META-HEURISTICS FOR DYNAMIC LOT SIZING:
A REVIEW AND COMPARISON OF SOLUTION APPROACHES

RAF JANS & ZEGER DEGRAEVE
Abstract: Proofs from complexity theory as well as computational experiments indicate that most lot sizing problems are hard to solve. Because these problems are so difficult, various solution techniques have been proposed to solve them. In the past decade, meta-heuristics such as tabu search, genetic algorithms and simulated annealing, have become popular and efficient tools for solving hard combinational optimization problems. We review the various meta-heuristics that have been specifically developed to solve lot sizing problems, discussing their main components such as representation, evaluation neighborhood definition and genetic operators. Further, we briefly review other solution approaches, such as dynamic programming, cutting planes, Dantzig-Wolfe decomposition, Lagrange relaxation and dedicated heuristics. This allows us to compare these techniques. Understanding their respective advantages and disadvantages gives insight into how we can integrate elements from several solution approaches into more powerful hybrid algorithms. Finally, we discuss general guidelines for computational experiments and illustrate these with several examples.

### Bibliographic Data and Classifications

<table>
<thead>
<tr>
<th>Library of Congress Classification (LCC)</th>
<th>5001-6182</th>
<th>Business</th>
</tr>
</thead>
<tbody>
<tr>
<td>5201-5982</td>
<td>Business Science</td>
<td></td>
</tr>
<tr>
<td>HB 143.7</td>
<td>Optimization Techniques</td>
<td></td>
</tr>
<tr>
<td>Journal of Economic Literature (JEL)</td>
<td>M</td>
<td>Business Administration and Business Economics</td>
</tr>
<tr>
<td></td>
<td>M 11</td>
<td>Production Management</td>
</tr>
<tr>
<td></td>
<td>R 4</td>
<td>Transportation Systems</td>
</tr>
<tr>
<td></td>
<td>C 61</td>
<td>Optimization Techniques</td>
</tr>
<tr>
<td>European Business Schools Library Group (EBSLG)</td>
<td>85 A</td>
<td>Business General</td>
</tr>
<tr>
<td></td>
<td>260 K</td>
<td>Logistics</td>
</tr>
<tr>
<td></td>
<td>240 B</td>
<td>Information Systems Management</td>
</tr>
<tr>
<td></td>
<td>255 G</td>
<td>Management science, operations research</td>
</tr>
<tr>
<td>Gemeenschappelijke Onderwerpsontsluiting (GOO)</td>
<td>85.00</td>
<td>Bedrijfskunde, Organisatiekunde: algemeen</td>
</tr>
<tr>
<td></td>
<td>85.34</td>
<td>Logistiek management</td>
</tr>
<tr>
<td></td>
<td>85.20</td>
<td>Bestuurlijke informatie, informatieverzorging</td>
</tr>
<tr>
<td></td>
<td>85.03</td>
<td>Methoden en Technieken, Operations Research</td>
</tr>
<tr>
<td>Keywords GOO</td>
<td>Bedrijfskunde / Bedrijfseconomie</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bedrijfsprocessen, logistiek, management informatiesystemen</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Operations research, series, optimalisatie, methoden</td>
<td></td>
</tr>
<tr>
<td>Free keywords</td>
<td>Dynamic lotsizing, algorithms, meta-heuristics, Dantzig-Wolfe decomposition, reformulations.</td>
<td></td>
</tr>
</tbody>
</table>
META-HEURISTICS FOR DYNAMIC LOT SIZING:
A REVIEW AND COMPARISON OF SOLUTION APPROACHES

Raf Jans
Rotterdam School of Management, Erasmus University
PO Box 1738, 3000 DR Rotterdam, The Netherlands
rjans@fbk.eur.nl

Zeger Degraeve
London Business School
Regent’s Park, London NW1 4SA, U.K.
zdegraeve@london.edu

April 19, 2004

Abstract

Proofs from complexity theory as well as computational experiments indicate that most lot sizing problems are hard to solve. Because these problems are so difficult, various solution techniques have been proposed to solve them. In the past decade, meta-heuristics such as tabu search, genetic algorithms and simulated annealing, have become popular and efficient tools for solving hard combinatorial optimization problems. We review the various meta-heuristics that have been specifically developed to solve lot sizing problems, discussing their main components such as representation, evaluation, neighborhood definition and genetic operators. Further, we briefly review other solution approaches, such as dynamic programming, cutting planes, Dantzig-Wolfe decomposition, Lagrange relaxation and dedicated heuristics. This allows us to compare these techniques. Understanding their respective advantages and disadvantages gives insight into how we can integrate elements from several solution approaches into more powerful hybrid algorithms. Finally, we discuss general guidelines for computational experiments and illustrate these with several examples.
1. Introduction

There is a wide variety of models for production planning and inventory management. Lot sizing models determine the optimal timing and level of production. At one end of the spectrum there are the continuous time scale, constant demand and infinite time horizon lot sizing problems. In this category we find the famous Economic Order Quantity model (EOQ) and the Economic Lot Scheduling Problem (ELSP). At the other end of this spectrum we have the discrete time scale, dynamic demand and finite time horizon lot sizing models. This type of planning is generally referred to as dynamic lot sizing and is the subject of this review.

A number of studies have provided a good introduction into the lot sizing literature. De Bodt et al. (1984) and Bahl et al. (1987) present some earlier reviews. Kuik et al. (1994) discuss the impact of lot sizing and production planning at different decision levels in the organization and respond to some criticisms on lot sizing. Wolsey (1995) reviews the history of the single item lot sizing problem. The integration of lot sizing and scheduling, both for discrete and continuous time models, is discussed in Drexel and Kimms (1997). Belvaux and Wolsey (2001) explain how to model many extensions of the basic lot sizing problem to obtain good formulations. Although these reviews have made important contributions to the literature, they do not provide an up-to-date overview of the existing algorithms in the much evolving field of lot sizing.

This paper has several contributions. First, it fills a gap by providing a comprehensive overview of the use of meta-heuristics for solving lot sizing problems (Section 3.1). To the best of our knowledge, no general analysis of meta-heuristics has been done specifically for lot sizing problem. Second, this paper aims to provide a general review on lot sizing algorithms (Section 3.2-3.6), whereas most tutorials or literature reviews focus mainly on one single approach or model. By discussing the many different solution techniques, we obtain insights in their relative strengths and weaknesses (Section 4) and the power of hybrid algorithms. This discussion is also of interest for researchers who apply these techniques to other problems. Third, we discuss guidelines for computational experiments, with a focus on meta-heuristics (Section 5). Many meta-heuristics are specifically developed for special cases of standard lot sizing problems. Therefore, we first briefly introduce the most important problems (Section 2).
2. Lot Sizing Models

2.1. The single item uncapacitated lot sizing problem

The *dynamic lot sizing problem* in its simplest form tries to find a production plan that satisfies deterministic demand and minimizes the total costs of production, set up and inventory for one item over a multi-period horizon. We can mathematically formulate this problem as follows:

\[
\begin{align*}
    \text{Min} & \quad \sum_{t=1}^{m} \left( vc_t x_t + sc_t y_t + hc_t s_t \right) \\
\text{s.t.} & \quad s_{t-1} + x_t = d_t + s_t \quad \forall t \in T \\
& \quad x_t \leq \sum_{k=t}^{m} d_k y_t \quad \forall t \in T \\
& \quad x_t, s_t \geq 0; y_t \in \{0,1\} \quad \forall t \in T
\end{align*}
\]

In this model, \( x_t \) represents the production variable for period \( t \), \( y_t \) is the set up variable and \( s_t \) is the inventory variable. There are costs associated with these decision variables: \( vc_t \), \( sc_t \) and \( hc_t \) are respectively the variable production cost, set up cost and holding cost in period \( t \). \( T \) is the set of all periods in the planning horizon and \( m \) denotes the last period. Demand for each period, \( d_t \), is given. The objective function (1) minimizes the total cost of production, set up and inventory. Constraint (2) is the demand balance equation: the inventory from the previous period and the production from the current period can be used to meet demand in this period or to build up inventory to satisfy later demand. In each period with non-zero production a set up is incurred. This is modeled in the fixed charge constraint (3). Since no ending inventory is allowed, production is limited by the remaining cumulative demand. This problem is usually referred to as the *single item uncapacitated lot sizing problem* (ULS) and was first discussed in the seminal paper by Wagner and Whitin (1958). Zangwill (1969) demonstrated that this problem is actually a fixed charge network problem.

