A look-ahead addition to the ant colony optimization metaheuristic and its application to an industrial scheduling problem

Caroline Gagné * Marc Gravel * Wilson L. Price †

* Département d’Informatique et de Mathématique, Université du Québec à Chicoutimi
555, Boul. de l’Université, Chicoutimi, Québec, Canada, G7H 2B1
Email: caroline.gagne@uqac.uquebec.ca, marc.gravel@uqac.uquebec.ca

† Faculté des Sciences de l’Administration, Université Laval
Ste-Foy, Québec, Canada, G1K 7P4
Email: wilson.price@fssa.ulaval.ca

1 Introduction

In many industrial situations, it is natural to turn towards the use of metaheuristics, such as the Ant Colony Optimization (ACO) algorithm, which have been shown to offer successful solution strategies for large and complex problems. Exact optimization algorithms require overlong solutions times and cannot produce an acceptable or even feasible solution in the time available. Further, it may be awkward to represent some constraints or objective function characteristics in the algebraic form required by classical optimization methods. In their review of solution techniques that have been used for scheduling flexible shops, Blazewicz et al. [2] note that methods such as simulated annealing, tabu search and genetic algorithms are frequently used and have been shown to be powerful techniques for this task. Elsewhere in the scheduling literature, we find the use of ant colony optimization ([1], [4], [5], [13], [15], [17], [20]).

In this paper we explore the addition of a look-ahead mechanism to the ACO algorithm. In constructing a candidate solution, ants use a transition rule that incorporates complete information on past decisions (the trail) and local information (visibility) on the immediate decision that is to be made. The look-ahead mechanism allows the incorporation of information on the anticipated decisions that are beyond the immediate choice horizon.

The industrial situation addresses the scheduling of a single machine for a known order book where setup times are sequence dependent. Superficially, this problem resembles the asymmetric traveling salesman problem, however there are a number of important differences. The scheduling problem must be solved with respect to multiple objectives which typically are related to throughput, to service level, and to logistics. In addition, a number of technological constraints related to the specific characteristics of the production machinery must be taken into account. In particular, the availability of specialized tooling and placement of downtime for required maintenance are important. Unfortunately, these technological constraints also increase the time needed for the function evaluation for a candidate solution.
2 Problem description

The scheduling problem is drawn from an Alcan aluminium foundry located in Québec. In this foundry, holding furnaces are charged with molten metal from a transfer crucible coming from the refiners. These furnaces serve to keep the aluminium in flux while the various ingredients are added and they feed liquid metal to the casting rig. A customer’s order has specific characteristics, such as the alloy type, the number of pieces to be produced, the dimensions of these pieces and the delivery date. The alloy specification is produced by adding the required ingredients and grain refiners to the molten aluminium in the holding furnaces. A change in the alloy being produced may require a draining and cleaning to prevent contamination of the alloy to be cast next. Changes in the dimensions of ingots to be produced requires installation of a mold having the crosssection of the desired ingots. These operations constitute the setup times incurred for the production of a customer’s order and are dependent on the scheduled sequence.

Should the horizontal casting machine lack molten metal for any reason, it will require a costly shut-down and cleaning even if ingots of the same characteristics are to be produced when operations resume. While the casting center is amply supplied with pure molten aluminum by several nearby electrolysis plants, care must be taken to manage the supply of metal in the holding furnaces to ensure that the casting process is not interrupted. This requirement constrains the problem and eliminates certain otherwise interesting candidate solutions.

Another technological constraint must take into account the use of previously refined solid metal which will slow the holding furnace preparation time. The scheduled sequence must seek to avoid stoppages on the casting rig by ensuring that, while one holding furnace supplies metal for the pour, the second is prepared and loaded.

Specialized tooling called a basin joins molds to the casting rig. Molds may be attached only to specific basins and these are in limited supply. The one required may be undergoing cleaning from a previous usage. For some dimensions, only one example of a basin is available and so some otherwise feasible sequences must be eliminated because successive pours using different molds require the same basin.

A feasible sequence of orders is one that ensures that sufficient pure metal is available for all pours, that basins and molds are available when required for each order and that draining and cleaning of the rig is done when required. A desirable feasible sequence takes into account the objectives of customer service and efficiency.

We model the objectives of the scheduler by treating the minimization of unweighted total tardiness for all orders, the minimization of unused production capacity over the planning horizon and we include a penalty function encouraging efficient transportation of the product. This function favors sequences where orders for the same destination are consecutive and penalizes sequences where this is not the case.

3 Ant colony optimization

The ant colony optimization metaheuristic ([3], [8]) was inspired by studies of the behavior of ants ([6], [7], [14]). Ants communicate among themselves through pheromone, a substance they deposit on the ground in variable amounts as they move about. It has been observed that the more ants use a particular path, the more pheromone is deposited on that path and the more it becomes attractive to other ants seeking food. If an obstacle is suddenly placed on an established path leading to a food source, ants will initially go right or left in a seemingly random manner, but those choosing the side that is in fact shorter will reach the food more quickly and will make the return journey more often. The pheromone on the shorter path will therefore be more strongly reinforced and will eventually become
the preferred route for the stream of ants. The works of Colorni et al. [3], Dorigo & Di Caro [9], Dorigo & Gambardella [10], Dorigo et al. [11], Dorigo et al. [12] offer detailed information on the workings of the algorithm and the choice of the various parameters.

4 An augmented ACO for industrial scheduling

We have proposed in [13] a number of additions to the basic ACO to take into account the characteristics of this industrial scheduling problem. These characteristics include:

- the insertion of multiple distance matrices to represent the multiple objectives;
- the incorporation of a look-ahead feature in an attempt to improve the performance of the transition rule;
- the incorporation of a local improvement method to better candidate solutions found by the ACO.

We believe the first two of these additions to be unique to our augmented ACO while the last has been tested by a number of authors.

4.1 Multiple distance matrices

Three objectives have been identified in the industrial problem. The first is to improve throughput in the plant by minimizing total setup time. The second is to minimize the total tardiness for the current order book. The third is to minimize lost transportation capacity for the finished products. Each of these objectives allows the definition of a characteristic distance matrix analogous to that of a TSP, and each distance matrix is used to define terms in the transition rule. Compared to that of the usual ACO, the transition rule has a number of new terms as shown by equations 1 and 2. In these expressions, the first term refers to the trail, the second refers to the setup penalty matrix $S$, the third to the slack-time matrix $M$ which is used in the tardiness minimization, the fourth is a penalty matrix $C$ related to loss of transportation capacity. We have found that the use of three separate matrices allows better adjustment of the exponent parameters ($\beta$, $\delta$, $\lambda$) and leads to increased control of the algorithm. We are able to produce solutions of better quality as compared to single distance matrix approaches.

4.2 Look-ahead feature

Usually, in the transition rule, the selection of the next order to place in the sequence is the result of a compromise between the trail, information on past choices, and the visibility, information derived from the problem parameters. We propose the addition of look-ahead information about the potential of the current partial solution. The idea that we propose differs in a number of respects from that proposed by Michel & Midderdorf [18] in a different problem context. These authors proposed a look-ahead function which works by evaluating the trail intensity resulting from a choice.

The look-ahead information that we propose to include estimates the potential quality of a partial solution by combining a surrogate function value, $T_{Q_h}$, for the current partial solution, $Q_h$, to which is added one of the candidate choices and a lower bound, $T_{Q'_h}$, on the remaining uncompleted portion, $Q'_h$. Note that computing the actual function value each time would be prohibitively costly. The computation of the lower bound for the uncompleted portion follows the method used in the branch and bound computations of Ragatz [19], and later in Tan et al. [21].
For example, the lower bound on the total tardiness of the partial solution, $B_{ij}$, including the order $j$ presumed to be the candidate chosen follow order $i$, is calculated according to the following:

$$B_{ij} = T_{Q_h} + T_{Q'_h}$$

The remaining two objectives are bounded in a similar fashion.

In order to represent the look-ahead information, equations 1 and 2 related to the transition rule include a fifth term, $B_{ij}$, and its accompanying exponent parameter $\phi$. In the construction of a solution, the next order $j$ is selected according to:

$$j = \begin{cases} \arg \max_{l/ \in \text{Tabou}_k} \left\{ [\tau_{il}(t)]^\alpha \cdot \left[ \frac{1}{s_{il}} \right]^\beta \cdot \left[ \frac{1}{m_{il}} \right]^\delta \cdot \left[ \frac{1}{c_{il}} \right]^\lambda \cdot \left[ \frac{1}{B_{il}} \right]^\phi \right\} & \text{if } q \leq q_0 \\ J & \text{if } q > q_0 \end{cases}$$ (1)

Where $J$ is chosen according to the probability:

$$p^k_{ij}(t) = \frac{[\tau_{ij}(t)]^\alpha \cdot \left[ \frac{1}{s_{ij}} \right]^\beta \cdot \left[ \frac{1}{m_{ij}} \right]^\delta \cdot \left[ \frac{1}{c_{ij}} \right]^\lambda \cdot \left[ \frac{1}{B_{ij}} \right]^\phi}{\sum_{l/ \in \text{Tabou}_k} [\tau_{il}(t)]^\alpha \cdot \left[ \frac{1}{s_{il}} \right]^\beta \cdot \left[ \frac{1}{m_{il}} \right]^\delta \cdot \left[ \frac{1}{c_{il}} \right]^\lambda \cdot \left[ \frac{1}{B_{il}} \right]^\phi}$$ (2)

### 4.3 Local improvement methods

A local improvement method suggested by Johnson & McGeoch [16] has been adapted for use in the augmented ACO. Called the restricted 3-opt method by these authors, it involves successive arc-exchanges in an attempt to improve a candidate solution. In the context of the industrial scheduling problem that we treat, it has been necessary to choose a limited number of arc-exchanges from among those available in order to avoid over-long computation times.

### 5 Results and discussion

The authors will present a series of numerical experiments in which allow us to distinguish the contributions of each of the elements of the augmented ACO for industrial scheduling.

In particular, we will discuss the results of the application of the look-ahead information to a set of simple test cases drawn from the literature. We will subsequently present cases closely based on an actual order book drawn from the industrial situation we have described. These latter cases require the inclusion of the complete set of technological constraints and the use of the surrogate evaluation function previously referred to. Our current results indicate that the look-ahead information improves solution quality at some cost in computation time.

Progress so far encourages us to outline avenues for the inclusion of look-ahead information in future versions of our industrial applications.

### References


