Genetic algorithms applied to scheduling and optimization of refinery operations

F. Oliveira\(^1\), M.R. Almeida\(^2\) and S. Hamacher\(^1\)

\(^1\)Department of Industrial Engineering, Pontifícia Universidade Católica , Rio de Janeiro, Brazil
\(^2\)PETROBRAS, Petróleo Brasileiro S/A, Brazil

Keywords: Genetic Algorithm, Scheduling, Petroleum Refinery, Multi-objective.

Abstract. This paper presents a Genetic Algorithm-based method to optimize the production schedule of the fuel oil and asphalt section in a petroleum refinery. Two Genetic Algorithm models were developed to establish the sequence and size of all production shares. A special mutation operator was also proposed to minimize the number of changes in the production. A multi-objective fitness evaluation technique was also incorporated to the Genetic Algorithm models. The obtained results confirm that the proposed Genetic Algorithm models, associated with the multi-objective energy minimization method, are able to solve the scheduling problem, optimizing the refinery’s operational objectives.

1 Introduction

The emergence of international markets and the development of global competition were the principal aspects that led the chemical processing industry towards reaching more efficient plant operations, with concomitant better economical performance (Pinto et al., 2000).

A key issue in the refinery context is its operations scheduling. Initially, scheduling research has focused on methods for obtaining optimal solutions to simplified problems using integer programming and "branch & bound" approaches. Blazewicz (1992) describes some operational research methods for solving scheduling problems. Göthe-Lundgren et al. (2002) show how mixed integer linear programming (MILP) can be applied to oil refinery’s scheduling. Pinto et al. (2000) also present a MILP approach, showing the application of some linearization techniques to deal with product quality constraints, which are strongly nonlinear. More recently, Karuppiah et al. (2007) have proposed an algorithm that focuses on effectively solving a mixed-integer linear programming using nonlinear relaxations.

Nevertheless, many authors have pointed out some limitations of using linear mathematical programming to optimize refinery productivity, especially regarding operations scheduling (Moro and Pinto, 2004; Ballintijn, 1993; Potts and Van Wassenhove, 1992).

Because of these limitations, several authors highlight the importance of taking other approaches using heuristic methods. Martin (2009) argues that methods such as Genetic Algorithms (GA) are more appropriate to deal with highly complex problems in an attempt to find reasonably good solutions in reasonable amounts of time. He and Hui (2007a) show that MILP is not suitable to solve large single-stage multi-product scheduling problems and suggest that such problems must be heuristically approached.

GAs became a subject of interest after the publication of Holland’s work (Holland, 1975). Nowadays, interesting articles can be found in the literature dealing with applications of GAs to scheduling problems. More recently, Simão et al. (2007) have proposed a parallel approach using cooperative coevolution in order to solve scheduling problems in refineries. He and Hui (2007b) have proposed a GA for solving large multi-stage batch plant scheduling with a penalty method for handling the constraints in the problem. The authors of both articles point out the advantages of GAs, especially due to their flexibility concerning real problem representation.
This article undertakes this goal. It shows the development of a methodology that applies GA to solve the scheduling problem of fuel oils and asphalt. This article also presents a new genetic operator specially developed for this problem, as well as briefly presenting a novel method to optimize multi-objective problems.

2 Description of the problem

The problem considered in this study is the optimization of scheduling of fuel oil and asphalts at the REVAP refinery, located in the State of São Paulo, Southeast Brazil. REVAPs total capacity is roughly around 1,000,000 $m^3$/month, and the plant produces 180,000 $m^3$/month of fuel oils and 40,000 $m^3$/month of asphalt. The refinery receives crude oil and intermediate feedstock from pipelines, and dispatches most of its finished products via pipelines as well. Despite this problem was solved by Joly et al. (2002) and Pinto et al. (2000) using mathematical programming, some important characteristics of the real problem were not take into consideration in such approach. Hereby we propose a new model that embraces these characteristics and therefore its representation is closer to REVAPs production planning problem.

The fuel oils and asphalt division of REVAP is a multi-product plant with two machine stages - a mixer and a set of 21 storage tanks - which operates continuously, without setup time and with resource constraints. During the horizon of production scheduling, asphaltic residue (RASF) is produced continuously by the deasphalting unit. The RASF is diluted with decanted oil and/or light cycling oil to produce four types of fuel oils and two types of ultraviscous fuel oils, and diluted with gas oil in order to produce asphalt.