2.2. The Capacitated Lot Sizing Problem (CLSP)

The capacitated multi-item lot sizing problem (CLSP) is a typical example of a *large bucket* model, where many different items can be produced on the same machine in one time period. There are \( n \) different items that can be produced and \( P \) is the set of all these items. They all make use of the same machine with a limited production capacity \( cap_c \). Producing one unit of
product $i$ consumes $vt_i$ units of capacity. Product specific variables and parameters now have an extra index $i$ to identify the item. The formulation is as follows:

$$\text{Min } \sum \sum \left( sc_{ai}y_{ai} + vc_{ai}x_{ai} + hc_{ai}s_{ai}\right)$$

(5)

s.t.  
\[ s_{i,t-1} + x_{ai} = d_{ai} + s_{ai} \quad \forall i \in P, \forall t \in T \]  
(6)

\[ x_{ai} \leq \min \left\{ \frac{cap_i}{vt_i}, \sum_{k=1}^{m} d_{ik} \right\} y_{ai} \quad \forall i \in P, \forall t \in T \]  
(7)

\[ \sum_{ai \in P} vt_i x_{ai} \leq cap_t \quad \forall t \in T \]  
(8)

\[ x_{ai}, s_{ai} \geq 0; y_{ai} \in \{0,1\} \quad \forall i \in P, \forall t \in T \]  
(9)

Each product consumes a specific variable production time, $vt_i$, and all the products are made on the same machine. In the set up constraint (7), production is now limited by both the capacity and remaining demand. The total production in each period is limited by the capacity, $cap_t$ (8). The problem can also be extended with set up times (Trigeiro et al. 1989).

2.3. Small bucket models

In small bucket models, the assumption is made that during each time period, at most one type of item can be produced on the same machine. This is the case for the Continuous Set Up Lot Sizing Problem (CSLP):

$$\text{Min } \sum \sum \left( g_{ai}z_{ai} + sc_{ai}y_{ai} + vc_{ai}x_{ai} + hc_{ai}s_{ai} \right)$$

(10)

s.t.  
\[ s_{i,t-1} + x_{ai} = d_{ai} + s_{ai} \quad \forall i \in P, \forall t \in T \]  
(11)

\[ \sum_{i=1}^{n} y_{ai} \leq 1 \quad \forall t \in T \]  
(12)

\[ vt_i x_{ai} \leq cap_t y_{ai} \quad \forall i \in P, \forall t \in T \]  
(13)

\[ z_{ai} \geq y_{ai} - y_{i,t-1} \quad \forall i \in P, \forall t \in T \]  
(14)

\[ x_{ai}, s_{ai} \geq 0; y_{ai}, z_{ai} \in \{0,1\} \quad \forall i \in P, \forall t \in T \]  
(15)

The new variable $z_{ai}$ is the start up variable and there is an associated start up cost of $g_{ai}$. A start up occurs when the machine is set up for an item for which it was not set up in the previous period. The objective function (10) now minimizes the total cost of start ups, set ups, variable production and inventory. We still have the regular demand constraints (11). Further, the single mode constraint (12) imposes that the machine can only be set up for at most one item in each period. For each item, production can be up to capacity if there is a set up (13).
The start up variables are modeled in constraint (14). There will only be a start up if the machine is set up for an item for which it was not set up in the previous period. Finally, the set up and start up variables are binary (15). Karmarkar and Schrage (1985) consider this problem without set up costs and called it the product cycling problem. Karmarkar et al. (1987) study the single item version of the CSLP, both for the uncapacitated and capacitated case. This problem is also referred to as lot sizing with start up costs (Wolsey 1989, Sandbothe 1991).

The Proportional Lot Sizing and Scheduling Problem (PLSP) relaxes the restriction of allowing production for only one product in each time period. In the PLSP at most two different items can be produced in each time period (Drexl and Haase 1995, 1996). There is still at most one set up in each period, but the set up from the previous period can be carried over to the next period.

Fleischmann (1990) proposes a generic model for the Discrete Lot Sizing and Scheduling Problem (DLSP). The difference with the CSLP is that a discrete production policy is assumed, implying that an item must be produced at full capacity. The generic model has a similar structure as the CSLP (10)-(15), except that the capacity and set up constraint (13) becomes an equality:

\[ \forall i \in P, \forall t \in T \quad (16) \]

Note that the production variable can be substituted out through this constraint. Jans and Degraeve (2004a) extended this model with industrial constraints such as general start up times and multiple alternative machines.

2.4. Multi-Level Lot Sizing

In a multi-level lot sizing problem, the production planning is not only done at the final level for the end product, but also for the components and subassemblies that are needed to make the end product. Production at one level leads to demand for components at a lower level (dependent demand). At the highest level, production is triggered by market demand (independent demand). Different kinds of product structures can be considered. Zangwill (1966b) analyses an uncapacitated multi-facility problem where the product flows through several facilities at different levels for a few special models such as facilities in series or parallel. Lambrecht and Vander Eecken (1978a) extend the model with a capacity constraint at the final level. Afentakis et al. (1984) reformulate the multi-stage problem with an assembly

3. Solution Approaches

Most lot sizing problems are hard to solve. Some complexity results can be found in Florian et al. (1980). They prove that the single item capacitated problem is NP hard for quite general objective functions. Problems with concave cost functions and no capacity limits (Wagner and Whitin 1958) or constant capacities (Florian and Klein 1971) are solvable in polynomial time. Also lot sizing with convex cost functions and no set up cost is polynomially solvable. Additional complexity results are given by Bitran and Yanasse (1982), Salomon et al. (1991), Brüggemann and Jahnke (1997) Vanderbeck (1998) and Webster (1999). Because these problems are so difficult, various solution techniques have been used to solve them. We provide a review of the various meta-heuristics that have been developed to solve lot sizing problems. Further, we briefly discuss other solution approaches. This allows us to discuss how these techniques relate to each other and list their respective advantages and disadvantages in Section 4.

3.1. Meta-Heuristics for lot sizing problems

In the past decade, meta-heuristics such as Tabu Search (TS), Simulated Annealing (SA) and Genetic Algorithms (GA) have become more and more popular for solving complex combinatorial problems. Readers who are not familiar with the basic concepts of these algorithms can find excellent introductions in Michalewicz and Fogel (2002), Glover and Laguna (1997) and Reeves (1997). General guidelines for the design of meta-heuristics are discussed in Hertz and Widmer (2003). One of the main reasons for the success of these meta-heuristics is their flexibility and ability to handle large and complex problems. As a consequence, these methods are usually developed for extensions of the standard lot sizing
problems for which no good special purpose algorithm exists and which are too difficult to solve with commercial integer optimization software:

1) Single level CLSP (Hindi 1996, Kohlmorgen et al. 1999), with set up times (Hung et al. 1999, Özdamar and Bozyel 2000), with set up times and overtime (Özdamar, Birbil and Portmann 2002), with set up times and multiple resources (Hung et al. 2003), with set up carry over (Gopalakrishnan et al. 2001), with sequence dependencies (Fleischmann and Meyr 1997, Laguna 1999, Meyr 2000), on parallel machines (Özdamar and Birbil 1998);

2) Multi level lot sizing without capacities (Kuik and Salomon 1990, Dellaert and Jeunet 2000, Dellaert et al. 2000, Tang 2004);

3) Capacitated multi level lot sizing (Kuik et al. 1993, Barbarosoğlu and Özdamar 2000), with set up times (Salomon et al. 1993, Hung and Chien 2000, Özdamar and Barbarosoğlu 2000, Xie and Dong 2002), and loading decisions (Özdamar and Barbarosoğlu 1999);

4) Multi level PLSP (Kimms 1996, 1999);

5) CSLP (Hindi 1995).

