlinear cost functions. The model is restricted to sharp splits and constant reflux ratio. In the work of Floudas and Anastasiadis (1988) and Aggarwal and Floudas (1990) the model is extended to processes with multiple feeds and non-sharp splits. These nonlinear models used simulation results to construct approximate models of mass and energy balances, and cost correlations. Paules and Floudas (1992) solve the problem of finding the optimal process for a non-constant feed composition. Floudas (1987) also addressed the problem of separating a multicomponent feed stream into several specified multicomponent product streams.

Novak et al. (1996) considered the synthesis of distillation sequences simultaneously with the synthesis of other process subsystems with heat integration. They proposed to use a smaller and more compact superstructure for the distillation sequence rather than the tree or network representation. In particular, for non-azeotropic mixtures of N components they used a superstructure with N-1 columns in which each column performs one specific cut between adjacent components.

All the previous models rely on short-cut methods or other simplifications, and the phase equilibrium is often described by assuming constant relative volatilities. The first general MINLP model for a rigorous distillation column design was presented by Viswanathan and Grossmann (1993a, b). They applied the MESH equations in a rigorous tray by tray calculation. The model allows the optimization of the number of stages, the optimal feed tray location, even for multiple feeds, and the optimal reflux. In their work the objective functions considered were the optimization of a combination of the number of trays and reflux, or the simplified function of Lucia and Kumar (1988). However, the model was not applied for the synthesis of distillation sequences. Smith (1995) proposed a superstructure representation, with a reduced number of columns, similar to that of Novak et al. (1996), with full connectivity among them, using the Viswanathan and Grossmann model. He used an enumeration procedure and a global optimization method for the NLPs. Recently Bauer and Stichlmair (1996, 1998) used the Viswanathan and Grossmann model with a superstructure for zeotropic and azeotropic distillation sequences based on the concept of preferred separation.

It is not until recently that greater attention has been paid to the systematic generation of superstructures and the interaction of superstructures and modeling for distillation systems. Bagajewicz and Manousiouthakis (1992) introduced the concept of state space representation for separation networks design. These authors claimed that this approach allows for a specific number of process operators and all possible interconnections among them, providing a framework for network design with minimal assumptions on process structure. In particular, they applied the methodology to distillation columns treated as a combination of heat and mass exchange networks through a pinch operator. Papalexandri and Pistikopoulos (1996) proposed a general framework for the synthesis of distillation sequences based on mass and heat exchange building blocks for constructing superstructures. This representation allows the construction of flexible superstructures with many alternatives for task operations. Sargent (1988) proposed recently a superstructure representation for the synthesis of both zeotropic and azeotropic systems, which is based on a state task network representation.
6.3 Heat-Exchanger Network Synthesis

Heat Exchanger Network Synthesis (HENS) is by far the most developed technique for which many methods and software packages are available. The last extensive review was given by Gundersen and Naess (1988). The discovery of the heat recovery pinch (Umeda et al., 1979; Limhöff et al., 1979), that is derived through thermodynamic analysis, provided the basis for advancing synthesis techniques for HENS. The most widely used method, commonly known as “pinch technology” (Linhöff and Hindmarsh, 1983) relies on the use of targets (energy, number units, area) and is based on a user-driven approach. SUPERTARGET and ADVENT are two major pieces of software implementing this approach. It should also be noted that ideas of pinch analysis are being expanded beyond HENS to total sites and to assessment of environmental problems (see Limhöff, 1993, for a review).

As for mathematical programming methods, there has been a gradual evolution from LP/MILP/NLP methods that are based on targets (Cerda and Westerberg, 1983; Papoulos and Grossmann, 1983; Floudas et al., 1986; Colberg and Morari, 1990; Gundersen and Grossmann, 1990) to simultaneous MINLP models in which networks are automatically synthesized and energy, area, and number of units optimized simultaneously (see Yee and Grossmann, 1990; Cerca and Floudas, 1991). Examples of software for this purpose include MAGNETS for the target-based methods and SYNHEAT for the simultaneous MINLP. As a recent example of the former approach, Galli and Cerda (1998) have proposed an MILP model for a heat exchanger network structure that explicitly accounts for relative location of heat transfer units, splitters and mixers. This sequential MILP model allows the designer to explicitly account for desired topology features.

The MINLP models that have been proposed to solve the HENS problem are large nonlinear and nonconvex models, and therefore may get trapped in sub-optimal solutions. Recent efforts have been aimed at addressing these limitations. Daichendt and Grossmann (1994b) proposed a preliminary screening procedure for heat exchanger networks. The strategy is based on convex aggregate models and thermodynamic insight to reduce the superstructure while preserving the optimality of the solution. Quesada and Grossmann (1993) developed a global optimization method for networks with fixed structure, and assuming linear cost and arithmetic mean driving force. Zamora and Grossmann (1998) proposed a global MINLP optimization algorithm for the synthesis of HENS, using the superstructure representation of Yee and Grossmann (1990). The algorithm is based on two new sets of convex underestimators for the heat transfer area, the first one is based on thermodynamic analysis and the second is a relaxation of the heat transport equation. The models developed are solved with a hybrid branch and bound/outer-approximation
search method. Yee and Grossmann (1991) expanded their staged representation MINLP method to retrofit design problems, where they considered the reconnection of existing equipment, the expansion of the available areas, and the inclusion of new heat exchangers in the superstructure. They also proposed a prescreening procedure to reduce the complexity of the MINLP model that is generated.

Several attempts have been made to integrate both the mathematical programming and pinch analysis approaches, in order to reduce the computational difficulties of the former and to improve the detailed synthesis of the latter. Duran and Grossmann (1986a) proposed an aggregated MINLP model that uses a pinch location method to calculate the minimum utility consumption of a process. This model has the advantage of being easy to embed in any mathematical programming synthesis model, to perform simultaneous flowsheet synthesis and heat integration. Recently, Grossmann et al. (1998) developed further the Duran and Grossmann model to rigorously account for isothermal streams. The proposed model is based on the big-M representation of a Generalized Disjunctive Programming (GDP) model, and was applied to the synthesis of heat integrated distillation sequences. Gundersen et al. (1996) extended the Vertical MILP model for heat exchanger network synthesis originally developed by Gundersen and Grossmann (1990). This vertical MILP model was based on the idea of selecting matches that transfer heat vertically between composite curves. The extended vertical MILP model accounts for curve shifting effects and stream pairing, by means of a penalty term in the objective function that is derived from film coefficient data. Also along the lines of an integrated mathematical programming/pinch analysis approach, Zhu (1997) proposed an automated design method for heat exchanger networks using block decomposition and heuristic rules. The basic idea is to simplify a design problem by decomposing it into a number of blocks (i.e. stages in Yee and Grossmann, 1990), where special properties are exploited through decomposition (Zhu et al., 1995). After decomposition, design is carried out with a simple MINLP model aided with heuristic rules. Briones and Kokossis (1996) proposed a rigorous and systematic method for retrofit design of heat exchanger networks. The approach uses both mathematical programming and pinch analysis techniques in three steps: (1) targeting of area and modifications, (2) structural optimization and (3) network optimization.

Other efforts have concentrated on increasing the robustness of the designs predicted by the MINLP models. Papalexandri and Pistikopoulos (1994a,b) developed a systematic framework for the synthesis and retrofit of heat exchanger networks, where issues of flexibility and controllability are addressed. The algorithm is based on a multiperiod hyperstructure network representation, where explicit structural controllability criteria are developed and included in a MINLP model. The outcome of the model is an economically optimal design that is able to operate within a specified range of uncertainty in flows, temperatures and heat transfer coefficients. They also address the identification of control variables and the selection of a control structure to implement. Koukman et al. (1995) have presented a similar approach for the design of controllable heat exchanger networks. The authors solve an optimization problem that considers the exchanger model equations and constraints simultaneously for all possible predefined disturbance directions. Regarding stochastic optimization approaches, Nielsen et al. (1996) proposed a modeling framework that describes each of the potential heat exchangers in terms of their functional behavior, in an object-oriented representation. The proposed network is solved with Simulated Annealing to find the best process alternative. Also, Aithier et al. (1997) developed a similar Simulated Annealing procedure for HEN synthesis.

6.4 Mass Exchange Networks

Motivated by applications in waste recovery systems, El-Halwagi and Manousiouthakis (1989a,b; 1990a) considered the problem of synthesizing mass exchange networks. For the simpler case when concentration targets are specified for single components, interesting analogies can be drawn with the heat exchanger network problem. Usually implicit in the formulation of the problem is the assumption that there are no temperature changes within the MEN. Finding the minimum utility consumption for a fixed target MEN synthesis task is formulated as a linear program. The solution of the LP determines the minimum cost and pinch points that limit the mass exchange between rich and lean streams (El-Halwagi and Manousiouthakis, 1990a). In a second stage, a MILP transshipment problem is solved to identify the minimum number of mass exchange units, in a similar manner than the MILP formulation for fixed target HEN synthesis (Papoulas and Grossmann, 1983). El-Halwagi et al. (1992) applied this approach to the specific problem of phenol treatment in petroleum refinery wastewater. The minimum utility cost problem becomes a MINLP for mass exchange networks that include regeneration (El-Halwagi and Manousiouthakis 1990b; Garrison et al., 1995). Papalexandri et al. (1994) have considered simultaneous optimization models in contrast to the sequential design strategies used by other authors. Lee and Park (1995) propose an alternative method for MEN synthesis, using the P-graph theory and a NLP formulation to find the optimal mass exchange network. Papalexandri and Pistikopoulos (1995, 1996) proposed a general process synthesis framework based on the representation of synthesis alternatives via mass and heat transfer mechanisms. They introduced a multipurpose mass/heat exchange module to represent the building block for any process operation involving
mass/heat exchange between two previously defined streams.

The problem with multiple component targets was addressed by El-Halwagi and Manousiouthakis (1989b), Gupta and Manousiouthakis (1994) and Wilson and Manousiouthakis (1998). Gupta and Manousiouthakis (1993) presented an approach where supply and target compositions of the rich and lean streams are allowed to vary within upper and lower bounds. The problem is solved as a MINLP in which the global optimum can be ensured. Later, Gupta and Manousiouthakis (1996) proposed a new formulation of the problem that leads to a linear program that yields the same solution that the previous MINLP formulation. The reduction in complexity of the new formulation allows to deal with larger problems. Wang and Smith (1994a) and Kuo and Smith (1997) approached the general problem for the design of the final disposal wastewater network using a graphical representation for the superstructure of design alternatives. Wang and Smith (1994b, 1995) also treated the problem of multiple contaminants, regeneration re-use, regeneration recycling, and flowrate constraints. Alva-Arregaz et al. (1998) proposed a solution approach based on a recursive MILP to optimize the Wang and Smith (1994a) model. With this approach optimality is not guaranteed. Galan and Grossmann (1998) have presented a global optimization strategy based on NLP and MINLP models to address the nonconvexities that arise in the mass balances of the superstructures presented by Wang and Smith (1994a).

El-Halwagi et al. (1996) proposed an approach that removes the pollutants from in-plant streams instead of dealing with the pollutants in the terminal waste streams. It also provided a framework for the simultaneous consideration of gas and liquid pollutants. The problem was formulated as a MINLP in which the objective was to determine the optimum interception locations, extents and separating agents through the plant. Chang and Hwang (1996) developed a multiobjective programming approach for cost minimization and global emissions minimization objectives, in the synthesis of utility systems of chemical processes. The problem of simultaneous design and control of MEN is presented in Huang and Edgar (1995) and Huang and Fan (1995) and the flexible performance was addressed by Papalexandri and Pistikopoulos (1994) and Zhu and El-Halwagi (1994, 1995).

El-Halwagi and Srinivas (1992) introduced the problem of synthesizing reactive mass exchange networks (REAMENs). The main objective of synthesizing REAMENs is to preferentially transfer certain species from a set of rich streams to a set of reactive mass separating agents (MSAs), whereby the undesirable species may be converted into other chemical forms. In this work the problem was restricted to linear or convex equilibrium relations. The problem of synthesizing REAMENs with general equilibrium relations was presented by Srinivas and El-Halwagi (1994a). This work used a sequential approach, in which first the minimum cost of the MSAs is identified without any commitment to the final network structure. The second stage is aimed at minimizing the number of mass exchangers while realizing the minimum cost of MSAs. Lakshmanan and Biegler (1995) developed reaction-network targeting strategies for waste minimization using multi-objective optimization.

The problem of simultaneous waste reduction and energy integration was studied by Srinivas and El-Halwagi (1994b). These authors introduced the combined heat and reactive mass exchange network (CHARMEN). Systematic design techniques have also been devised for other separation systems that can be used in recyle/reuse networks. These include the design of heat-induced separation networks (HISEN) in which the removal of the pollutants is accomplished via heating/cooling so as to affect a phase change (Dye et al., 1995; Richburg and El-Halwagi, 1995; El-Halwagi et al., 1995; Dium et al., 1995) and pressure driven membrane separation (Srinivas and El-Halwagi, 1993; El-Halwagi, 1992). Finally, Dantus and High (1996) used an approach that combines the MINLP techniques with the capabilities of simulators (in particular ASPEN PLUS) for the retrofit of chemical processes through waste minimization and process integration. A good compendium of mass exchange networks and their applications can be found in the book of El-Halwagi (1997).

6.5 Utility Systems

The objective of a utility plant is to supply energy demands to industrial process plants in form of electrical, mechanical and steam demands. The first papers that used mathematical programming were based in LP models such as that of Petroulas and Reklaitis (1984). Papoulias and Grossmann (1983) introduced the MILP formulation for the structural and parametric optimization of utility systems under fixed steam and power demands. Fixing the operating conditions such as pressures and temperatures, yields linear energy balance equations. The MILP approach of Papoulias and Grossmann has been recently used for the multiperiod optimization of utility plants by Hui and Natori (1996), and Iyer and Grossmann (1997), and in multi-objective approaches for waste minimization in utility plants (Chang and Hwang, 1996).

Colmenares and Seider (1989) proposed an NLP model for the design of a utility plant integrated with a chemical process. It was based on the temperature interval method and the development of a superstructure of Rankine cycles. However, due to the nature of the NLP model there is no possibility of choosing among different turbine configurations, or for selecting electric motors for mechanical power demand. Kalivantzeff (1991), and Diaz and Bandoni (1996) used MINLP techniques to optimize the operation of the plant, but their models are not applicable to the synthesis of new utility plants.
Marechal and Kalivtsev (1991) proposed a MILP formulation that allows to tackle the optimal integration of the utilities to satisfy the energy requirement of the process at minimum cost. Later these authors (Marechal and Kalivtsev, 1996, 1997a) proposed a sequential approach for integrating utility systems. The procedure is divided in three steps. In the first one, pinch analysis concepts are used to determine the possible utilities to be used. In the second the MILP formulation is used to select the optimal configuration of the utility system that minimize the cost of energy. In the third step they try to convert the numerical results into practical solutions. In a more recent paper, these authors propose a slightly different approach to the selection of the optimal utility systems (Marechal and Kalivtsev, 1998) using a sequential procedure consisting of three steps. In the first one a generic utility system superstructure is used to identify what are the technology requirements of the process. The model of the superstructure is based on the Effect Modeling and Optimization (Marechal and Kalivtsef, 1997b) concepts that use linear models of the technologies, and then integrate these models in a MILP method to identify the best solutions. In the second step an expert system is used to identify the available technologies able to satisfy the requirements of step 1. The objective of the third step is to target the optimal process configuration.

Mavromatis and Kokossis (1998a,b) proposed a methodology that combines the target objectives with the optimization. They proposed a procedure at various levels. They introduce the turbine hardware model (THM) that accounts for the variation of efficiency with the turbine size. The procedure is applicable to any type of units and can provide accurate estimates of its performance over the entire operation range. It can be used for realistic targets at the early stages of design as well as for selecting the steam levels that maximize the potential of power cogeneration. They divide the synthesis procedure into three stages. The first one is the development of the design components to be used for the synthesis structure. In the second stage is the optimization of the previous structure in order to minimize the losses due to the variation in operation. The optimum model is a result of the application of the THM model formulated as an MILP. And finally the analysis and synthesis of complex turbines as the results of the optimization stage can be further processed to systematically reveal compact utility networks that facilitate the objectives and analyze the results for competitive alternatives.

Bruno et al. (1998) proposed a rigorous MINLP for the synthesis and operation of utility systems that can be implemented in actual industrial problems. The optimal solution is selected from a superstructure similar to the one proposed by Papoulias and Grossmann (1983), containing conventional utility plant equipment specified by the designer for each demand. The retrofit of alternatives is addressed by fixing some of the options available in the model to match the equipment options considered. The resultant MINLP model has been implemented in the interactive computer program STEAM that automatically generates the model. Maia et al. (1995) also proposed an approach in which a superstructure very similar to that proposed by Papoulias and Grossmann is optimized. The main difference with previous approaches is that the authors consider only the equipment available in standard capacities, by handling discrete variables and discontinuous cost functions. In this work the authors used simulated annealing instead a deterministic technique to solve the mathematical programming problem. Maia and Qassim (1997) extended the previous formulation to systems with time-varying utility demands. Wilkendorf et al. (1998), also used simulated annealing for the automatic synthesis of a complete utility system. Their approach is also based in a superstructure similar to the originally proposed by Papoulias and Grossmann (1983), but in this case the authors extend the formulation to time dependent processes.

6.6 Process Flowsheet Synthesis.

The current state of flowsheet synthesis is represented by two major approaches: (1) hierarchical decomposition (Douglas, 1985; Douglas 1988; Douglas 1990) and its computer implementation PIP (Process Invention Procedure, Kirkwood et al., 1988); (2) mathematical programming (Grossmann 1985; Grossmann 1990a,b) and its initial computer implementation in PROSYN-MINLP (PROcess SYNthesis, Kravanja and Grossmann, 1990). These approaches can be regarded as complementary to each other (Rippin, 1990).

The hierarchical decomposition technique breaks the synthesis procedure into five decision levels: (1) Batch versus Continuous, (2) Input-Output Structure of Flowsheet, (3) Recycle Structure and Reactor Considerations, (4) Separation Systems and (5) Heat Exchanger Network. At each decision level beyond the first, the economic potential of the project is evaluated and a decision is made whether or not further work on the project is justified. This method utilizes heuristics, short-cut design procedures, and physical insight to develop an initial base-case design. The approach is motivated by Douglas's claim that only 1% of all designs are ever implemented in practice, and thus this screening procedure avoids detailed evaluation of most alternatives. Relying on heuristics, this approach cannot rigorously produce an optimal design, and, although heuristics often lead to good designs, they are fallible (see Papoulias and Grossmann, 1983; Colmenares and Seider, 1987; Fonyő and Muzsey, 1990). Furthermore, due to the sequential nature of the flowsheet synthesis, interactions among the design variables at the various decision levels may not be properly accounted for, as it is necessary to solve for them simultaneously. For
instance, Duran and Grossmann (1986a) and later Lang et al. (1988) have shown that simultaneous optimization and heat integration of flowsheets generally produces significant improvements in the profit compared to the sequential approach. Despite these shortcomings, hierarchical decomposition provides a useful approach for generating an initial flowsheet and alternatives (i.e., a base-case design and superstructure). It also provides a framework, when coupled with the concept of simultaneous synthesis of the complete flowsheet, for decomposing the synthesis problem into a hierarchy of detailed and aggregated models, that is then simpler to solve than the entire flowsheet, while still reflecting the presence of downstream tasks.

Other attempts that have been made to implement the hierarchical decomposition include the software package PROSYN (Schmebeckel et al., 1994). It employs heuristic rules to derive flowsheet configurations and uses detailed analysis of different unit operations. The heuristic rules are available in the form of expert systems for particular types of flowsheet elements. Also along the lines of Artificial Intelligence, Bieszcad et al. (1998) have developed MODELLA, a phenomena-based modeling language for process systems engineering. This language is fully declarative, and allows model development at the level of chemical engineering knowledge (i.e., a phenomena-based mechanistic description), not mathematical equations. Another tool that can aid the process synthesis approach based on hierarchical design is ICAS (Integrated Computer Aided System), which was developed by Gani et al. (1998). It is based on process simulation with an equation-oriented approach. ICAS is composed of a model generator, a simulator, a problem definition interface and toolboxes for tackling particular problems.

The mathematical programming approach utilizes optimization techniques to select the configuration and parameters of the processing system (Grossmann, 1985, 1990a,b, 1996). A superstructure containing alternative processing units and interconnections are modeled as discrete, binary variables (0-1) to denote the existence (1) or nonexistence (0) of that unit. An Outer-Approximation (OA) algorithm for solving MINLP problems has been developed and successively refined (Duran and Grossmann, 1986b; Kocis and Grossmann, 1987, 1989a,b; Viswanathan and Grossmann, 1990) and made widely available in the program DICOPT++ within the modeling system GAMS (Brooke et al., 1988). This algorithm partitions the problem into two parts: (1) an NLP subproblem, where initially the continuous variables for a single flowsheet configuration are optimized and the remaining alternative substructures are then suboptimized for the given flows, and (2) linearization of the nonlinear equations, through which an MILP master problem is obtained, that then determines a new optimal flowsheet configuration (i.e., new set of binary variables) for the next NLP subproblem. The Outer Approximation algorithm is guaranteed to obtain the global optimal solution for convex problems, but it does not guarantee global optimality for non-convex problems, and can get trapped in poor, local solutions (Daichendt and Grossmann, 1994a,b). This problem is to a large extent due to "disappearing" units whose variables become undefined when the corresponding streams take zero values for their flows.

In order to address the MINLP optimization of process flowsheets more effectively, Kocis and Grossmann (1989b) developed the Modeling/Decomposition (M/D) strategy in which the basic objective is to solve NLP subproblems pertaining only to the existing part of the superstructure. This not only avoids the solution of NLP problems of larger dimensionality, but reduces numerical difficulties such as singularities that arise in the case of non-existing units which have zero flows. The M/D strategy, which was successfully illustrated with the synthesis of the HDA flowsheet problem by Douglas (1988), has been implemented in PROSYN-MINLP and extended in various ways by Kravanja and Grossmann (1994). Because simulators are widely used to model a process, Kravanja and Grossmann (1996) proposed some modifications to the original M/D strategy to substitute the use of the equation-based solvers (i.e. GAMS) with a simulator for solving the non-linear synthesis subproblems. Diwekar et al. (1992) also proposed an implementation of the M/D strategy in ASPEN, and Reneaume et al. (1995) developed a new formulation of a MINLP optimization problem that handles equations in a modular environment. The formulation of Reneaume et al. introduces the concepts of pseudo torn streams and pseudo variables, to link the information provided by the simulator ProSim with a MINLP optimization algorithm. Diaz and Bandoni (1996) and Diaz et al. (1997) have solved large scale chemical plants synthesis problems by applying the Outer Approximation algorithm using a process simulator in the solution of the NLP subproblems.

As for methods that combine different approaches for flowsheet synthesis, Bagajewicz and Manousiouthakis (1992) introduced the state-space representation of synthesis alternatives, where a superstructure of alternatives is constructed with two blocks: a stream mixing/splitting block and a process unit block. The units in the unit block are modeled either by a superstructure operator (unit operation) or a pinch analysis operator. Kovac and Glavic (1995) addressed the optimal design for complex and energy intensive processes by combining thermodynamic (pinch) methods and MINLP techniques. The approach consists of two steps: (1) eliminate unpromising structures from a superstructure by studying an Extended Grand Composite Curve. The superstructure obtained is then optimized with MINLP techniques. Daichendt and Grossmann (1998) developed a combined hierarchical decomposition method with MINLP optimization. The basic idea in this method is
to rely on "black box" aggregated models for reaction, separation and heat integration that are optimized within a tree enumeration to avoid solving a single large-scale MINLP model. Kravanna and Grossmann (1997) also explored an alternative search procedure in which the tree enumeration is avoided by iteratively solving MINLP problems at various levels using integer cuts. Both these methods were successfully applied to the synthesis of the HDA flowsheet problem.

In terms of exploring new representations for flowsheet synthesis, Papalexandri and Pistikopoulos (1996) proposed a generalized modeling framework for process synthesis based on fundamental mass/heat-transfer principles. They introduced a mass/heat-transfer module as the building block of the framework, and developed block-superstructure rules to represent conventional and unconventional flowsheets. With this strategy, it is possible to formulate an MINLP model. Because this algorithm describes the problem in terms of tasks rather than unit operations, it is an aggregated representation of the synthesis problem. Following a more conventional approach to represent a synthesis problem, Smith and Pantelides (1995) proposed a state-equipment representation, where they have stream properties and unit operations as the superstructure building blocks. Their major goal was to avoid the combinatorial complexity of the synthesis problem by allowing full connectivity of all the unit operations in the superstructure. The model generated from this representation was then solved with global optimization techniques. The major drawback of this approach is the introduction of zero flows in most of the streams of the representation, which lead to singularities. Recently, Yeomans and Grossmann (1998) proposed a systematic modeling framework for process synthesis problems that relies on two different types of superstructure representations: State Task Network (STN) and State Equipment Network (SEN). To understand the close relationship between modeling and these representations, GDP models were developed in order to analyze their structure. Interestingly, numerical results have shown that neither representation is superior over the other. Friedler et al. (1993, 1995) have concentrated their research efforts in the automatic generation of superstructures for linear process networks. They proposed a graph-based algorithm that automatically generates structures suitable for synthesis with a guarantee that the search space is sufficient to include the optimal solution. The proposed method is polynomial in time.

Perhaps one of the major recent advances in the modeling and solution of flowsheet synthesis problems are the logic-based techniques. Raman and Grossmann (1994) have proposed the Generalized Disjunctive Programming (GDP) model for the mathematical formulation of flowsheet synthesis problems. For the linear case these problems can be effectively tackled with branch and bound methods that symbolically integrate logic to greatly reduce the enumeration of the nodes in the tree (Raman and Grossmann, 1993). For flowsheets with linear models the Logic Based Outer Approximation Algorithm was developed by Türkay and Grossmann (1996a). This algorithm extends the original Outer Approximation Algorithm of Duran and Grossmann (1986b) for the GDP model, and also represents a formalization of the M/D strategy by Kocis and Grossmann (1989b). The NLP subproblems of the algorithm are constructed with the equations and constraints of disjunctions that correspond to equipment units that are selected at each iteration of the algorithm. The Master problem of the Logic-Based OA algorithm, which predicts these choices, is constructed by the application of the convex hull of disjunctions to the linearized disjunctive model. Türkay and Grossmann (1996b, 1998) further extended the application of the Logic Based OA algorithm to the design of equipment with discontinuous cost functions, and illustrated its application with the synthesis of a vinyl chloride plant.

7. Concluding Remarks

From the review presented in this paper, it should be clear that mathematical programming has become a major methodology in the area of process synthesis. Advances in algorithms and modeling systems for solving various types of optimization problems, better understanding of issues related to superstructures, models and solution strategies have greatly helped to advance this field. From the review it should also be clear that over the last decade there has been extensive development of mathematical programming models for subsystems such as reactor networks, distillation systems, heat and mass exchange networks, utility plants, and total process flowsheets. All these models have the feature that they can be used as a basis for developing automated design tools that can effectively help to support design engineers. Also, they can be combined with approaches that are based on physical insights, such as pinch analysis, and attainable regions. In terms of future research directions, we can expect that new advances in global optimization and generalized disjunctive programming will have a great impact for improving the capability of optimizing synthesis models. This should also help to promote the development of synthesis models that are based on rigorous performance models, and to expand their scope for handling issues related to operability and uncertainty.

Finally, although there has been significant progress, another important challenge that remains is for engineers in industry to become more knowledgeable and proficient about mathematical programming techniques and their application in process design and operation. Availability of tools and modeling systems is no longer an issue. The major issue is the skill for formulating problems. A recent textbook that includes this topic for undergraduate and
graduate courses is the book "Systematic Methods for Chemical Process Design" by Biegler et al. (1997). It is hoped that more books like these will promote a modern approach to process design and synthesis. We are confident that engineers in the 21st century will no longer have to question the value of mathematical programming because they do not understand it.

Acknowledgments
The authors would like to acknowledge financial support from the National Science Foundation under Grant CTS-97103, Consejo Nacional de Ciencia y Tecnología (CONACYT) México, and from the Fulbright Program (MEC-Fulbright 98).

References
Mathematical Programming

MINLP Methods


Reactor Networks


Distillation


Barbosa, D. and Doeherty, M.F., "The Influence of Equilibrium Chemical Reactions on Vapor Liquid
Seader, J.D. and Westerberg, A.W., “A Combined Heuristic and Evolutionary Strategy for Synthesis


Heat Exchanger Networks


Mass Exchange Networks


Latin American Applied Research


Utility Systems


**Process Flowsheets**


Latin American Applied Research


Received August 30 1999;
accepted for publication April 14 2000
Recommended by Bandoni, Díaz and T. Pinto