Several constraints are inherent to this problem: (i) no tank can ever be loaded and unloaded simultaneously; (ii) production must be continuous, because RASF never stops feeding the mixer; (iii) demand must be supplied based on a previously set schedule, i.e., it is not the scheduler’s task to determine the best schedule to supply the demand; (iv) there is a minimum volume that must remain in finished and intermediate product tanks (due to the particular characteristics of the floating-roof tanks); (v) there is a preparation and a quality inspection time for both intermediate and final products.

The central objective here is to find a feasible solution that can: (i) fulfills the demand; (ii) minimizes the number of operational changes (i.e., maximize lot sizes); (iii) minimizes average inventory levels; (iv) minimizes give-away (i.e., minimize the occurrence of unnecessarily high specifications of the products, which means avoiding the production of high-valued products in order to supply low-spec product demand). The following section describes the approach taken to reach such objective.

3 The genetic algorithm

In this study, the proposed GA aims at minimizing the demand that cannot be supplied, minimizing the production that cannot be allocated to the tanks, and minimizing the number of operational changes. Indirectly, by conducting experiments with different inventory strategies, this study intends also to minimize the average product volumes in stock.

The most important aspects concerning genetic algorithms are the manner in which the solution of a problem is represented by a chromosome and which operators must be applied in order to assure an appropriate solution. Classically, one can find in the literature two distinct approaches: direct and indirect representation.

In the direct approach, a complete solution is encoded into a chromosome and the GAs are responsible for evolving this solution into a better one.
On the other hand, in the indirect representation approach a sequence of dispatching rules that assign the job are encoded into a chromosome and the GAs are used to evolve them in order to reach a better sequence of dispatching rules.

For the problem considered here, we propose two GA models. The first uses a direct representation of the production schedule, dividing the scheduling horizon into discrete intervals of one hour. The second model uses an indirect representation which must be decoded into a production scheduling. The following subsections present some aspects of both approaches as well as the genetic operators used with them and the methodology to evaluate multi-objective solutions.

3.1 The direct approach

The representation adopted here follows what was done by Joly (1999), Sikora (1996) and Lee et al. (1997). The weekly production schedule is divided into intervals of one hour, which means that the suggested chromosome always carries 168 genes. The chromosome is built by adding the demand for each product and it always tries to produce the maximum of every product at each time interval, respecting resource constraints and the limitations of the mixer. Then, a chromosome with two-element list genes is obtained; the first element represents the product and the second one its production volume for each one-hour time interval.

3.2 The indirect approach

In this approach, each gene of the chromosome represents a demand for a certain product existing in the scheduling horizon. The aim here is to reach a solution in which a job sequence can be a feasible (and hopefully good) solution.

After obtaining the task sequence proposed by the GA, each chromosome must be decoded according to the necessary time to prepare the product for the mixing process; the time (hours) it takes to produce the volume demanded at the maximum production rate (within quality restrictions and available raw material volume); when the product must be ready (taking into account preparation and quality inspection time); and when the product must start to be delivered. These values represent the ideal time in which each product should be produced to fulfill the demand, without considering any inventory that might exist.

In several occasions the decoding process cannot allocate the production of a specific product or even fulfill an order on time, due to the problem’s constraints. To deal with these issues we have proposed a set of rules to allocate the production, dividing the lots whenever necessary.

3.3 The genetic operators

Three genetic operators are used in this study. Two of them already exist in the literature, namely uniform order-based crossover (Davis et al., 1991) and mutation by shuffling (Davis et al., 1991). Besides these two operators, a new mutation operator type was developed, which has been called neighborhood mutation (NM).