6) DLSP with batch availability (Brüggemann and Jahnke 1994, 2000)

Other meta-heuristics such as Neural networks (NN) (Aarts et al. 2000, Gaafar and Choueiki 2000) and Treshold Accepting (TA) (Meyr 2000) have only been used on a limited scale for solving lot sizing problems.

There is a lot of freedom in the design of these meta-heuristics. Therefore we do not discuss the details of each specific implementation but we refer the interested reader to the specific papers. We restrict ourselves mainly to the discussion of the general components of these meta-heuristics and how they are implemented for dynamic lot sizing. In the design of a SA and TS algorithm, the major questions that need to be answered are (Michalewicz and Fogel 2002): 1) how do we represent a solution; 2) how do we evaluate a solution; and 3) how do we define a neighbourhood. Further, we need an initial solution and there are also many control parameters that need to be set such as the initial temperature, cooling ratio and a stopping criterion for SA and the memory structure and length, an aspiration criterion and a stopping rule for TS. For a GA, the major decisions concern: 1) the representation; 2) evaluation; 3) the construction of genetic operators to generate offspring; and 4) the choice of a selection mechanism to determine the next population. Further implementation issues are for example the initial solution, generation size and repair operators.
In the **representation** of a solution, whether it is for TS, SA or GA, we have two basic options for a **direct representation** as we are dealing with Mixed Integer Programming problems containing both integer variables for the set ups and possibly the sequencing decisions and continuous variables for the production quantities. In the first option, we explicitly include both the integer variables and the production quantities in the representation. (Özdamar and Birbil 1998, Özdamar and Barbarosoğlu 1999, Barbarosoğlu and Özdamar 2000, Özdamar and Barbarosoğlu 2000, Özdamar and Bozyel 2000, Gopalakrishnan et al. 2001, Özdamar, Birbil and Portmann 2002). The second option is to include only the integer variables and discard the production quantities. Each setting of the integer variables corresponds with an optimal value for the production variables. This solution can be found by solving a Linear Programming model, which is the original MIP problem with the integer variables fixed (Kuik and Salomon 1990, Hindi 1995, Hindi 1996, Hung et al. 1999, Hung and Chien 2000, Meyr 2000, Hung et al. 2003). As this calculation is repeated many times, it is of paramount importance that the LP can be solved efficiently. If this is not the case, the production quantities can be chosen heuristically (Kuik et al. 1993, Salomon et al. 1993, Fleischmann and Meyr 1997, Xie and Dong 2000). Another solution is to reoptimize the latest LP as a move usually only induces a small change (Hindi 1995, Hindi 1996, Meyr 2000, Hung et al. 2003). Using a dual method for calculating optimal production values for a given set up schedule has the advantage that expensive solutions can be detected early and the problem does not have to be solved to optimality (Meyr 2000). In case of the uncapacitated multi-level problem (Dellaert and Jeunet 2000, Dellaert et al. 2000, Tang 2004), the set up decisions automatically determine the optimal production quantities through the zero-switch property. Compared to the first option, this representation with only integer variables is less complex, but we have to solve additional optimization problems to obtain a complete solution.

There is also the possibility of an **indirect representation** of the solution. Kimms (1999) present a genetic algorithm for the multi-level multi-machine PLSP. The representation of a solution is quite different from the general bitstrings in the previously discussed algorithms. The chromosome is a two-dimensional matrix and each entry represents a rule for selecting the set up state for a machine at the end of the period. Choosing the item with maximum holding costs, minimum set up cost, maximum depth, maximum number of predecessors are examples of such selection rules. To compute the fitness value of chromosome, a construction scheme is called which uses the rules that define the chromosome to build a solution. Kimms
(1996) represents the solution of the same problem for a single machine by a graph structure on which a tabu search is performed and the production plan is generated by a specific construction scheme. Kohlmorgen et al. (1999) experiment with a parallel genetic algorithm for the CLSP. A solution is represented by a string of real values, which control a heuristic method for generating a production plan. For the CLSP with set up times, Özdamar and Bozyel (2000) apply genetic operators to priority vectors which are used to construct a feasible solution.

There are several options available for evaluating a solution. The most obvious choice for the evaluation function is the objective function. However, in a GA it is possible to obtain an infeasible solution after applying the genetic operators and for TS and SA a move can also lead to an infeasible neighbor. A main issue is how to treat such infeasible solutions. One option is to discard all infeasible solutions or attach an infinite cost to them (Kimms 1999, Dellaert et al. 2000). A second option is to impose some penalty cost proportional to the infeasibility. An example is to impose a backlog cost for demand which cannot be met in time (Barbarosoğlu and Özdamar 2000, Özdamar and Barbarosoğlu 2000, Xie and Dong 2002), a penalty for capacity violation (Özdamar and Birbil 1998, Özdamar and Barbarosoğlu 1999, Barbarosoğlu and Özdamar 2000, Özdamar and Barbarosoğlu 2000, Gopalakrishnan et al. 2001) or a high cost for initial inventory (Meyr 2000). A third option is to devise some repair operators which are invoked when a solution is infeasible (Özdamar and Birbil 1998, Özdamar and Barbarosoğlu 1999, Dellaert et al. 2000, Xie and Dong 2002). Finally, a two-phase algorithm can be applied (Brüggemann and Jahnke 2000) where first an initial feasible solution is constructed and next the optimization within the feasible region is started. For GA, some sort of scaling is usually done where the fitness of a chromosome is calculated relative to the best solution in the population (Hung et al. 1999, Hung and Chien 2000, Dellaert and Jeunet 2000, Dellaert et al. 2000, Özdamar and Bozyel 2000, Xie and Dong 2002). Other options such as ranking or tournament selection (Reeves 1997) are not used for the lot sizing problem.

The definition of the neighborhood for SA and TS depends on the specific problem at hand and the representation of the solution. In a representation with both integer and continuous variables, moves can be defined on both parts (Gopalakrishnan et al. 2001). Neighbor solutions can also be generated by partially or completely shifting the lot of one item (Özdamar and Birbil 1998, Özdamar and Barbarosoğlu 1999, Özdamar and Barbarosoğlu
2000, Barbarosoğlu and Özdamar 2000, Özdamar, Birbil and Portmann 2002). For the representation with only the integer variables, the simplest mechanism only alters the set up variable for one product-period combination. This is the most widely used neighborhood definition (Kuik and Salomon 1990, Kuik et al. 1993, Salomon et al. 1993, Hindi 1995, Hindi 1996, Hung et al. 1999, Hung and Chien 2000, Hung et al. 2003). More sophisticated transition mechanisms have also been proposed (Kuik and Salomon 1990, Meyr 2000).

