A ROBUST COMBINATORIAL APPROACH BASED ON P-GRAPH FOR SUPERSTRUCTURE GENERATION IN DOWNSTREAM BIOPROCESSES

Xiaohui Xu, Chunying Zhu*, Youguang Ma and Haihua Song

School of Chemical Engineering and Technology, State Key Laboratory of Chemical Engineering, Tianjin University, Tianjin 300072, China.
Phone: + 86 (0) 22 2740 4772
E-mail: zhchy971@tju.edu.cn

(Submitted: November 12, 2013 ; Accepted: June 2, 2014)

Abstract - In the production of chemicals from fermentation, the cost of downstream purification constitutes the major portion of the total production cost. Since the bioprocess generally contains a large amount of biochemical separation units, its flowsheets are complex. How to generate the rigorous superstructure of the downstream bioprocess is a primary problem and key step. In this work, a robust combined approach based on the P-graph was proposed to generate the rigorous superstructure of the downstream process of fermentation. This method integrates the hierarchical decomposition of the heuristics with P-graph-based algorithm MSG with the advantages including: (1) Different design flowsheets could be evaluated simultaneously; (2) An unique bipartite graph, P-graph, was introduced, which could represent the maximum structure clearly and intuitively; (3) The rigorous superstructure could be automatically generated. The validity of the present method was verified with two practical bioprocesses. Results show that the effective processes and the adequate operation units could be determined in the beginning stage of the design, and the tedious reforming steps during process synthesis could be avoided.

Keywords: Process synthesis; P-graph; Rigorous superstructure; Separation network synthesis; Bioprocess.

INTRODUCTION

In process synthesis, heuristics and the algorithmic method are two major classical methods. Heuristics are established on the basis of past experience and existing knowledge or databases, and the procedure for implementing heuristic methods is usually direct. Since experiences always have the limitation of subjectivity, pure heuristic results are uncertain. Heuristics have the deeper possibility of missing the optimal solution, while it is a powerful selection means at the beginning of process synthesis. The mathematical programming methods, i.e., algorithms, which are recently the research priority of process synthesis, can provide a rigorous systematic framework for modeling the problem and automatically implementing synthesis. However, the so-called rigorous superstructure, also defined as the maximal structure, must be constructed in advance, before the algorithmic method is initiated. The maximal structure is the largest and the simplest combined structure, in which all feasible solutions must be covered. So how to automatically and correctly generate the maximal structure is the development direction of the algorithm.

Early algorithmic methods were only used to treat simpler problems and solve mixed integer non-linear programming (MINLP). Because the computational...
difficulty increases almost exponentially with the size of the problem, the algorithm could not guarantee that the superstructure of complex problems was generated explicitly, and the optimal solution was always missed in the result. Even now, it is difficult for algorithmic methods to generate the maximal structure during the process synthesis of large complicated problems.

Friedler et al. (1992; 1993; 2006) introduced a novel algorithm, i.e., the P-graph theory algorithm, which is efficient for solving large industrial process synthesis problems. The superstructure generated by the P-graph theory algorithm is rigorous superstructure, i.e., the maximal structure (Friedler, 2006; Friedler et al., 1993; Kovács et al., 2000). Recently, the P-graph theory algorithm has been widely used to generate the maximal structure in conceptual design (Varga et al., 2010; Bertok et al., 2013a; Bertok et al., 2013b; Fan, 2003). However, the necessary conditions, the scopes of feeds \( R \), products \( P \), all material \((m)\) and all operating units \((o)\), must be determined in advance before the P-graph method is performed. But how to determine these conditions exactly and integrally is also pivotal for the generation of the maximum structure. Since these necessary conditions usually were decided via heuristic methods, a combination of two classical methods is actually needed in the process synthesis. Rippin (1990) concluded that the two methods can be regarded as complementary each other rather than competitive. The combination has the advantages of the heuristic method and the algorithm. Therefore, using the combined method to cope with complex problems is the development trend of process synthesis.

Various combined methods appeared in the early literature for total flowsheet process synthesis (Mizsey and Fonyo, 1990; Daichendt and Grossmann, 1998). Among these combined methods, heuristics usually was used to produce initial flowsheets, and then the initial flow structures were modeled and solved to find the optimal solution by the algorithm method. For example, Mizsey and Fonyo (1990) introduced the hierarchical methods into the conceptual-design phase to create the preliminary flowsheets, and imported the user-driven synthesis technique to examine the implication of all constraints, then applied algorithmic methods to find the optimal solution of the superstructure. Daichendt and Grossmann (1998) used the combined method of hierarchical decomposition and mathematical programming techniques to optimize the entire flowsheet at the input-output subsystem level, reaction subsystem level, separation subsystem level and heat-integration subsystem level. They utilized at each level an abstract black-box model to represent the structure of the flowsheet and perform the optimization, which was mainly employed to solve a large and solvable MINLP. Sarkozi (2003a; 2003b) suggested the combined method of P-graph algorithms to synthesize reactions or azeotropic-distillation systems, which merged the means-ends analysis (Sirola, 1996) and the P-graph technique.

Replacing traditional chemical processes with bioprocesses is one vital development direction of the green chemical industry. However, the tedious subsequent separation of downstream bioprocesses often restricts the application of the biochemical method. With development of biochemical technology, numerous separation units were demanded, the biochemical products stream was more complex and, until now, the biochemical separation process could not be automatically synthesized by means of computers. In this paper, a new combined method based on the P-graphs theory algorithm is suggested to automatically generate the rigorous superstructure of biochemical separation process. In the combined method, the physical insights of analysis, the level-by-level strategy of hierarchical design procedure, and the goal-oriented-analysis were introduced to gather all in-process materials; then the idea of phenomena-driven design was combined with the hierarchical decomposition approach, which could guarantee that all correlative operation-units were found; finally, the MSG (maximal structure generation) algorithm was used to generate the maximal structure. The usefulness and accuracy of the new combined method was verified through the separation of butanol (B) (Liu et al., 2004; Liu et al., 2006; Fan et al., 2008) and the purification of penicillin (Swartz, 1985; Steffens et al., 1999).

**THE NOVEL COMBINED METHOD BASED ON P-GRAPH**

**The Algorithm for the Maximal Structure**

The algorithmic approach was an important part of the new combined method. The classical algorithmic method could not guarantee that superstructure was produced explicitly, since the computational difficulty almost always increased exponentially with the size of the process synthesis problem. The ambiguity of superstructure would make missing the optimal solution possible. However, Friedler et al. (1992; 1993) applied the combinatorics concept to facilitate
the rigorous superstructure generation. Here a special directed bipartite graph, the P-graph, has been used to represent the uniqueness of the structures of process systems explicitly, which avoids the simple graphs uncertainty. The P-graph approach was just the algorithmic approach improved mathematically by Friedler et al. (1993). The approach relied heavily on both graph theory and combinatorial techniques. Armed with a set of exact algorithms of the P-graph, it is advantageous to generate rigorous superstructure for solving large industrial process synthesis problems. Since the P-graph approach was applied to process synthesis by Friedler et al. (1993), the popularization of P-graph occurred, especially because separation-network synthesis (SNS) showed encouraging results.

The foundation of the P-graph method includes three cornerstones: a novel graph representation, in terms of P-graphs; a set of rigorous axioms, portraying the unique structural features of process networks; and a set of exact algorithms, the MSG algorithm for maximal structure generation, the SSG algorithm for solution structure generation, and the ABB algorithm for generating the optimal and near optimal solutions. Because the P-graph-based approach contained a set of rigorous axioms generated by the mass-conservation law and stoichiometric principle, the validity of a MSG algorithm for maximal structure generation (rigorous superstructures) could be guaranteed.

As the scopes of feeds, products, and all potential operating units were determined, the MSG algorithm could create superstructures explicitly by computer to solve large industrial process synthesis problems. But only the product (P) and the raw material (R) conditions are only known in process synthesis problems. Here, heuristic method should be applied first to analyze problems and determine all intermediate materials and operating units, and then an algorithmic construction of the superstructure performed by the P-graph method.

The Heuristics for the Maximal Structure

(1) Hierarchical Decomposition Approach

Although the introduction of the P-graph approach could simplify complex process synthesis problems and guarantee the generation of superstructures explicitly by computer, the scopes of feeds, products, and all operating units must be determined in advance by heuristics. For example, when the P-graph approach was used to generate all feasible sequences for a distillation system, the space of residue curve map was partitioned into lumped materials through the thermodynamic boundaries and pinches, and the operating units should be identified on the basis of these lumped materials (Feng et al., 2000; Feng et al., 2003). Thus the residue curve maps gave insights into how to carry out these separations.

The idea of the hierarchical decomposition approach was proposed by Douglas (1988). He suggested that the design process had broad choices at one level, transferring to subsidiary choices at lower levels, and increasing detail at each level, as illustrated in Figure 1, which shows the hierarchical selection procedure for a separation system. The top layer in the frame is a broad task definition of “a separation system”. The second layer is the possible approaches for realizing separation on the basis of the mixture properties, such as decantation, distillation, absorption, extraction or membrane permeation and so on. The down layer considers alternative types of separation process, and identifies the range of process conditions for the feasibility of each type. As a design activity, separation-network synthesis (SNS) was characterized by a particularly large number of alternatives. But the conventional separation-network synthesis methods almost always supposed that the available separators were of the same class, which was unreasonable. (Heckl et al., 2009).

![Figure 1: The hierarchical selection procedure for one separation system.](image-url)
The hierarchical decomposition was a hierarchical decision process regarding tasks. The hierarchical decomposition was a target-directed approach and also a domain analysis methodology to overcome the complexity of large engineering tasks, which guaranteed that it was especially effective for providing a framework for process synthesis. Separation-network synthesis (SNS) was strongly hierarchical, and the larger problem became manageable when an order was imposed on decision making. Due to its sequential nature, the major limitation of hierarchical decomposition is that it is impossible to manage the interactions between different design levels. Here hierarchical decomposition was introduced only to analyze the problem and determine the necessary conditions for the P-graph approach, which can effectively avoid its major limitation.

(2) Introduction of Phenomena-Driven Design

Heckl et al. (2007) proposed an algorithmic synthesis of an optimal separation network, in which various classes of candidate separators perform separations by different mechanisms. Because various classes of separators could enlarge the search space, the separation network generated by various classes of separators was far superior to that generated by a single class of separators. The idea of Heckl was similar to phenomena-driven design. Phenomena-driven design proposed that reasoning should not start at the level of building blocks but at a low level of the phenomena that occur in those building blocks. This design method takes the occurring phenomena as the ‘heart’ of the process, and the design tasks were decomposed at various levels by asking the following questions in sequence: what is desired, where can it be achieved (in which unit), when (under which conditions), and how can it be achieved (Gavrila and Iedema, 1996)? In the decomposition process, phenomena-driven design offered a systematic way to generate the desired phenomena and favorable conditions in order to achieve the design objective. It could explore new units and processes to support creative design, because the unit operation was not limited (Li and Kraslawski, 2004).

So the idea of phenomena-driven design and the hierarchical decomposition approach were also introduced into the new combined method. Thus, the search space could be greatly enlarged and a vital operating unit would not be missed. Therefore, the necessary conditions for the P-graph approach could be determined.

Outline of a Novel Combined Method Based on P-Graph

The new combined method took advantage of the insights provided by physicochemical properties and the idea of hierarchical decomposition, and simultaneously utilized the power of mathematical programming to deal with complex interactions and trade-offs.

The framework of the new combined method based on P-graph is shown in Figure 2. The interaction between process designers and the P-graph-based algorithm in the new combined method could reduce the arduousness of large process synthesis problem and avoid the missing of vital operation units. The outline of the new combined method based on P-graph is listed below.

First, we define the problem space scope. The raw materials are considered to be an initial state. The desired products are the goal state.

Second, according to the hierarchy of separation process design, the intermediate materials can be specified by the difference in attributes between raw materials and product. The purpose of an industrial chemical process is to apply some recognized rules

![Figure 2: The framework of the P-graph-based combined method.](image)
in sequence so that the raw materials are transformed into the desired products, i.e., their property differences in identity, amount, concentration, phase, temperature, pressure and form were systematically eliminated (Douglas, 1985). All intermediate materials were determined by the target-directed hierarchical decomposing tasks.

A pre-requisite for any synthesis method is the ability to determine the state (i.e., composition, temperature etc) of the streams which leave each unit. Based on different mechanisms including mainly volatility, solubility, permeability, adsorbability, density, etc., all operating units must be involved in option sets. The operating unit sets were determined by the idea of phenomena-driven design.

Third, the maximal structure was generated by the MSG algorithm, after all material and all operation units were determined. Unlike other approaches, the rigorous graph-theoretic method (P-graphs) could ascertain a rigorous and efficient superstructure to solve a MINLP or MILP problem through the concept of holistic analysis.

In all, the new combined methods based on P-graph were likely to involve a greater degree of interaction with other parts of the process innovation enterprise, including all various classes of candidate operating units. Since the new combined methods based on P-graph provide a transparent design procedure for a concept design in the spirit of hierarchical methods and the MSG algorithm, it was highly efficient to generate the superstructure of a complex industrial separation-network synthesis automatically by computer.

### THE COMPARISON WITH OTHER COMBINED METHODS

Compared with the earliest combination method (Mizsey and Fonyo, 1990; Daichendt and Grossmann, 1998), the main improvement of the new combined method is the introduction of the MSG algorithm of P-graph theory, which can automatically and efficiently generate a rigorous superstructure of the process synthesis problem. Among early combination methods, heuristics were used to generate the initial structure, and then a mathematical programming method was applied to solve the structure for the optimal solution. In the new combined method, heuristics is only used as an analysis to search for the material set \( \{m\} \) and operating unit set \( \{o\} \) of the separation process, and then the P-graph MSG algorithm can automatically generate the maximal structure, i.e., the rigorous super-structure. The MSG algorithm was highly efficient to automatically generate the superstructure of the separation-network synthesis for a complex industrial process synthesis by computer. Moreover, in early combined methods, the hierarchical decomposition method was used for dividing a total process into input-output layer, reactor layer, separating layer, and heat-integration layer. In the new combined method, the hierarchical decomposition method is only used to analyze the separation process for finding the necessary conditions of the MSG algorithm. Therefore, some disadvantages of the hierarchical decomposition method, such as subjectively producing super-structure and incongruously optimizing among layers, can be avoided.

The new combined method also differs from Sarkozi’s (2003a; 2003b) method, though the P-graph theory algorithm is contained in both methods. Firstly, different P-graph algorithms were used for different purposes. In Sarkozi’s method, the SSG algorithm was applied to generate all single flowsheets. In this paper, the MSG algorithm was introduced into the new combined method to generate the maximal structure. Secondly, different ideas are used to identify the candidate operating units. In Sarkozi’s method, the operating units were identified by the Means-ends analysis. It was especially convenient to synthesize azeotropic-distillation systems. The lumped materials and the candidate operating units could be initially and easily identified on the basis of the feasible inputs and outputs in the regions of the residue curve maps (RCM). But the usage of means-ends analysis cannot guarantee the generation of a feasible flowsheet, and even made the problem complicated. In our method, the new combined method was used to determine intermediate materials by hierarchically decomposing tasks and the operating unit sets \( \{o\} \) by the idea of phenomena-driven design. The new method used the full hierarchical nature of the separation process, by which different kinds of class separation units could be evaluated during process synthesis. Then the maximal structure of the separation process could be automatically generated by computer.

The main advantage of the new combined method was that the P-graph approach was used to reduce the complexity of the problem, after the materials and all the operating units were specified by hierarchical decomposition. Thus, the new combined method based on P-graph was highly efficient to generate the rigorous superstructure of larger practical separation-network synthesis problems without sacrificing optimality.
APPLICATION OF THE NEW COMBINED METHOD IN BIOPROCESSES SEPARATION NETWORKS

Example One

(1) Process Description

The case analysis was the separation process of a fermentation product – butanol (B). In the fermentation fluid, there were other valuable byproducts, ethanol (E) and acetone (A), which also need to be recovered.

This separation process had been continuously studied and improved step by step (Liu et al., 2004; Liu et al., 2006). In the early stage, the products were obtained through two process routes. One route was that the fermentation fluid (S00) became in-process product (S05) via a gas stripper (G1), and the products (B, E, A) were obtained through a series of distilling operating-units. The other was that feed stock passed through the extractor (E1) and solvent stripper (S1) and the products (B, E, A) were obtained via distillation. Later, one mature separation technology, an absorption unit, was introduced as a selection method of process reforming. The fermentation broth contains 2 wt % insoluble solids, whose densities are much higher than those of the aqueous portion of the broth comprising butanol (B), ethanol (E), acetone (A), and water (W). Hence, the insoluble solids can be readily removed from the fermentation broth by centrifugation prior to adsorption, and then the in-process product (S08) was obtained via the absorption unit. Recently, another developing technology, pervaporation, was also introduced. The new route was that suspended and undesirable substances were removed by a hyperfiltration unit and then S08 was obtained via a pervaporation unit. The evolution of this separating process perfected the process step by step.

When dealing with S08, all the posterior parts of the different routes were the same – a series of distilling operating-units, so only the early part of the separation process was studied, namely, the process synthesis of the route from feed stock (S00) to S08.

(2) Synthesis Procedure

The above procedure of process synthesis via the new combined method based on P-graph is interpreted below.

Initially, the relevant properties of raw materials and products were defined. To better understand the combined method, S08 was regarded as the desired product, as previously mentioned. After raw materials and products were determined, the relevant properties of raw materials (S00) and products (S08) are defined. Besides a lot of water and suspended solids, S00 contains B, E and A at the concentrations of 1.5, 0.2, 0.6 wt%, respectively. S08 consists of 20 wt% A, 5.7 wt% E and 74.3 wt% B. And S05 consists of 36 wt% B, 14 wt% A, 3 wt% E, and 47 wt% W (Liu, et al., 2004; Liu, et al., 2006). The problem space scope was determined by hierarchical decomposition. S00 was considered to be an initial state and S08 the goal state.

Secondly, according to the state differences between the raw materials (S00) and the desired products (S08), all potential intermediate products and the operating units were heuristic-decided step by step. Figure 3 shows the hierarchical structure of the separation process for product S08.
Differing from the products (S08), S00 contained a lot of water and few solids. And S05 consisted of less water than S08. Currently, simple distillation tended to be the first choice for liquid mixture separation. Here we chose gas stripping (G1), which is as simple as or even simpler than distillation; thus S05 and waste water S01 could be obtained. Extraction could be a viable alternative to distillation. If S00 was fed to the extractor (E1), S03 (the raffinate phase) and S16 (the extract phase) were obtained, and then S08 was obtained from the entrainer through the operating unit S1 (stripping tower). Adsorption (B1 and B2) is a major means of separating fluid mixtures. However, feed stock S00 was first centrifugated by the operating unit (C1) to obtain a clear supernatant (S51) and byproduct (S52) that prevented the adsorption of insoluble solids, including cells, onto the adsorbents’ surfaces, and facilitated the adsorption of the desired B, E, and A. Finally, S08 and byproduct S53 were obtained. The other attractive alternative in chemical or biochemical processing was pervaporation, which is a membrane separation process. S55 and byproduct S56 were obtained after the suspended solids were removed by a hyperfiltration unit (U1), and then S08 and byproduct S57 were obtained via a pervaporation unit (P1). We also noted that S08 can be obtained from S05 by the adsorption units (B3, B4), while byproduct S54 was obtained.

Finally, the superstructure was generated by the MSG algorithm. Due to the efforts of two former steps, all potential materials sets \( \{m\} \) and the operating units set \( \{o\} \) have been determined, which were the necessary conditions for P-graph theory. The superstructure, i.e., the maximal structure, was generated by the MSG algorithm with a computational time of 3 minutes. S01, S03, S56, S54, S52, S57 and S53 were regarded as by-products in the downstream processing system for biochemical production. The biggest structure of the separation process for product S08 is shown by P-graph in Figure 4.

The superstructure was the same as that developed by Fan et al. (2008). The optimal separating process and operating units could be determined at the beginning of the process synthesis; therefore, the combined method was obviously efficient for saving time.

Example Two

(1) Process Description

The production of penicillin was another bioprocess synthesis case to evaluate the feasibility of the new combined method. The penicillin production process is relatively large scale. The traditional penicillin production flowsheet is as follows (Swartz, 1985).

Production of penicillin usually begins with aerobic fermentation in batch fermenters. When the penicillin reached a concentration of around 20 to 35 g/L, the broth was then filtered using a rotary vacuum drum filter to remove the cells. The filtered broth was acidified and the penicillin was extracted into the organic phase using a solvent (e.g., butyl or amyl acetate). Next an extraction step would be used to return the penicillin to the aqueous phase. Depending on product specifications, pigments and other impurities would be removed in the carbon adsorption step. Then excess potassium or sodium acetate was added in a mixing tank to induce crystallization. Finally, the crystals were collected by a centrifuge or drum filtration unit and washed and dried to produce the final product.

Swartz (1985) has identified potential flowsheet modifications to reduce the operating costs of the traditional penicillin production process: first, membrane filtration instead of rotary drum filtration of the broth; second, whole broth processing via organic extraction (i.e., no initial filtration step); and then ion-exchange in place of two-phase extraction. Steffens (1999) gave the three best flowsheets, in which the carbon adsorption and filtration process were not included. In addition, a membrane filtration unit was used instead of rotary drum filtration.

(2) Synthesis Procedure

The synthesis procedure of the penicillin production process via the new combined method based on P-graph was as follows.

Initially, the relevant properties of raw materials and products were defined. Penicillin was regarded as the desired product, and the broth (F0) was the raw material. Besides a lot of water and the suspended solids, F0 contains penicillin (product) at the concen-
tations of 20-35 wt%. The problem space scope was therefore determined.

Secondly, according to state differences between the raw materials (F0) and the desired products (penicillin), all potential intermediate products and the operating units were decided step by step. Table 1 summaries the unit design methods and assumption used in the penicillin bioprocesses separation network synthesis (Steffens, 1999). Figure 5 shows the hierarchical structure of the separation process for penicillin.

**Table 1: unit design methods and assumptions.**

<table>
<thead>
<tr>
<th>Design method and assumptions</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fed-batch operation</td>
<td>Fermentation</td>
</tr>
<tr>
<td>Incompressible cake, assume</td>
<td>Rotary drum filtration</td>
</tr>
<tr>
<td>solids cone.</td>
<td>Membrane filtration</td>
</tr>
<tr>
<td>Small key is completely</td>
<td>Organic extraction</td>
</tr>
<tr>
<td>permeable, large key</td>
<td>Aqueous extraction</td>
</tr>
<tr>
<td>impermeable</td>
<td>Centrifugation</td>
</tr>
<tr>
<td>Solvent: butylacetate</td>
<td>Ion-exchange</td>
</tr>
<tr>
<td>Phases: PEG/Phosphate</td>
<td>Crystallization</td>
</tr>
<tr>
<td>Assume slurry concentration</td>
<td></td>
</tr>
<tr>
<td>Assume a residence time</td>
<td></td>
</tr>
<tr>
<td>Assume purity and yield</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 5:** The hierarchical structure of the separation process for penicillin.

**Figure 6:** The biggest structure of the separation process for penicillin.

**CONCLUSION**

Since separation process of biochemicals are usually complex and tedious, the design heuristics and extensive laboratory procedures are continually used during process synthesis. However, it is very difficult for biochemical process synthesis to generate the rigorous superstructure. A new combined method based on P-graph was applied to determine the necessary conditions for the MSG algorithm by the use of hierarchical decomposition and the idea of phenomena-driven design. This method greatly expanded the search space and avoided missing important operating units. The P-graph algorithm facilitates complicated process synthesis problems and guarantees the generation of superstructures without losing the optimality.

The efficacy of the new combined method based on P-graph was verified with two bioprocesses examples. It indicated that the method is robust to automatically generate a rigorous superstructure for downstream biochemical processes.

**REFERENCES**


Bertok, B., Barany, M. and Friedler, F., Generating and analyzing mathematical programming models