NM was developed aiming to cope with the minimization of operational mode changes. The steps to accomplish such mutation are: (i) choose a gene randomly; (ii) search for an equal gene to the left and to the right; (iii) keep searching until a different gene is found and replace it with the one found in (ii). By approximating lots of identical products, the NM manages larger lots, thus minimizing the number of operational changes.
3.4 Evaluation of solutions and the multi-objective optimization technique

The production scheduling evaluation has been accomplished by means of a rule-based simulation system, which emulates the fuel oil and asphalt division behavior. Such evaluation involves:

1. Demand fulfillment: every time it becomes impossible to cope with the customer’s need, the solution (i.e., the scheduling) receives a penalty, which is proportional to the volume that was not delivered to the customer;
2. Operational changes: the number of mode changes is also penalized;
3. Production allocation: the allocation can be exposed to two different problems: the first happens when the tank is completely full and the production cannot be allocated into it; the second happens when it is impossible to load and unload the tanks simultaneously.

In order to optimize the achievement of such objectives, the Energy Minimization (EM) methodology (Zebulum et al., 2000) was used. This method solves the main difficulty of scalar aggregation techniques, namely to choose the most adequate weights associated with each objective. EM, which was designed to be used within a GA, is capable to adaptively update the weights throughout the evolutionary process. Therefore, higher priorities are constantly shifted between the objectives that are less satisfied among the solution population. In addition, the EM method incorporates user specifications, which is not trivially done using techniques that seek the Pareto-optimal set.

The main idea here is to minimize the “energy” of the system, which consists straightforwardly of the weighted sum of each objective-function, while the weights are dynamically updated during the evolutionary process.

4 Results

To analyze the experimental results obtained from the GA, real data were used, consisting of a demand scenario with 165.32, which represents 98.4 percent of production of 7 days with 24 hours each at maximum operation capacity. In other words, this scenario illustrates a problem that is hard to schedule manually. The three types of representations and mutations presented in Sections 3.1 and 3.2 - namely Direct Representation with Neighborhood Mutation (NM), Direct Representation without Neighborhood Mutation, and Indirect Representation - have been tested. The results obtained are shown in Figure 1, representing the average results of 6 GA simulations. As one can see in figure 1, the NM operator has improved the results of direct representation, but this kind of representation is far from attaining the problem’s goals, especially the operational change minimization. Indirect representation has better performance, achieving better results since the first generation. However, the GA has not fully accomplished its objective, since there still are some unfulfilled demands. This is directly correlated with the inventory levels adopted by the company. We designed and conducted experiments in order to define finished product tank level scenarios aiming to find the best inventory level. To meet such goal, we have tested 4 scenarios with different finished-product starting level in the tanks.

Table 1 summarizes the average results obtained in 6 simulations using indirect representations. The goal of the first two objectives is to reach zero, which means that all of the production should be allocated to the tanks and all of the demand should be supplied. The third objective (operational mode changes) aims at attaining the lowest possible number. As can be seen, the best result has been obtained in Scenario II. Using the indirect approach, all simulations led to solutions that supply the demand without exceeding the capacity of the tanks as well as keeping the stocks of the finished products at low levels.
This paper has proposed two genetic algorithm models to solve the optimization of a lotsizing and sequencing problem in a multi-product plant with two-stage serial machines. Indirect representation has achieved outstanding performance levels concerning demand fulfillment and production allocation; a satisfactory performance level (according to the refinery’s real production scheduling) was observed when it comes to operational mode change minimization. As can be seen in Figure 1, direct representation achieves some interesting results, which are further improved with the use of the NM operator. Nevertheless, when it comes to the number of operational mode changes, using such production schedule becomes impossible due to the high number of changes.

Some comparisons can be made with the MILP proposed by Pinto et al. (2000). In the GA approach, a wider time horizon can be represented. Also, the use of a time slice procedure is not necessary using the GA, which can actually compromise the solution’s quality. Another aspect that must be pointed out is the amount of operational change considered. The MILP approach often obtains solutions that are not feasible in the refinery’s reality with respect to operational changes. The GA approach presented here takes such changes into consideration, penalizing them in order to converge to more stable solutions (with bigger lots).

Another conclusion that can be pointed out concerns the starting tank levels. Both high and low levels directly influence the performance of the GA, as they might turn the scheduling infeasible. Using the right experimentation methodology, we were able to define a better tank level policy.

Another contribution concerns the implementation of the modified EM method. The EM
allowed the GA to find better solutions due to weight updates during the evolutionary process avoiding meeting some specific objectives while neglecting others.

References