In a GA, we use genetic operators to explore the solution space. The mutation operator changes a single solution. The most popular mutation operator is the single bit flip (Xie and Dong 2002, Dellaert and Jeunet 2000, Dellaert et al. 2000), where the set up is changed for a randomly chosen item and period. In the case lot sizes are incorporated into the solution, the mutation operator can change the lot size by a random amount (Özdamar and Birbil 1998, Özdamar and Barbarosoğlu 1999, Özdamar, Birbil and Portmann 2002). A cross-over combines elements from several solutions into a new one. In the one point cross-over (Dellaert and Jeunet 2000, Xie and Dong 2002) the bit strings of the two parents are cut in two at some random point and are recombined into one new solution. More elaborate crossovers have also been devised (Özdamar and Birbil 1998, Hung et al. 1999, Özdamar and Barbarosoğlu 1999, Hung and Chien 2000, Dellaert et al. 2000, Özdamar, Birbil and Portmann 2002). The selection operator determines which chromosomes from the current population and the newly created offspring will form the new population. In a popular random strategy the probability of selection of each chromosome is proportional to its fitness value (Hung et al. 1999, Dellaert et al. 2000, Hung and Chien 2000, Xie and Dong 2002). An elitist strategy also keeps the best chromosomes from the current population (Dellaert et al. 2000, Özdamar and Bozyel 2000). Selection can also be done deterministically, choosing the chromosomes with the highest fitness value (Kimms 1999, Özdamar, Birbil and Portmann 2002). A genetic algorithm can also be applied to multiple parallel populations (Özdamar, Birbil and Portmann 2002). The migration operator allows for a cross-over between chromosomes from different populations.

3.2. Dynamic Programming

Wagner and Whitin (1958) proposed a Dynamic Programming (DP) recursion for the single item uncapacitated lot sizing problem. They prove that there exists an optimal solution that
satisfies the following property: \( s_{t-1} x_t = 0, \forall t \in T \). This property implies that there exists an optimal solution in which one never produces in a period and at the same time has inventory coming in from the previous period. As a consequence, production in one period satisfies the demand for an integral number of consecutive periods. Based on this special property of the optimal solution, Wagner and Whitin formulate a dynamic programming recursion of \( O(m^2) \).

The properties of an optimal solution can also be proven by network theory (Zangwill 1969). Diaby (1993) develops an efficient post-optimization procedure to recompute the optimal schedule starting from the Wagner-Whitin solution for the case where some set ups are imposed or prohibited. A recent milestone in the history of the single item uncapacitated lot sizing problem is the independent discovery of an \( O(m \log m) \) algorithm for this problem by Federgruen and Tzur (1991), Wagelmans, Van Hoesel and Kolen (1992) and Aggarwal and Park (1993). The algorithms run in \( O(m) \) time for special cases.

Also for other single item lot sizing problems, DP recursions have been developed based on the structure of an optimal solution. The extension with backlogging is discussed in Zangwill (1966a), Van Hoesel et al. (1994) and Federgruen and Tzur (1993). Florian and Klein (1971) consider the single item lot sizing problem with a constant capacity and concave costs, whereas Lambrecht and Vander Eecken (1978b) extend this analysis by looking at an arbitrary capacity structure. Van Hoesel and Wagelmans (1996) propose an \( O(m^2) \) DP recursion for the case of constant capacities, concave production costs and linear holding costs. Shaw and Wagelmans (1998) improve the DP recursion of Florian and Klein for general costs structures, as long as the production cost function is piecewise linear. Many other researcher have proposed DP algorithms for the single item lot sizing problem and special cases (Lippman 1969, Florian et al. 1980, Bitran and Yanasse 1982, Chung and Lin 1988, Kirca 1990, Pochet and Wolsey 1993). Further, Van Hoesel and Wagelmans (2001) show how DP algorithms can be used to construct a fully polynomial approximation scheme for the single item capacitated lot sizing problem. DP results for the DLSP can be found in Cattrysse et al. (1993), Van Hoesel and Kolen (1994), Van Hoesel et al. (1994) and for the CSLP in Vanderbeck (1998).

3.3. Polyhedral Results and Strong Valid Inequalities

Much attention is devoted to the polyhedral study of the convex hull of a problem. The goal of this line of research is to describe the convex hull of a problem by valid inequalities. Next,
one can apply this knowledge to a substructure of a larger problem to obtain improved lower bounds. Excellent overviews on polyhedral approaches for lot sizing problems are given in Pochet and Wolsey (1995), Pochet (2001) and Wolsey (2002). The research into strong valid inequalities for lot sizing problems was initiated by Barany, Van Roy and Wolsey (1984\textsuperscript{a,b}) who describe the convex hull of the single item uncapacitated lot sizing problem. They define the $(l, S)$ inequalities as follows:

$$
\sum_{i \in S} x_i \leq \sum_{i \in S} s_{ij} y_i + s_l \quad \forall l \in T, \forall L = \{1, \ldots, l\}, \forall S \subseteq L \tag{17}
$$

These cutting planes can be used to strengthen the formulation of other problems such as the capacitated problems. Pochet and Wolsey (1988) present strong valid inequalities for the case with backlogging, which describe the convex hull of the single item uncapacitated problem. Pochet (1988) and Leung, Magnanti and Vachani (1989) derive several valid inequalities for the capacitated problem. Polyhedral results and cutting planes for problems with a Wagner-Whitin cost structure, i.e. $\nu_c + h_c \geq \nu_{c_{i+1}}$, are presented in Pereira and Wolsey (2000) and Pochet and Wolsey (1994). Other valid inequalities for the capacitated lot sizing problem and variants are discussed in Pochet and Wolsey (1993), Constantino (1998), Miller, Nemhauser and Savelsbergh (2000\textsuperscript{a,b}), Loparic, Marchand and Wolsey (2003), Van Vyve and Ortega (2003) and Van Vyve (2003). The polyhedral structure and valid inequalities for a single period relaxation of the CLSP with set up times are presented in Miller et al. (2003\textsuperscript{a,b}). They study the ‘Single-Period Relaxation with Preceding Inventory’, which includes the joint capacity constraint, the demand constraints with only the preceding inventory variables and the set up constraints for multiple items in a specific period $t$. Miller et al. (2000\textsuperscript{b}) present an efficient branch-and-cut algorithm for the CLSP with set up times using these valid inequalities of the single period relaxation. A good overview of several inequalities for various models, such as capacitated models, lot sizing with start ups and multi-stage problems, is given in Pochet and Wolsey (1991). Belvaux and Wolsey (2000) report on an efficient branch-and-cut system that includes preprocessing and cutting planes for a variety of lot sizing models.

Polyhedral results for the Discrete Lot Sizing and Scheduling problem can be found in Magnanti and Vachani (1990), Van Eijl and Van Hoesel (1997), Magnanti and Sastry (2002) and Loparic et al. (2003). Vanderbeck and Wolsey (1992) show how to solve the Lasdon-Terjung model more efficiently with cutting planes. Valid inequalities and polyhedral results

3.4. Lagrange Relaxation and Dantzig-Wolfe Decomposition

Dantzig-Wolfe decomposition and Lagrange relaxation are two other techniques for finding improved lower bounds. The basic idea is to divide the problem into smaller subproblems which are much easier to solve and a coordinating mechanism ensures that we obtain a good approximation for the overall problem. It is well known that the optimal values resulting from Dantzig-Wolfe decomposition and Lagrange relaxation are the same (Geoffrion 1974, Fisher 1981). One formulation is the dual of the other. In the Dantzig-Wolfe decomposition, the linking or complicating constraints are put in the master whereas in the Lagrange relaxation they are dualized in the objective function. The subproblems that result from these two approaches are equal.

