Combinatorial Algorithms of the S-Graph Framework for Batch Scheduling

Máté Hegyháti and Ferenc Friedler*

Department of Computer Science and Systems Technology, University of Pannonia, 10 Egyetem u., Veszprém, H-8200, Hungary

ABSTRACT: Methods for solving batch process scheduling problems have gone through a vast development in the last 2 decades. Most of the published approaches are based on a mixed integer programming formulation. Since the difficulty of scheduling is originated from its combinatorial nature, graphs and combinatorial algorithms are more adequate to represent and solve the problem. Although, combinatorial algorithms and data structures have an enormous literature, these algorithms cannot be directly applied to scheduling and further elaboration is needed. In the present work, the combinatorial nature of batch scheduling problems is analyzed. Several combinatorial algorithms are listed that can be considered for the scheduling of batch processes. Their proper adaptation is illustrated via the S-graph framework, in which the main emphasis lies on the combinatorial tools. Furthermore, Place Petri Nets and Timed Automata are also briefly described. An S-graph algorithm has been extensively compared with well-known MILP formulations.

COMBINATORIAL NATURE OF BATCH PROCESS SCHEDULING

Scheduling is a key problem in the operation of batch plants. The industry generates a wide range of batch scheduling problems, where the goal in general is to allocate the tasks of the process to the available equipment units in the most favorable way. An ordinary batch scheduling problem is given by the master recipe of the process, the objective, and the intermediate storage policy. The most common objectives are the minimization of the whole processing time, i.e., makespan, or the maximization of the throughput or profit over a fixed time horizon. According to different problems, the storage policy can vary between unlimited intermediate storage (UIS), finite intermediate storage (FIS), common intermediate storage (CIS), nonintermediate storage (NIS), and zero-wait (ZW). Problem specification may include further parameters, e.g., transfer times, changeover times, or variable processing times. The recipe defines the set of products to be produced, the network of tasks to produce the desired products, the available equipment units, processing times, stoichiometric data, etc. In the case of a complex recipe, i.e., when the process does not have sequential characteristics, the unambiguous representation of the network of the tasks is not evident. In batch process scheduling, mostly directed graphs, e.g., State-Task-Network (STN), Resource-Task-Network (RTN), State-Sequence-Network (SSN), S-graph, Timed Place Petri Net (TPPN), or Priced Timed Automata (PTA) are applied for this purpose. Despite the wide range of available graph representations, most of the approaches consider them only as a graphical representation and not as the model for the optimization.

The combinatorial nature of batch scheduling problems derives from the two main decisions to be made during the optimization process: (i) which processing unit is assigned to a task (if more than one is available) and (ii) what is the order of the tasks to be performed in an equipment unit. Moreover, if the objective is to maximize the throughput, the optimal number of batches has to be also identified, which is an additional computational issue.

COMBINATORIAL ALGORITHMS

Combinatorial algorithms that operate on graphs and sets rather than continuous variables or functions can usually address problems with discrete decisions in a straightforward way, which is often beneficial in terms of the computational efforts. This type of algorithms has an enormous literature, with an excessive introduction to the basic algorithms is given by Cormen et al. The algorithms and data structures presented in this book are usually not directly adaptable to real life problems, but they can be applied as subroutines in an algorithm that has been constructed to meet the specific requirements. Some of the combinatorial algorithms also appearing in batch process scheduling are shortly described below.

Cycle Detection Algorithm. For a given directed graph, the existence of a cycle among the vertices can be decided by the so-called colored depth-first-search algorithm. The algorithm is based on the well-known depth first search originally developed by Tarjan; however, it also assigns three different colors to each vertex, denoting whether the vertex is (a) undiscovered, (b) discovered but not finished, or (c) finished. With the appropriate data structures, the algorithm can detect a cycle or prove its nonexistence in as many steps as the sum of the number of the vertices and arcs.

Special Issue: Puigjaner Issue
Received: August 27, 2010
Accepted: March 8, 2011
Revised: February 26, 2011
Published: April 27, 2011
Longest Path Algorithm. If the arcs of a directed graph have weights, it is a common aim to find a path, in which the sum of the weights is maximal. The problem is hard to solve in its general form; however, a slightly modified version of Dijkstra’s algorithm\(^\text{14}\) can efficiently return the longest path, if the graph is known to be acyclic.

Branch-and-Bound Algorithm. When the optima is to be found in a given set, branch-and-bound algorithms\(^\text{15}\) usually reduce the computational needs compared to an exhaustive enumeration. A branch-and-bound algorithm consists of two functions.\(^\text{17}\) The branching function splits a set of candidate solutions into smaller subsets until singleton sets are achieved; thus, the solution of the original problem is decomposed to the solution of smaller subproblems. The bounding function provides a lower bound for the candidates in a certain subset, assuming that the objective is to be minimized. If the bound is worse than the best solution found so far, the corresponding subset is pruned from the set of open subproblems in the proven absence of optimal solution. The efficacy of this type of algorithms strongly depends on the selection of these functions and the selection strategy of unexamined subproblems.

Further Combinatorial Algorithms. Specific combinatorial algorithms are widely applied for scheduling problems of other areas.\(^\text{8,15}\) Moreover, there is a long list of additional combinatorial algorithms that can be applied for solving specific problems or for acceleration of the search. To mention a few of these algorithms: the so-called Hungarian method\(^\text{20}\) for the assignment problem,\(^\text{21}\) Johnson’s algorithm\(^\text{22}\) for flow-shop problems with two machines, or Jackson’s algorithm\(^\text{23}\) for job-shop problems with two machines.

**COMBINATORIAL ALGORITHMS FOR BATCH PROCESS SCHEDULING**

The idea of applying combinatorial algorithms for the scheduling of batch processes arose at a scientific discussion initiated by professor Luis Puigjaner, which resulted in the establishment of the S-graph framework appearing in 1998. Since then, several Ph.D. works further developing the framework have been prepared in both Barcelona\(^\text{24-26}\) and Veszprem.\(^\text{27,28}\)

At its birth, the S-graph framework has been presented to solve makespan minimization problems with NIS policy, fixed processing time, and batch sizes. The originality of this approach was its new, graph-based mathematical model and a problem specific solution procedure. In the following subsections, the application of the previously mentioned combinatorial algorithms for batch process scheduling is illustrated via the S-graph framework.

**Mathematical Model: The S-Graph.** The S-graph framework employs a special directed graph for the optimization, called the S-graph.\(^\text{10}\) The problem is represented by an S-graph, where a vertex is assigned to each task and product of the process. The precedence rules defined by the recipe are expressed by the so-called recipe-arcs, whose weight is the processing time of the task represented by the starting node of the arc.

The S-graph representing the recipe of an example is given in Figure 1. The process consists of two products (P1 and P2) that are produced through five processing steps (T1 – T5). Product P1 can be performed by two sequential operations T1 and T2, which have processing times 3 and 4, respectively. For the sake of simplicity, in this illustrative example, the processing times of the tasks do not depend on the selection of the plausible equipment units. The first task of the first product can be performed either in equipment unit E1 or E2. The second task has also two plausible units, namely, E2 and E3. The production of the second product (P2) has two initiative tasks, T3 and T4, which provide different inputs for task T5 that produces the final product. Task T4 is identical to task T1, and tasks T3 and T5 can be performed in 4 h by the units E3 and E2, respectively. In this example one batch is to be produced from both of the products. In the case of multiple batches, the graph corresponding to the product has to be copied to construct the S-graph representing the recipe. Note, that recipe-arcs represent timing precedences, i.e., if a recipe-arc leads from task Ti to Tj with the weight of PTij, the execution of task Tj has to start at least PTi, time units later than the execution of task Ti.

During the optimization procedure, the discrete decisions described in the first section are made and the graph is modified accordingly. If a unit is assigned to a task, the set of plausible units is replaced by that single unit. Moreover, the ordering of tasks assigned to the same unit is expressed by zero-weighted arcs, so-called schedule-arcs. In Figure 2, a partial schedule is represented by an S-graph. Unit E1 has been assigned to task T4; moreover, it has been decided that task T2 will be performed in unit E3 right after it has finished performing T3 and loaded its output to unit E2 for task T5. This latter decision is expressed by the zero-weighted schedule-arc leading from task T5 to task T2, since E3 cannot undertake task T2 until it has been emptied.
Since both the schedule and the recipe-arcs express timing precedence between the execution of the tasks, the existence of a cycle would mean an infeasible schedule. Hence, after each decision, the resultant S-graph is tested for containing a cycle with the colored depth-first-search algorithm. If the graph is not acyclic, the decision is disapproved.

If all the necessary decisions have been made, the resultant S-graph uniquely defines a schedule and the corresponding Gantt-chart. In Figure 3, the remaining decisions have been made in such a way that both tasks T1 and T4 will be performed in unit E1, starting with T4. For a scheduled S-graph, the longest path represents the critical path in the schedule, thus its value is the makespan itself. In this schedule the longest path is $T3 \rightarrow T5 \rightarrow T1 \rightarrow T2 \rightarrow P1$ with a length of 11 h.

An other possible schedule including the decisions of the partial schedule in Figure 2 is given with its Gantt-chart in Figure 4 that has the makespan of 15 h. The optimal schedule with the makespan of 8 h is given in Figure 5.

With minimization of the makespan, the optimization algorithm explores a branch-and-bound search tree, in which one S-graph corresponds to each node. The root of the tree is the S-graph that corresponds to the recipe, while the S-graphs at the leafs represent the possible schedules of the problem. The intermediate nodes correspond to partial schedules that are an extension of the S-graph corresponding to their parent node.

As it was mentioned, the efficiency of a branch-and-bound algorithm strongly depends on the selection of the branching and the bounding function and the node selection strategy. Two different types of branching functions have been published in the literature: the equipment-based method, in which the decision is made to select the next task to be performed by a certain unit; and the task-based method, in which an equipment unit is to be selected for a certain task. Both methods have their advantages, and the properties of scheduling problems determining which branching is more efficient to apply have been identified. The most simple bound that can be considered is the longest path of a partial schedule. Although it is easy to compute, it can provide considerably sharp bounds in many cases.

**Additional Applications of S-Graphs.** The previously described algorithm minimizes the makespan for a certain demand. Majozi Friedler has introduced a new concept to apply S-graphs for throughput maximization. Laínez et al. has adapted this concept to develop a new algorithm for maximizing the expected profit for stochastic scheduling problems with uncertain demand and price. Romero et al. have extended the original method to address batch scheduling problems with FIS policy that often appears in...
the industry. Adonyi et al.\textsuperscript{32} have developed an S-graph based algorithm addressing scheduling problems with cleaning costs and cleaning times that is common in paint production facilities. Adonyi et al.\textsuperscript{33} have extended the original approach to take into account streams with various temperatures. In their approach, utility cost can be minimized for a given time horizon or the makespan can be minimized with an upper bound on the utility cost. Romero et al.\textsuperscript{34} introduced a methodology based on the S-graph framework to provide optimal schedule by exploiting the inner flexibility of the batch process at the plant level.

A critical modeling issue has been recognized by Ferrer-Nadal et al.\textsuperscript{35} and Hegyh\H{a}ti et al.\textsuperscript{36} in the literature, which made MILP models result in infeasible solutions. The infeasibility arised as a situation, where a set of units must exchange their materials without available storage. It has been shown\textsuperscript{36} that the combinatorial meaning of this infeasibility is a zero-weighted cycle in the ordering of tasks, which can be detected by a cycle search as any other infeasibility when applying the S-graph framework. In Figure 6, an infeasible schedule of the former example is given with its Gantt-chart and S-graph, where the zero-weighted cycle is highlighted. At 4 h of the time horizon, units E2 and E3 try to exchange their materials, which is inapplicable in practice. This issue highlights the advantage and importance of using straightforward combinatorial tools for the modeling and optimization of the scheduling of batch processes.

\section*{AUTOMATA AND PETRI NETS FOR SCHEDULING}

Finite automata and Petri nets are widely applied in the analysis of discrete event systems.\textsuperscript{37} These mathematical models can be especially useful to explore the reachable set of states of the system to find deadlocks and other potential undesired behaviors of a system. Several extensions exist to address timing in these models.

Timed automaton were introduced by Alur and Dill\textsuperscript{38} extended to Priced Timed Automaton (PTA) by Behrmann et al.\textsuperscript{39} On the basis of PTA and the zone abstraction technique, the formulation of multiproduct batch scheduling problems was given by Panek et al.\textsuperscript{12} Ghaeli et al.\textsuperscript{11} proposed a modeling of batch scheduling problems based on Timed Place Petri nets. Both approaches apply branch-and-bound algorithms to explore the set of reachable states and find the goal state with the smallest value.

The Petri net representation of the problem given in Figure 1 is shown in Figure 7. In this Petri net, a place corresponds to the execution of a task in a unit. Moreover, a place is defined for each product and equipment unit. Transitions refer to the starting or ending of a task in a particular unit. For the sake of visibility, the arcs connected to the places representing units E2 and E3 are not shown in the figure and only the arcs related to equipment unit E1 are given. Note, that automata and Petri net based approaches also avoid cross transfer, which appears in these models as a deadlock.

\section*{COMPUTATIONAL COMPARISON}

To illustrate the expediency of combinatorial algorithms in batch process scheduling, an exhaustive comparison has been made for a multiproduct scheduling problem,\textsuperscript{40} where 4 products, A, B, C, and D are to be produced in three consecutive steps. The recipe with the five available equipment units (U1 – U5) and processing times is shown in Figure 8. The objective is to minimize the makespan for the given number of batches. In this study, 28 different cases were examined: starting from 1 batch up to 7 batches from each

---

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Infeasible schedule represented by an S-graph and its Gantt-chart.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure7.png}
\caption{Petri net representation of the problem given in Figure 1. (Note that the arcs connected to places E2 and E3 are not shown for the sake of visibility.)}
\end{figure}
product. Each case was solved by four different approaches: a global time point based formulation by Maravelias and Grossmann, a unit specific time point based formulation by Shaik et al., a precedence based formulation by Mendez and Cerda, and the S-graph based algorithm by Sanmarti et al.

Each test was performed on the same computer with an AMD ATHLON 1.8 GHz processor and 1.5 GB of memory. For the S-graph based approach, a recent implementation by Smidla and Heckl was applied. The MILP models were solved by CPLEX version 11.1.1. The time limit for the optimization was 1000 CPU seconds if an approach failed to finish in this time limit and the current best solution was reported. In the case of time point based formulations, the iteration on time points was performed until the second occurrence of the same objective value. For the three-index formulation, the $\Delta n$ parameter was considered to be 0.

The results of the comparison are given in Table 1. The S-graph based solver and the precedence based formulation could provide the best results. The more general time point based formulations reached the 1000 s limit at much smaller batch numbers.

### CONCLUDING REMARKS

The combinatorial nature of batch process scheduling has been investigated. Problem representation and modeling that fits the scheduling is crucial both in guaranteeing the optimal solution and reducing the computational effort. S-graph framework initiated by Puigjaner and Friedler proved to be appropriate for problem representation and for involving combinatorial algorithms in the solution procedure. Several additional combinatorial algorithms have been presented here as an illustration.

Because of the industrial requirements, scheduling problems are getting more and more complicated; therefore, the efficacy of the solution procedures is crucial. The present work illustrated that combinatorial algorithms and tools play important role in adequate modeling and in the effective solution of batch scheduling problems. Combinatorial algorithms can further be exploited to solve industrial problems with additional requirements, e.g., simultaneous heat integration and scheduling. It is assumed that S-graph framework provides an appropriate base for that.

![Figure 8. Recipe of the example for the comparison study.](image)

Table 1. Comparison Results for the Example

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obj.</td>
<td>CPU time</td>
<td>Obj.</td>
<td>CPU time</td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>25 h</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>31 h</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>37 h</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>39 h</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>41 h</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>46 h</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>52 h</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>55 h</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>57 h</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>62 h</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>67 h</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>71 h</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>73 h</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>78 h</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>82 h</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>87 h</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>89 h</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>94 h</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>98 h</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>103 h</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>$\leq 105 h$</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>$\leq 110 h$</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>$\leq 114 h$</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>$\leq 119 h$</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>$\leq 123 h$</td>
</tr>
</tbody>
</table>
AUTHOR INFORMATION

Corresponding Author
*E-mail: friedler@dcs.uni-pannon.hu. Phone: +3688424483. Fax: +3688428275.

REFERENCES