As an example, we consider the Capacitated Lot Sizing Problem (5)-(9) with set up times. Manne (1958) proposed an innovative formulation for solving this problem. He explicitly models all the possible schedules with different set up sequences. For a problem with a planning horizon of $m$ periods, there are $2^m$ different set up schedules possible for each product, as for each period we either have a set up or not. Manne only considers ‘dominant’ schedules, which have the property that in each period demand will be met by production in that period if there is a set up or otherwise from the nearest preceding period with a set up. The model that Manne proposes is actually the full blown master of the Dantzig-Wolfe reformulation (Dantzig and Wolfe 1960) with the capacity constraints as the linking constraints. Dzielinski and Gomory (1965) propose the use of column generation to counter the difficulty of the large amount of variables in Manne’s formulation. Column generation starts with a feasible restricted master with only a few columns and new columns are added iteratively as they are needed. At each iteration of the column generation procedure, we solve a separate single item uncapacitated subproblem for each item $i$, where the objective is to minimize the reduced cost. If the new column has a negative reduced cost, we add it to the master. Solving the master again provides new dual prices which are used to generate new columns. If we cannot find any new column with a negative reduced cost, we have solved the master’s LP relaxation to optimality with an objective value of $v_{dw}$. All the columns that are
generated as such will have the Wagner-Whitin property. Manne’s dominant schedules are in fact all the schedules that satisfy the Wagner-Whitin property. Manne further proves that the LP solution of his formulation will naturally have most variables at zero or one if the number of items is much larger than the number of time periods in the planning horizon and as such it will be a good approximation to the optimal IP value. There is however a structural problem with Manne’s formulation. The formulation with integrality constraints on the variables will not necessarily give the optimal solution to the multi-item lot sizing problem as the schedules in the optimal solution usually do not have the Wagner-Whitin property. Lambrecht and Vanderveken (1979) and Bitran and Matsuo (1986), among others, note that the solution of Manne’s formulation with integrality constraints is only a subset of the solutions of the original IP problem. Degraeve and Jans (2003) present a new Dantzig-Wolfe reformulation model of this problem which does not suffer from this structural deficiency. They separate the integer set up decisions and the continuous production decisions. Bitran and Matsuo (1986) provide an error bound for the solution obtained by using the production schedule proposed by Manne’s formulation and properly accounting for the set ups. The Dantzig-Wolfe decomposition for the DLSP is discussed in Lasdon and Terjung (1971), Cattrysse et al. (1993), and Jans and Degraeve (2004a) and for the CSLP in Vanderbeck (1998).

In the Lagrange relaxation of the CLSP (Kleindorfer and Newson 1975, Thizy and Van Wassenhove 1985, Trigeiro et al. 1989, Diaby et al. 1992b), the capacity constraint (8) is dualized into the objective function (5) with a specific set of positive multipliers $p = \{p_1, p_2, \ldots, p_m\}$:

$$
\bar{v}_{LR}(p) = \min \sum_{i \in P} \sum_{t \in T} \left( sc_i y_{it} + vc_i x_{it} + hc_i s_{it} \right) - \sum_{t \in T} p_t \left( cap_t - \sum_{i \in P} st_i y_{it} - \sum_{i \in P} vt_i x_{it} \right)
$$

(18)

The Lagrange problem also decomposes into separate single item uncapacitated subproblems for each item $i$. The Lagrange Dual problem consists of finding the maximum lower bound $\bar{v}_{LD}$:

$$
\bar{v}_{LD} = \max_{p \geq 0} \bar{v}_{LR}(p).
$$

In iterative steps, the multipliers are adapted in the hope of finding a better lower bound. The most popular method for updating the dual prices is the subgradient method. Du Merle et al. (1997) provide a competitive alternative by using an interior point cutting plane method for finding new dual prices. Similar Lagrange relaxation schemes have been proposed for the DLSP (Fleishmann 1990, 1994) and the CSLP (Karmarkar and Schrage 1985, Karmarkar et al. 1987)
Most of the decomposition schemes relax the capacity constraint. However, other decompositions are possible. Diaby et al. (1992) discuss a new approach in which they dualize the demand constraint. Here the subproblem consists of a combination of set up and knapsack constraints for which they derive a specialized branch-and-bound algorithm. From computational experiments they conclude that the capacity relaxation is much more efficient than the demand relaxation. Jans and Degraeve (2004) study yet another decomposition. They start from the network reformulation, which is discussed in the next section, and dualize the demand constraints. This decomposition results in improved lower bounds, as the network formulation already gives the same lower bound as the relaxation of the capacity constraint. For a multi-machine DLSP, De Matta and Guignard (1994) also propose to dualize the demand constraint.

For multi-level problems, relaxation of the capacity constraints results in a single uncapacitated multi-level, multi-item problem. Relaxation of both the capacity constraints and demand constraints results in single-item uncapacitated subproblems (Tempelmeier and Derstroff 1996). Also the reformulation in terms of echelon stock (Afentakis et al. 1984, Afentakis and Gavish 1986, Clark and Armentano 1995, Belvaux and Wolsey 2001) allows the problem to be decomposed into uncapacitated single item subproblems. The echelon stock for an item $i$ (Clark and Scarf 1960) is defined as all the units of item $i$ that are held in inventory in the whole system, either as item $i$ or as a component in an assembly.

Dantzig-Wolfe decomposition and Lagrange relaxation are essentially techniques for finding tighter lower bounds. They need to be extended with other heuristics if we want to find good upper bounds or they can be integrated within a B&B scheme to find an optimal solution. Starting from the solutions of the Lagrange subproblems, heuristic solutions can be obtained by shifting production quantities (Trigeiro et al. 1989, Tempelmeier and Derstroff 1996). Fixing the set up variables according to the Lagrange solution and next solving a transportation problem to obtain a feasible solution is also done (Thizy and Van Wassenhove 1985, Diaby et al. 1992). Cattrysse, Maes and Van Wassenhove (1990) discuss the set partitioning formulation of the problem for which schedules are obtained by heuristics for the multi-item capacitated lot sizing problems. Hence the individual schedules do not necessarily have the Wagner-Whitin property. They find feasible solutions either by rounding up the fractional set up variables and improving the solution or by solving a smaller problem heuristically for the items with a fractional value. During column generation, Degraeve and

3.5. Reformulations of the Lot Sizing Problem

It is well known that for (mixed) integer programming problems, some formulations are better than others in the sense that they provide tighter bounds. For the lot sizing problem, three such reformulations have been proposed. Krarup and Bilde (1977) propose the simple plant location reformulation:

\[
\text{Min} \sum_{t=1}^{m} s c_t y_t + \sum_{t=1}^{m} \sum_{l=1}^{m} c_t d_t w_{tl}
\]

s.t. \[ \sum_{l=1}^{T} w_{tl} = 1 \quad \forall l \in T \]  
\[ w_{tl} \leq y_t \quad \forall l, t \in T, t \leq l \]  
\[ y_t \leq 1 \quad \forall t \in T \]  
\[ w_{tl}, y_t \geq 0 \quad \forall l, t \in T, t \leq l \]

The variable \( w_{tl} \) is the fraction of the demand of period \( l \) produced in period \( t \). In the objective function (19) we find the set up cost and an adjusted production cost, \( c_t = vc_t + \sum_{l=1}^{m} hc_t \), as this is a formulation without inventory variables. Constraint (20) imposes that the demand in each period is satisfied from production in that period or from a previous period. Further we have the fixed charge constraint (21) and the simple upper bound on the \( y_t \) variables (22) and non-negativity constraints (23). This reformulation gives integer solutions to the single item uncapacitated lot sizing problem. Pochet and Wolsey (1995) discuss the multi-commodity flow reformulation, which is equivalent to the facility location formulation after variable transformation.

Eppen and Martin (1987) describe another reformulation, which is in fact the network or shortest path formulation of the DP recursion used by Wagner and Whitin. They introduce a
variable $z_{vk}$ which represents the fraction of the demand from period $t$ to $k$ that will be satisfied by production in period $t$. Based on the property of an optimal solution, they impose that the $z_{vk}$ variable is binary for a single item uncapacitated problem. If $z_{vk}$ equals one, then we produce in period $t$ the demand for period $t$ until $k$. As such we can describe the lot sizing problem as a shortest path network. On every arc, a cost $cv_{vk}$ is defined as the total cost for producing in period $t$ the demands for period $t$ until $k$ and the according inventory holding cost: $cv_{tk} = vc_s d_{sk} + \sum_{s=t+1}^{k} \sum_{u=t}^{s-1} hc_u d_s$. The single item uncapacitated lot sizing problem can then be reformulated as follows:

$$\text{Min } \sum_{t=1}^{m} sc_t y_t + \sum_{t=1}^{m} \sum_{k=1}^{m} c_{vk} z_{vk}$$

s.t. 

$$1 = \sum_{k=1}^{m} z_{vk}$$ 

$$\sum_{s=t}^{t-1} z_{v_{s,t-1}} = \sum_{k=t}^{m} z_{v_{t,k}} \quad \forall t \in T \setminus \{1\}$$

$$\sum_{t=1}^{m} z_{v_{tm}} = 1$$

$$\sum_{k=t}^{m} z_{vk} \leq y_t \quad \forall t \in T$$

$$z_{vk} \geq 0; y_t \in \{0,1\} \quad \forall t, k \in T, k \geq t$$

Constraints (25) to (27) are the conservation of flow equalities for the shortest path network. Constraint (28) is the set up forcing constraint. Eppen and Martin show that the LP relaxation of this formulation has an integer solution for ULS. This network formulation can also be used to obtain tighter lower bounds for multi-item capacitated problems.

As the reformulations provide improved lower bounds, their LP solution can also be used to construct heuristic solutions. Maes et al. (1991) discuss such LP-based rounding heuristics for the multi-level problem, but the ideas can also be applied to single level problems (Alfieri et al. 2002). The single pass branching algorithm fixes the largest non integer set up variable to one and resolves the problem until an overall integer solution is obtained. In a variant, sets of variables are set to one and zero at the same time. In the curtailed branch-and-bound algorithm, all the variables which are integer in the LP solution are fixed and a limited branch-and-bound is performed on the remaining set up variables. The curtailed branch-and-bound results in a smaller gap compared to the single pass branching algorithm, but at the expense of a much higher CPU time. The fixing of sets of variables appears to be more efficient than branching on one variable at the time.

3.6. Special-Purpose Heuristics for Lot Sizing Problems

Several dedicated heuristics have been proposed for solving the uncapacitated single item lot sizing problem. These heuristics are often used as lot size rules in MRP systems instead of the Wagner-Whitin algorithm. A discussion of many of these lot size rules such as the Economic Order Quantity, Period Order Quantity, Least Period Cost (Silver and Meal 1973), Least Unit Cost, Part Period Balancing, Least Total Cost, Marginal Cost Difference (Groff 1979) and the Technique for Order Placement and Sizing (Coleman and McKnew 1990) can be found in Baker (1993) or in most textbooks on production planning (Silver et al. 1998). A comparison of many of these heuristics can be found in Coleman (1992) and Simpson (2001). Bitran et al. (1984), Axssäter (1985) and Vachani (1992) present worst-case bounds for some of these heuristics for several demand classes.

Much research has been done on the more challenging problem of multi-item capacitated lot sizing. Some heuristics are based on a Silver-Meal criterion of the average cost per period (Eisenhut 1974, Lambrecht and Vanderveken 1979). Dixon and Silver (1981) present another heuristic where the criterion is to select that item for which a one period increase in the supply results in the largest decrease in average cost per unit time (Silver Meal cost criterion) per unit of capacity absorbed. The Silver-Meal criterion has also been extended for the case of regular and overtime production (Dixon et al. 1983) and set up times (Trigeiro 1989). Maes and Van Wassenhove (1986) use in their ABC heuristic several cost criteria to determine whether or
not to include a next period and several rules for determining the order of the items. Dogmaraci et al. (1981) develop a forward-sweep algorithm and a greedy search starting from the lot-for-lot solution. In an extensive computational experiment, Maes and Van Wassenhove (1986a,b) compare three of the heuristics discussed above: Lambrecht and Vanderveken (LV), Dixon and Silver (DS) and Dogmaraci et al. (DPA). On average, DPA is slightly better than the two others, but at the expense of considerable higher CPU times. LV and DS are comparable both in the quality of solution and CPU times. The cost structure and tightness of the capacity constraint seem to have a large impact on the performance. Van Wassenhove and De Bodt (1983) apply the Lambrecht and Vanderveken and Dixon and Silver heuristics to solve a real live planning problem in the injection moulding industry. Another period-by-period heuristic is discussed in Günther (1987). Starting from a lot-for-lot schedule, Günther attempts to smooth the capacity profile using marginal cost savings as a priority index. Karni and Roll (1982) start from the WW schedules and try to achieve feasibility and improve the cost by shifting production. Van Nunen and Wessels (1978) present a separate lot size determination and capacity smoothing procedure. Kirca and Kökten (1994) develop a very effective heuristic for the multi-item capacitated lot sizing problem. They use an item-by-item approach in which items are selected one at a time. Next a single item problem is solved with adapted capacities and extra bounds on production and inventory to ensure feasibility of the overall problem. The best results are obtained by selecting the item with the largest average cost per unit demand, where this cost is based on the EOQ formula. The computation times are comparable with those for the Lambrecht and Vanderveken and Dixon and Silver heuristics, but they have a superior solution quality. These dedicated heuristics have also been extended to solve more complex lot sizing problems. An example is the lot sizing problem with sequence-dependent set up times, batch production and storage constraints (Selen and Heuts 1990) for which a cost savings based heuristic is developed.

Another kind of heuristic is based on a mathematical programming approach. We have already discussed how the improved LP relaxation from Dantzig-Wolfe decomposition, Lagrange relaxation and reformulations can be used as a starting base to obtain good heuristic solutions. Maes and Van Wassenhove (1988) conclude that mathematical programming based heuristics are more flexible and perform better in general compared to the special-purpose heuristics. Good solutions are also obtained by solving the problems in a rolling horizon way (Belvaux and Wolsey 2000, Stadler 2003). Set up decisions are only considered for a specific time window, which is moved forward in iterative steps. Set up decisions before the time
Hung and Hu (1998) start from a solution where a setup is done in each period. The algorithm solves LP’s iteratively and decides in which periods it is better not to have a setup based on the shadow price information. Newson (1975) sequentially deletes arcs in an uncapacitated shortest path network as long as some capacity restrictions are violated.

For the multi-level case, efficient heuristics follow a level-by-level approach, but modify the setup and inventory costs at lower levels to model the interdependencies (Blackburn and Millen 1982, Bookbinder and Koch 1990, Billington et al. 1994, Tempelmeier and Helber 1994). Heuristics based on a similar idea of iteratively solving LP’s with modified cost and constraint coefficients are proposed by Harrison and Lewis (1996) for the capacitated case and by Katok et al. (1998) for the capacitated case with setup times. A recent overview of various heuristics for multi-level lot sizing problems is provided in Dellaert and Jeunet (2003).

4. Comparing methodologies

Lot sizing problems are hard to solve, as the regular IP formulations usually have a large gap. We have discussed several solution approaches that aim to find tighter lower bounds or optimal solutions. Dynamic programming recursions have been developed for several special cases, based on insights into the structure of the optimal solution. The drawback of these DP algorithms is that they are developed for single item problems and cannot be directly extended to multi-item problems. However, these single item problems are very important as they appear as core structures in more complex problems. Many efficient algorithms exploit the specific knowledge about such a substructure to find tighter lower bounds. Dantzig-Wolfe decomposition and Lagrange relaxation are examples of such algorithms. However, in practice these methodologies exhibit convergence problems. Lagrange relaxation is usually done only for a limited number of iterations and there is no guarantee that we have found the optimal bound. Dantzig-Wolfe decomposition provides both an upper and lower bounds for the optimal relaxation value, but the method suffers from a slow convergence towards the end, known as the tailing off effect. A further disadvantage of Lagrange relaxation is that it does not provide a primal solution. In Dantzig-Wolfe decomposition, the intermediate primal solution of the current master can be exploited in a LP fixing heuristic to find an integer solution. Cutting planes also make use of specific knowledge about the polyhedral
(sub)structure. Usually a class of cutting planes has an exponential number of members, but only a small number of them are generated as they are needed. Therefore, a customized separation algorithm is usually developed. A disadvantage is that there is no structured way of finding good cuts. A general drawback of DW, LR and CP is that they only provide a lower bound and have to be extended with heuristics or complemented by branch-and-bound to obtain a feasible solution. Yet another way of finding tighter models is variable redefinition. The main advantage of the variable redefinition is that commercial optimization software can be used instead of a custom-tailored algorithm (Wolsey 2002). The techniques described above have the same goal: obtaining strong lower bounds by approximating the convex hull of the problem. As an illustration, Table 1 compares several lower bounds for 6 test instances of the Capacitated Lot Sizing Problem with Set Up Times (Trigeiro et al. 1989, Belvaux and Wolsey 2000, Jans and Degraeve 2004a), with 6, 12 and 24 products and 15 and 30 periods.

Table 1. Comparison of lower bounds

<table>
<thead>
<tr>
<th></th>
<th>LP</th>
<th>1 period</th>
<th>WW</th>
<th>LAG1</th>
<th>DW</th>
<th>BW</th>
<th>MNS</th>
<th>LAG2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tr6-15 (G30b)</td>
<td>13,023.8</td>
<td>32,042.2</td>
<td>36,793</td>
<td>37,198.6</td>
<td>37,201.2</td>
<td>37,213.3</td>
<td>37,319</td>
<td>37,380.3</td>
</tr>
<tr>
<td>Tr6-30 (G62)</td>
<td>16,530.2</td>
<td>47,072.2</td>
<td>60,351</td>
<td>60,946.1</td>
<td>60,946.2</td>
<td>60,979.4</td>
<td>61,150</td>
<td>61,189.1</td>
</tr>
<tr>
<td>Tr12-15 (G53)</td>
<td>24,738.5</td>
<td>61,703.0</td>
<td>69,412</td>
<td>73,756.8</td>
<td>73,847.9</td>
<td>73,858.2</td>
<td>73,929</td>
<td>73,942.7</td>
</tr>
<tr>
<td>Tr12-30 (G69)</td>
<td>29,179.9</td>
<td>91,317.7</td>
<td>129,377</td>
<td>130,176.9</td>
<td>130,177.2</td>
<td>130,177</td>
<td>130,292</td>
<td>130,330.4</td>
</tr>
<tr>
<td>Tr24-15 (G57)</td>
<td>43,021.4</td>
<td>110,831.2</td>
<td>135,843</td>
<td>136,363.7</td>
<td>136,365.7</td>
<td>136,366</td>
<td>136,388</td>
<td>136,417.9</td>
</tr>
<tr>
<td>Tr24-30 (G72)</td>
<td>57,136.1</td>
<td>202,958.9</td>
<td>287.235</td>
<td>287,753.2</td>
<td>287,753.4</td>
<td>287,753</td>
<td>287,811</td>
<td>287,824.3</td>
</tr>
</tbody>
</table>

The regular LP relaxation is given in the column LP. Next, we add a limited number of \((l,S)\) inequalities with only one period: \(x_i \leq d_i y_i + s_i, \forall t \in T\). This already substantially improves the lower bound. Also disregarding the capacity constraint and just solving the single item uncapacitated Wagner-Whitin problems result in a better lower bound (WW). Relaxing the capacity constraint in the Lagrange relaxation (LAG1) results theoretically in the same lower bound as keeping it as the linking constraint in the master of the Dantzig-Wolfe reformulation (DW). In practice, however, the Lagrange bound will usually be smaller because of convergence issues. This Dantzig-Wolfe bound (DW) is also equal to the ones obtained by the Plant Location and Shortest Path reformulation and the regular formulation plus all the \((l,S)\)-inequalities. Belvaux and Wolsey (2000) (BW) present a branch-and-cut algorithm, consisting of both general cutting planes and specific lot sizing inequalities. Compared to DW, they obtain better bounds for the smaller problems and the same bounds for the larger problems.
Miller et al. (2000b) (MNS) derive cuts for the ‘Single Period relaxation with Preceding Inventory’, while Jans and Degraeve (2004b) apply a Lagrange relaxation on the shortest path reformulation (LAG2).

Research into optimal algorithms is part of the process of continuous improvement. Problems that are unsolvable today, may become solvable tomorrow due to new research developments. Heuristics are important as well and there is a clear interaction between these heuristics and optimal algorithms. Heuristics usually provide a good starting solution for optimal algorithms and allow for early pruning in branch-and-bound based algorithms. Further, heuristics are necessary when optimal algorithms fail to solve very large or complex problems. The literature review shows that some of the best heuristics are also based on optimal algorithms. For the standard problems, traditional method or special purpose heuristics seem to outperform the meta-heuristics (Gopalakrishnan et al. 2001). On the other hand, meta-heuristics provide good results for the more difficult problems such as multi-level or sequence dependent problems, for which the traditional methods fail.

A major disadvantage of these meta-heuristics is the fact that no lower bound is available unless it is calculated separately. As such it is difficult to assess the solution quality. The flexibility also implies that the meta-heuristics are used to solve a variety of problems which all have (slightly) different assumptions. Hence it is very difficult to compare these algorithms as they are customized for a specific case and they are usually not tested on standard test sets. Although one often quoted advantage is that meta-heuristics are easy to understand, most of these algorithms are fairly complex because of all the special adaptations that are needed to make them work better.

Hybrid systems have been developed to combine the strengths of different methodologies. Solutions obtained by LP-based heuristics provide good starting solutions for meta-heuristics (Hindi 1996, Kuik et al. 1993). Elements from SA and TS have also been used as a diversification strategy in GA’s (Özdamar and Birbil 1998, Hung et al. 1999, Özdamar and Barbarosoğlu 1999, Özdamar, Birbil and Portmann 2002). A local search with TS or SA is done on randomly selected chromosomes of the current population and they are replaced by the best found solutions. This also helps to restore feasibility in chromosomes. Meta-heuristics have also been used to obtain primal solutions within Lagrange relaxation algorithms (Özdamar and Barbarosoğlu 1999, Özdamar and Barbarosoğlu 2000). Hung et al. (2003) use
information from the simplex sensitivity analysis to rank candidate neighbors within a tabu search framework. The best neighbor is more likely to be among the candidates with the highest rank. In a first modification, this ranking is used to reduce the neighborhood. A second algorithm searches the sorted neighbors sequentially and takes the first solution with a better objective function as the new solution. Computational experiments demonstrate that especially this last combination is very effective.

Vanderbeck (1998) combines branch-and-price and cutting planes for solving the CLSP. Dantzig-Wolfe decomposition can be combined with Lagrange relaxation to speed up the column generation process and obtain faster convergence, either by using Lagrange relaxation to solve the master (Cattrysse et al. 1993, Jans and Degraeve 2004a) or by using Lagrange relaxation on the original problem to generate new columns (Degraeve and Jans 2003). To obtain stronger bounds, Dantzig-Wolfe decomposition is done on a variable redefinition reformulation for the CLSP (Jans and Degraeve 2004b). There is an advantage in using DP for solving the subproblems in the Dantzig-Wolfe decomposition. Not only the column with the minimum reduced cost is available, but also the suboptimal solutions and their costs. This can be exploited by adding several columns at the same time to accelerate convergence (Jans and Degraeve 2004a).

5. Computational experiments for meta-heuristics

In order to increase the quality of computational experiments, researchers should take into account some guidelines. These guidelines hold for any algorithm, but are particularly relevant in the experimental evaluation of meta-heuristics as they are usually developed for a specific extension of a standard problem for which no other algorithms are available yet. For an excellent tutorial on computational experiments for heuristics, we refer the reader to Rardin and Uzsoy (2001). As a general remark, they state that once several algorithms for a specific problem have been tested in the literature, simply demonstrating that an algorithm is possible is not a valid research question anymore. In order to have a valid contribution, the research must either produce a new inside, or demonstrate the superiority of the new algorithm. First, the trade-off between the speed and solution quality must be evaluated. The aim of the heuristics is to provide a good quality solution within a reasonable computation time. Generally, large test sets with a wide variety of instances are necessary to
evaluate the algorithmic performance. It is important that you check the robustness of your meta-heuristic with respect to different problem characteristics, such as number of items, capacity utilization, time between orders, demand variability and set up time which may affect the performance. Usually, the meta-heuristics are fine-tuned to provide good solutions. A characteristic of a good heuristic is that it remains robust with respect to the setting of the various control parameters. This also has to be tested. Second, there are several ways of comparing the solution quality and speed. A comparison can be done either with other heuristic approaches or with an optimal algorithm. This latter approach is usually only an option for small problems. However, the conclusions should be interpreted with caution: heuristic behavior on small instances does not necessarily extend to larger instances. When comparing several approaches, you can use previously published heuristics or develop several different new heuristics. Also, lower bounds can be used to evaluate the solution quality. However, good lower bounds are usually difficult to obtain. When working with published benchmark libraries, the best known solutions are available for comparison. When no other results are available from the literature, best known solutions could also be obtained with the meta-heuristic itself by either using different starting points or allowing for long runs.

As an illustration of these guidelines, we finally discuss some papers which have performed comprehensive computational experiments. However, this is not an exhaustive list of papers with elaborate computational experiments. Gopalakrishnan et al. (2001) developed a TS heuristic for the CLSP with set up carryover. For the problem with set up carryover, they developed a lower bound to estimate the optimality gap of their TS solutions. They also tested their TS heuristic on the CLSP without set up carryover and compared their results to the solution from another heuristic (Trigeiro et al. 1989). Algorithms can indeed be tested on several data sets: if you consider sequence dependent set up times, you can test your algorithm on a data test set without the sequence dependency or an algorithm for problems with set up times could be tested on problems without set up times as well. Likewise, Kimms (1999) compares a GA for the multi-machine PLSP also with a TS algorithm for the single machine problem using previously published data sets. A comparison is made with optimal solutions and he also tests the effect of problem parameters such as number of machines, demand pattern, cost structure and capacity utilization and algorithm parameters such as the population size. Özdamar, Birbil and Portmann (2002) compare their results on the data set consisting of 90 four-item, six-period problems used in Özdamar and Bozyel (2000) and extend their results with another test on 90 fourteen-item, six-period problems. They compare
their solutions to the optimal solution on the small data set and the best solution found among several algorithms for the larger data set. Hung et al. (2003) compare several versions of their Tabu Search to a LP-based heuristic (Hung and Hu 1998) and to B&B optimization. The gaps are calculated and compared to the best solution found among all the algorithms. They also investigate the effect of several problem parameters such as number of products, time period, set up cost, set up time and demand-to-capacity ratio. In order to have a fair basis for the comparison, they introduce a time limit.

6. Conclusions

The numerous extensions of the basic lot sizing problem demonstrate that it can be used to model a variety of industrial problems. Lot sizing problems are challenging because many extensions are very difficult to solve. We have discussed several techniques to tighten the formulations and obtain good quality solutions. The development of algorithms based on the combination of these techniques has already led to promising results. Many more opportunities for combining algorithms are still largely unexplored. The contribution of a new heuristic should be supported by extensive computational experiments using large and standard data set, making a comparison with other algorithms or lower bounds and testing insights into the effect of the algorithm parameters and problem structure.

From the literature review, it is clear the meta-heuristics have been applied to a wide variety of lot sizing models. Yet it seems that there is an absence of research on meta-heuristics for the regular DLSP. This is odd as a binary encoding is natural for this problem and this could be exploited by such algorithms as the production quantities can be calculated without having to solve an LP. Also, meta-heuristics should be applied to real-life problems and can then be compared to current practices. Until now, meta-heuristics are usually tested on artificial data sets only. Yet the power of these algorithms is exactly the ability to deal with complex, industrial considerations.

Finally, the interaction between modeling and algorithms will play an important role in future research. Many extensions of the general lot sizing problems have been proposed and the research expands in two opposite directions. The first line of research focuses on modeling the operational aspects in more detail, such as the sequence dependencies, different kinds of
set ups and inventory aspects. The second direction is towards more tactical and strategic models such as integrated production-distribution planning and supplier selection. Solution approaches for such integrated models will be based on previous research in the separate areas. Decomposition methods such as Dantzig-Wolfe decomposition and Lagrange relaxation seem a natural way to coordinate decisions by separating the different stages. Existing knowledge about the structure and properties of a specific subproblem can be exploited in solving integrated models. Additionally, meta-heuristics will play an important role in solving such complex problems, either as stand-alone algorithms, but more likely as part of a hybrid optimization model.
References


Diaby, M., Bahl, H.C., Karwan, M.H., Zionts, S., 1992b. A Lagrangean Relaxation Approach


Fleischmann, B., 1994. The Discrete Lot-Sizing and Scheduling Problem with Sequence-
Spektrum* 19, 11-21.
Florian, M., Lenstra, J.K., Rinnooy Kan, A.H.G., 1980. Deterministic Production Planning:
Florian, M., Klein, M., 1971. Deterministic Production Planning with Concave Costs and
Programming Study* 2, 82-113.
for the Capacitated Lot-Sizing Problem with Set-Up Carryover. *Management Science* 47
(6), 851-863.
Inventory Management* 20, 47-53.
Harrison, T.P., Lewis, H.S., 1996. Lot Sizing in Serial Assembly Systems with Multiple
Hindi, K.S., 1995. Solving the single-item, capacitated dynamic lot sizing problem with
startup and reservation costs by tabu search. *Computers and Industrial Engineering* 28 (4),
701-707.
Research Society* 47, 151-161.
candidate list to solve production planning problems with setups. *Computers & Industrial
Engineering* 45, 615-634.
of the Operational Research Society* 51, 1309-1318.


Publications in the Report Series Research* in Management

ERIM Research Program: “Business Processes, Logistics and Information Systems”

2004

Smart Pricing: Linking Pricing Decisions with Operational Insights
Moritz Fleischmann, Joseph M. Hall and David F. Pyke
ERS-2004-001-LIS
http://hdl.handle.net/1765/1114

Mobile operators as banks or vice-versa? and: the challenges of Mobile channels for banks
L-F Pau
ERS-2004-015-LIS
http://hdl.handle.net/1765/1163

Simulation-based solution of stochastic mathematical programs with complementarity constraints: Sample-path analysis
S. Ilker Birbil, Gül Gürkan and Ovidiu Listeş
ERS-2004-016-LIS
http://hdl.handle.net/1765/1164

Combining economic and social goals in the design of production systems by using ergonomics standards
Jan Dul, Henk de Vries, Sandra Verschoof, Wietske Eveleens and Albert Feilzer
ERS-2004-020-LIS
http://hdl.handle.net/1765/1200

A Review Of Design And Control Of Automated Guided Vehicle Systems
Tuan Le-Anh and M.B.M. De Koster
ERS-2004-030-LIS
http://hdl.handle.net/1765/1323

Online Dispatching Rules For Vehicle-Based Internal Transport Systems
Tuan Le-Anh and M.B.M. De Koster
ERS-2004-031-LIS
http://hdl.handle.net/1765/1324

Generalized Fractional Programming With User Interaction
S.I. Birbil, J.B.G. Frenk and S. Zhang
ERS-2004-033-LIS
http://hdl.handle.net/1765/1325

Meta-heuristics for dynamic lot sizing: A review and comparison of solution approaches
Raf Jans and Zeger Degraeve
ERS-2004-042-LIS

* A complete overview of the ERIM Report Series Research in Management:
https://ep.eur.nl/handle/1765/1

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship