Polyhedral Techniques in Combinatorial Optimization I: Theory

Karen Aardal

Department of Computer Science
Utrecht University
P.O. Box 80089
3508 TB Utrecht, The Netherlands
aardal@cs.ruu.nl

Stan van Hoesel

Department of Quantitative Economics
Limburg University
P.O. Box 616
6200 MD Maastricht, The Netherlands
s.vanhoesel@ke.rulimburg.nl

Abstract

Combinatorial optimization problems appear in many disciplines ranging from management and logistics to mathematics, physics, and chemistry. These problems are usually relatively easy to formulate mathematically, but most of them are computationally hard due to the restriction that a subset of the variables have to take integral values. During the last two decades there has been a remarkable progress in techniques based on the polyhedral description of combinatorial problems, leading to a large increase in the size of several problem types that can be solved. The basic idea behind polyhedral techniques is to derive a good linear formulation of the set of solutions by identifying linear inequalities that can be proved to be necessary in the description of the convex hull of feasible solutions. Ideally we can then solve the problem as a linear programming problem, which can be done efficiently. The purpose of this manuscript is to give an overview of the developments in polyhedral theory, starting with the pioneering work by Dantzig, Fulkerson and Johnson on the traveling salesman problem, and by Gomory on integer programming. We also present some modern applications, and computational experience.

Combinatorial optimization deals with maximizing or minimizing a function subject to a set of constraints and subject to the restriction that some, or all, variables should be integers. A well-known combinatorial optimization problem is the *traveling salesman problem*, where we want to determine in which order a "salesman" has to visit a number of "cities" such that all cities are visited exactly once, and such that the length of the tour is minimal. This problem is one of the most studied combinatorial optimization problems because of its numerous applications, both in its own right and as a substructure of more complex models, and because it is notoriously difficult to solve.

The computational intractability of most core combinatorial optimization problems has been theoretically indicated, i.e. it is possible to show that most of these problems belong to the class of *NP-hard problems*, see Karp (1972), and Garey and Johnson (1979). No algorithm with a worst-case running time bounded by a polynomial in the size of the input is known for any NP-hard problem, and it is strongly believed that no such algorithm exists. Therefore, to solve these problems we have to use an *enumerative algorithm*, such as dynamic programming or branch and bound, with a worst-case running time that is exponential in the size of the input. The computational hardness of most combinatorial optimization problems has inspired researchers to develop good formulations and algorithms that are expected to reduce the size of the enumeration tree. To use information about the structure of the convex hull of feasible solutions, which is the basis for polyhedral techniques, has been one of the most successful approaches so far. The pioneering work in this direction was done by Dantzig, Fulkerson and Johnson (1954), who invented a method to solve the traveling salesman problem. They demonstrated the power of their technique on a 49-city instance, which was huge at that time.

The idea behind the Dantzig-Fulkerson-Johnson method is the following. Assume we want to solve the problem

$$\min\{cx \ \text{ subject to } x \in S\}, \tag{1}$$

where S is the set of feasible solutions, which in this case is the set of traveling salesman tours. Let $S = P \cap \mathbb{Z}^n$, where $P = \{x \in \mathbb{R}^n : Ax \le b\}$ and where $Ax \le b$ is a system of linear inequalities. Since S is difficult to characterize, we could solve the problem

$$\min\{cx \ \text{ subject to } x \in P\} \tag{2}$$

instead. Problem (2) is easy to solve as linear programming problems are known to be polynomially solvable, but since it is a relaxation of (1) it may give us a solution x^* that is not a tour. More precisely, the following two things can happen if we solve (2): either the optimal solution x^* is a tour, which means that x^* is also optimal for (1), or x^* is not a tour, in which case it is not feasible for (1). If the solution x^* is not feasible for (1) it lies outside the convex hull of S which means we can cut off x^* by identifying a hyperplane separating x^* from the convex hull of S, i.e. a hyperplane that is satisfied by all tours, but violated by x^*. An inequality that is satisfied by all feasible solutions is called a *valid inequality*. When Dantzig, Fulkerson and Johnson

1

solved the relaxation (2) of their 49-city instance they indeed obtained a solution x^* that was not a tour. By looking at the solution they identified a valid inequality that was violated by x^*, and added this inequality to the formulation. They solved the resulting linear programming problem and obtained again a solution that was not a tour. After repeating this process a few times a tour was obtained, and since only valid inequalities were added to the relaxation, they could conclude that the solution was optimal.

Even though many theoretical questions regarding the traveling salesman problem remained unsolved, the work of Dantzig, Fulkerson and Johnson was still a break-through as it provided a methodology that was actually not limited to solving traveling salesman problems, but could be applied to any combinatorial optimization problem. This new area of research on how to describe the convex hull of feasible solutions by linear inequalities was called *polyhedral combinatorics*. During the last decades polyhedral techniques have been used with considerable success to solve many previously unsolved instances of hard combinatorial optimization problems, and it is still the only method available for solving large instances of the traveling salesman problem. The purpose of this paper is to describe the basic theoretical aspects of polyhedral techniques and to indicate the computational potential.

A natural question that arises when studying the work by Dantzig, Fulkerson and Johnson is whether it is possible to develop an algorithm for identifying valid inequalities. This question was answered by Gomory (1958), (1960), (1963) who developed a *cutting plane algorithm* for general integer linear programming, and showed that the integer programming problem (1) can be solved by solving a finite sequence of linear programs. Chvátal (1973) proved that all inequalities necessary to describe the convex hull of integer solutions can be obtained by taking linear combinations of the original and previously generated linear inequalities and then applying a certain rounding scheme, provided that the integer solutions are bounded. Schrijver (1980) proved the more general result that it is possible to generate the convex hull of integer solutions by applying a finite number of operations to the linear formulation containing the integer solutions, starting with P, if P is rational but not necessarily bounded. The results by Gomory, Chvátal, and Schrijver are discussed in Section 1. Here we will also address the following two questions: When can we expect to have a concise description of the convex hull of feasible solutions? How difficult is it to identify a violated inequality? These questions are strongly related to the computational complexity of the considered problem, i.e. the hardness of a problem type will catch up with us at some point, but we shall also see that certain aspects of the answers make it possible to hope that a bad situation can be turned into a rather promising one.

The results of Gomory, Chvátal and Schrijver were very important theoretically, but they did not provide tools for solving realistic instances within reasonable time. Researchers therefore began to develop problem specific classes of inequalities that contain inequalities that can be proved to be necessary in the description of the convex hull of feasible solutions. Based on the various classes of valid inequalities it is then necessary to develop *separation algorithms*, i.e. algorithms for identifying vio-

lated inequalities given the current solution x^*. In Section 2 we begin by describing families of valid inequalities and the corresponding separation problems for two basic combinatorial optimization problems. These inequalities are important as they are often useful when solving more complex problems as well, either directly, or as a starting point for developing new, more general families of inequalities. Moreover, they represent different arguments that can be used when developing valid inequalities. We then discuss two applications: the capacitated facility location problem, and the economic lot-sizing problem.

Next to the theoretical work of developing good classes of valid inequalities and algorithms for identifying violated inequalities, there is a whole range of computational issues that have to be considered in order to make polyhedral methods work well. These issues, together with some alternative approaches to solving integer and combinatorial optimization problems, and an extensive list of problems for which polyhedral results are known, will be discussed in the accompanying Part II of this paper.

Research carried out in the Netherlands involves both theoretical and more problem specific results. Gerards and Schrijver have considered several important theoretical issues, see e.g. Schrijver (1980,1981), Grötschel, Lovász and Schrijver (1981), Cook, Gerards, Schrijver and Tardos (1986). Here we also want to mention the result of H.W. Lenstra (1983) that the integer programming problem (1) can be solved in polynomial time for a fixed number of variables. Although not specifically a result in polyhedral combinatorics, it is central in integer programming and combinatorial optimization. If we consider more problem specific results, *lot-sizing problems* have been considered by Van Eijl, Van Hoesel, Kolen and Wagelmans, see Van Hoesel and Kolen (1994), Van Hoesel, Wagelmans, and Wolsey (1994), Van Eijl and Van Hoesel (1995), Van Hoesel, Kolen and Van Eijl (1995). Gerards and Schrijver characterize graphs for which the *node packing polytope* is described completely by certain constraints, see Gerards and Schrijver (1986), Gerards (1989), and Gerards and Shepherd (1995). Inequalities for the node packing problem have been used to solve problems such as the *radio link frequency assignment problems* by Aardal, Hipolito, Van Hoesel and Jansen (1995), and the *uncapacitated facility location problem* by Aardal and Van Hoesel (1995). Various more complex *facility location problems* have been studied by Aardal, see Aardal, Pochet and Wolsey (1993), Aardal (1994), and Aardal, Labbé, Leung and Queyranne (1994). Results on *scheduling problems* have been obtained by Nemhauser and Savelsbergh (1992), Crama and Spieksma (1995), Van den Akker, Van Hoesel and Savelsbergh (1993), and Van den Akker, Hurkens and Savelsbergh (1995). Savelsbergh has also considered the single-node fixed charge flow model (Gu, Nemhauser and Savelsbergh (1995)), and he is one of the researchers behind the commercial mixed-integer programming software MINTO, see Savelsbergh, Sigismondi and Nemhauser (1994).

1   Theoretical background

The integer linear programming problem (ILP) is defined as

$$\min\{cx : x \in S\},$$

where $S = P \cap \mathbb{Z}^n$ and $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. We call P the *linear formulation* of ILP. A polyhedron P is *rational* if it can be determined by a rational system $Ax \leq b$ of linear inequalities, i.e., a system of inequalities where all entries of A and b are rationals. The convex hull of the set S of feasible solutions, denoted $\text{conv}(S)$, is the smallest convex set containing S. A facet-defining valid inequality is a valid inequality that is necessary to describe $\text{conv}(S)$, i.e. it is the "strongest possible" valid inequality. In Figure 1 we give an example of sets P, S and $\text{conv}(S)$.
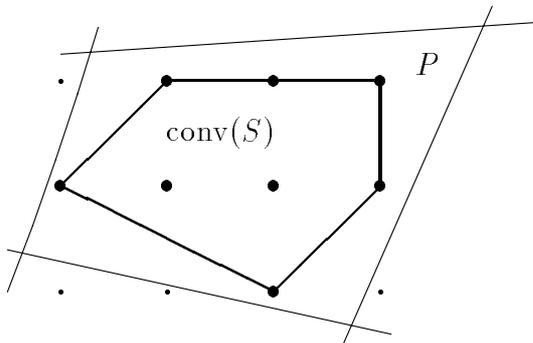

Figure 1: P, S, and $\text{conv}(S)$.

If we know the linear description of $\text{conv}(S)$ we can solve the linear programming problem $\min\{cx : x \in \text{conv}(S)\}$ which is computationally easy. In this section we shall primarily address the issue of how difficult it is to obtain $\text{conv}(S)$. First we show that for rational polyhedra, and for not necessarily rational, but bounded, polyhedra, we can generate $\text{conv}(S)$ algorithmically in a finite number of steps. In general however, there is no upper bound on the number of steps in terms of the dimension of S. We also demonstrate that it is very unlikely that $\text{conv}(S)$ of any NP-hard problem can be described by concise families of linear inequalities. Finally, we relate the complexity of the problem of finding a hyperplane separating a vector x^* from $\text{conv}(S)$ or showing that x^* belongs to $\text{conv}(S)$, to the complexity of optimizing over S. In general these two problems are equally hard, but if we restrict the search of a separating hyperplane to a specific class, this problem might be polynomially solvable even if the underlying optimization problem is NP-hard.

1.1   Solving Integer Programming Problems by Linear Programming

Gomory's Cutting Plane Algorithm. What was needed to transform the procedure of Dantzig, Fulkerson and Johnson (1954) into an algorithm was a systematic

procedure for generating valid inequalities that are violated by the current solution. Assume that we want to solve the variant of ILP where the integer vectors in S are bounded and where all entries of the constraint matrix A and the right-hand side vector b are integers. Gomory (1958), (1960) and (1963) developed a *cutting plane algorithm* based on the simplex method, for solving integer linear problems on this form. This was the first algorithm developed for integer linear programming that could be proved to terminate in a finite number of iterations. The basic idea of Gomory's algorithm is similar to the approach of Dantzig, Fulkerson and Johnson, i.e. instead of solving ILP directly we solve the linear programming (LP) relaxation (2) by the simplex method. If the optimal solution to LP is integral, then we are done, and otherwise we need to identify a valid inequality cutting off x^*. Gomory developed a technique for automatically identifying a violated valid inequality and proved that after adding a finite number of inequalities, called *Gomory cutting planes*, the optimal solution is obtained. We shall illustrate Gomory's technique by an example. Assume we have solved the LP-relaxation (2) of an instance of ILP by the simplex method, and that one of the rows of the final tableau reads

$$x_1 - \tfrac{1}{11}x_3 + \tfrac{2}{11}x_4 = \tfrac{36}{11},$$

where x_1 is a basic variable and variables x_3 and x_4 are non-basic, i.e. at the current solution $x_1 = 36/11$ and $x_3 = x_4 = 0$. We now split each coefficient in an integer and a fractional part by rounding down all coefficients. The integer terms are put in the left-hand side of the equation and the fractional terms are put in the right-hand side. Since all coefficients are rounded down, the fractional part of the variable coefficients in the right-hand side becomes nonpositive, giving

$$x_1 - x_3 - 3 = -\tfrac{10}{11}x_3 - \tfrac{2}{11}x_4 + \tfrac{3}{11}.$$

In any feasible solution to ILP, the left-hand side should be integral. Moreover, all variables are nonnegative. Since the variables in the right-hand side appear with nonpositive coefficients we can conclude that

$$\tfrac{3}{11} - \tfrac{10}{11}x_3 - \tfrac{2}{11}x_4 \le 0 \text{ , and integer.} \tag{3}$$

We have argued that inequality (3) is valid, i.e. it is not violated by any feasible integer solution. It is easy however to see that it does cut off the current fractional solution as $x_3 = x_4 = 0$. Let $\lfloor x \rfloor$ denote the integer part of x.

Outline of Gomory's cutting plane algorithm.

1. Solve the linear relaxation (2) of ILP with the simplex method. The current number of variables is k. If the optimal solution x^* is integral, stop.

2. Choose a source row i_0 in the optimal tableau with a fractional basic variable. Row i_0 reads $\bar{a}_{i_0,1}x_1 + \bar{a}_{i_0,2}x_2 + \ldots + \bar{a}_{i_0,k}x_k = \bar{b}_{i_0}$. Let $a'_{ij} = \bar{a}_{ij} - \lfloor \bar{a}_{ij} \rfloor$, and $b'_i = \bar{b}_i - \lfloor \bar{b}_i \rfloor$.

3. Add the equation $-a'_{i_0,1}x_1 - a'_{i_0,2}x_2 - \ldots - a'_{i_0,k}x_k + x_{k+1} = -b'_{i_0}$, where x_{k+1} is a slack variable, to the current linear formulation, and reoptimize. If the optimal solution x^* is integral, stop, otherwise $k \leftarrow k + 1$, go to 2.

In the outline above we have not specified how to choose the source row. To be able to prove that the algorithm terminates in a finite number of steps we have to make sure that certain technical conditions are satisfied. The technical details are omitted here but can be found in Gomory (1963) who gives two proofs of finiteness, and in Schrijver (1986), page 357.

Theorem 1 Gomory (1963). *There exists an implementation of Gomory's cutting plane algorithm such that after a finite number of iterations either an optimal integer solution is found, or it is proved that $S = \emptyset$.*

A recent discussion on Gomory cutting planes can be found in Balas et al. (1994) who incorporate the cutting plane algorithm in a branch-and-bound procedure and report on computational experience.

Chvátal's Rounding Procedure. Chvátal (1973) studied the more general version of ILP, where the integer vectors of S are bounded and where the entries of A and b are real numbers. He showed that if one takes linear combinations of the linear inequalities defining P and then applies rounding, and repeats the procedure a finite number of times, $\text{conv}(S)$ is obtained. After each iteration of the procedure we get a new linear formulation containing more inequalities. We again illustrate the procedure by an example. Note that this example will be referred to frequently in the sequel. Let $G = (V, E)$ be an undirected graph where V is the set of vertices and E is the set of edges. A *matching* M in a graph is a subset of edges such that each vertex is incident to at most one edge in M, see Figure 2. In the figure thick lines represent edges belonging to the matching. An early application of matching appears in so-called sets of distinct representatives, where a family of sets is given. From each set one element should be chosen, such that all chosen elements are different. The matching problem is also a substructure in many school timetabling problems. For more technical details, see Gerards (1995).
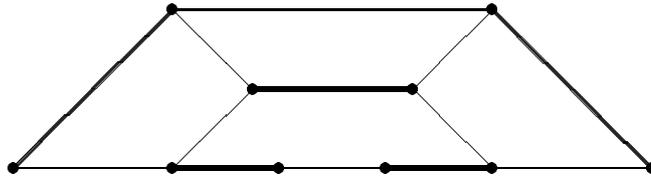

Figure 2: A maximum matching.

Let x_e be equal to one if edge e belongs to the matching M and zero otherwise, and let $\delta(v) = \{e \in E : e \text{ is incident to } v\}$. The maximum cardinality matching problem can be formulated as the following linear integer programming problem.

$$\max \sum_{e \in E} x_e \tag{4}$$

$$\text{s.t.} \quad \sum_{e \in \delta(v)} x_e \leq 1 \qquad \text{for all } v \in V, \tag{5}$$

$$0 \leq x_e \leq 1 \qquad \text{for all } e \in E, \tag{6}$$

6

$$x_e \text{ integer} \qquad\qquad \text{for all } e \in E. \qquad (7)$$

Let U be any subset consisting of k vertices, where $k \geq 3$ and odd, and let $E(U)$ be the set of edges with both endvertices in U. By adding inequalities (5) for all $v \in U$ we obtain $2 \sum_{e \in E(U)} x_e \leq |U|$, or equivalently

$$\sum_{e \in E(U)} x_e \leq \frac{|U|}{2}. \qquad (8)$$

Since each x_e is an integer, the left-hand side of (8) has to be integral. As $|U|$ is odd, the right-hand side of (8) is fractional, and hence we can round down the right-hand side of (8) giving the valid inequality

$$\sum_{e \in E(U)} x_e \leq \left\lfloor \frac{|U|}{2} \right\rfloor, \qquad (9)$$

which we call an *odd-set constraint*. It is easy to show that the odd-set constraints are necessary to describe the convex hull of matchings in G. We also note that there are exponentially many odd-set constraints as there are exponentially many ways of forming subsets U. We shall now give a more formal description of Chvátal's procedure.

An inequality $\sum_{j=1}^{n} a_j x_j \leq b$ is said to belong to the *elementary closure* of a set P of linear inequalities, denoted $e^1(P)$, if there are inequalities $\sum_{j=1}^{n} a_{ij} x_j \leq b_i$, $i = 1, \ldots, m$ defining P, and nonnegative real numbers $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that

$$\sum_{i=1}^{m} \lambda_i a_{ij} = a_j \quad \text{with } a_j \text{ integer}, \quad j = 1, \ldots, n, \quad \text{and} \quad \left\lfloor \sum_{i=1}^{m} \lambda_i b_i \right\rfloor \leq b.$$

For integer values of $k > 1$, $e^k(P)$ is defined recursively as $e^k(P) = e(P \cup e^{k-1}(P))$. The *closure* of P is defined as $c(P) = \cup_{k=1}^{\infty} e^k(P)$.

Theorem 2 Chvátal (1973). *If S is a bounded polyhedron, then $conv(S)$ can be obtained after a finite number, k, of closure operations.*

An interesting question is if k can be bounded from above by a function of the dimension of S. Chvátal called the minimum number of closure operations required to obtain $conv(S)$, given a linear formulation P, the *rank* of P. If we return to the matching problem (5)–(7), it was proved by Edmonds (1965) that the convex hull of the matching polytope is determined by inequalities (5), (6) and (9). As the odd-set constraints (9) can be obtained by applying one closure operation on the linear formulation, the rank of the set of inequalities (5) and (6) is one. In general however, there is no upper bound on k in terms of the dimension of P as the two-dimensional polytope $P = \{x \in \mathbb{R}_+^2 : x_1 + 2tx_2 \leq 2t, \ x_1 - 2tx_2 \leq 0\}$ illustrates. One closure operation reduces the polytope only slightly, i.e., the point $(t-1, \frac{1}{2})$ belongs to $e^1(P)$, and similarly, the point $(t-k, \frac{1}{2})$ belongs to $e^k(P)$ for $k < t$. Only if $S = \emptyset$ does there exists an upper bound on k that is a function of the dimension of P. This was proved by Cook et al. (1987).
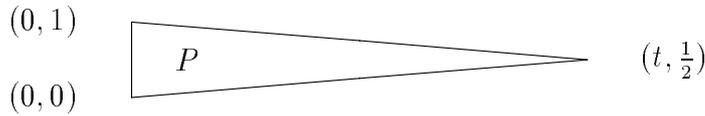
7

$(0, 1)$

$(0, 0)$

P

$(t, \frac{1}{2})$

Figure 3: The number of closure operations is of order t.

There is a clear relation between Chvátal's closure operations and Gomory's cutting planes in the sense that every Gomory cutting plane can be obtained by a series of closure operations, and every inequality belonging to the elementary closure can be obtained as a Gomory cutting plane. It would be possible to prove Theorem 2 using Gomory's algorithm, but then one would first need to get rid of the inequalities $x_j \geq 0$, $j, \ldots, n$, and the assumption that the entries of A and b have to be integer. For further details, see Chvátal (1973).

Schrijver's Rounding Procedure. Schrijver (1980) studied the version of ILP where S is not necessarily bounded, and where P is defined by a rational system of linear inequalities. The operations carried out on P to obtain the convex hull of feasible solutions is quite different from the linear combination and rounding schemes developed by Gomory and Chvátal. The key component of Schrijver's procedure is the formulation of a *totally dual integral* (TDI) system of inequalities. A rational system $Ax \leq b$ of linear inequalities is TDI if for all integer vectors c such that $\max\{cx : Ax \leq b\}$ is finite, the dual, $\min\{yb : yA = c, \; y \geq 0\}$, has an integer optimal solution. Note that if $Ax \leq b$ is TDI, and if b is integral, then $P = \{x : Ax \leq b\}$ is an integral polyhedron, i.e. all extreme points of P are integral. TDI systems were introduced by Edmonds and Giles (1977).

Each iteration of Schrijver's procedure consists of the following two steps.

1. Given a rational polyhedron P, find a TDI system $Ax \leq b$ defining P, with A integral.

2. Round down the right-hand side b.

It has been proved by Giles and Pulleyblank (1979) and Schrijver (1981) that there exists a TDI system as in step 1 of Schrijver's procedure for every rational polyhedron P, and that the TDI system is unique if P is full-dimensional. Finding such a TDI system can be done in finite time. After one iteration of the above procedure we get a polyhedron $P^{(1)}$ strictly contained in P unless P is integral. Given the polyhedron $P^{(1)}$ we repeat the steps 1 and 2. This continues until $\mathrm{conv}(S)$ is obtained.

Theorem 3 Schrijver (1980). *For each rational polyhedron P, there exists a natural number k, such that after k iterations of Schrijver's procedure $\mathrm{conv}(S)$ is obtained.*

The results presented above are of significant theoretical importance as they give algorithmic ways of generating the convex hull of feasible solutions. All three approaches are finite, but from a practical point of view finite in most cases does not imply that computations can be done within reasonable time. One apparent question is whether for some problem classes it is possible to write down the linear description

8

of the convex hull in terms of concise families of linear inequalities. If that is possible we could apply linear programming directly. This is the topic of the following subsection.

1.2 Concise Linear Descriptions

We mentioned in the previous subsection that the convex hull of matchings in a general undirected graph G is given by the defining inequalities (5), (6), and the exponential class of inequalities (9). Assume now that G is *bipartite*, i.e. that we can partition the set V of vertices into two sets V_1, V_2, such that all edges have one endvertex in V_1 and the other endvertex in V_2. For bipartite graphs the convex hull of matchings is described by the defining inequalities (5) and (6) only, which is a polynomial system of linear inequalities. This means that for bipartite graphs the integrality condition (7) is redundant. In contrast, there is no concise linear description known for the traveling salesman problem, even if we allow for exponential families of inequalities. The reason why the bipartite matching problem is so easy is that the constraint matrix is *totally unimodular* (TU). A matrix A is TU if each subdeterminant of A is equal to 0,1 or -1.

Theorem 4 *If A is a TU matrix the polyhedron $P = \{x : Ax \leq b\}$ is integral for all integer vectors b for which P is not empty.*

Seymour (1980) provided a complete characterization of TU matrices yielding a polynomial algorithm for testing whether a matrix is TU. For a thorough discussion on TU matrices we refer to Schrijver (1986), and Nemhauser and Wolsey (1988).

It is interesting to observe here that the bipartite matching problem is polynomially solvable as its linear description is polynomial in the dimension of the problem. For the matching problem in general undirected graphs there is a polynomial combinatorial algorithm due to Edmonds (1965), but the traveling salesman problem is known to be NP-hard. The following theorem confirms that there is a natural link between the computational complexity of a class of problems and the possibility of providing concise linear descriptions of the convex hull of feasible solutions. Before stating the result we need to introduce the following decision problems:

The *lower-bound feasibility problem*. An instance is given by integers m, n, an $m \times n$ matrix A, vectors b and c, and a scalar δ. The question is: $\exists\, x \in \mathbb{Z}^n : Ax \leq b, \; cx > \delta$?

The *facet validity problem*. An instance is given by the same input as for the lower-bound feasibility problem. The question is: Does $cx \leq \delta$ define a facet of $\mathrm{conv}(\{x \in \mathbb{Z}^n : Ax \leq b\})$?

Note that if the lower-bound feasibility problem for a family of polyhedra is NP-complete then optimizing over the same family of polyhedra is NP-hard.

Lemma 5 *If any NP-complete problem belongs to co-NP, then NP=co-NP.*

Theorem 6 Karp and Papadimitriou (1980). *If lower-bound feasibility is NP-complete, and facet validity belongs to NP, then NP=co-NP.*

The way to prove Theorem 6 is to show that if facet validity belongs to NP, then lower-bound feasibility belongs to co-NP. If lower-bound feasibility is NP-complete we can through Lemma 5 conclude that NP=co-NP. It is extremely unlikely that NP=co-NP, as this implies that all NP-complete problems have a compact certificate for the no-answer. Hence, if we believe that NP$\neq$co-NP, and if $\min\{cx : x \in S\}$ is NP-hard, then there are classes of facets of conv(S) for which there is no short proof that they are facets.

1.3   Equivalence Between Optimization and Separation

We have seen that if a problem is NP-hard we cannot expect to have a concise linear description of the convex hull of feasible solutions. Moreover, for the matching problem, which is polynomially solvable and which has a concise linear description of the convex hull of feasible solutions, this description is exponential in the dimension of the problem. These observations do not necessarily have to be negative since what we primarily need is a good description of the area around the optimal solution. The question is then whether there exists an efficient way to identify a violated inequality whenever needed, i.e. if we can find, in polynomial time, a hyperplane separating a given fractional solution from the convex hull, or prove that no such hyperplane exists.

The *separation problem for a family FP of polyhedra.* Given a polyhedron $P \in FP$, and a solution x^*, find an inequality $cx \leq \delta$, valid for P, satisfying $cx^* > \delta$, or prove that $x^* \in P$.

The *optimization problem for a family FP of polyhedra.* Given is a polyhedron $P \in FP$. Assume that $P \neq \emptyset$ and that P is bounded. Given a vector $c \in \mathbb{R}^n$, find a solution x^0 such that $cx^0 \leq cx$ for all $x \in P$.

Theorem 7 Grötschel, Lovász and Schrijver (1981). *There exists a polynomial time algorithm for the separation problem for a family FP of polyhedra, if and only if there exists a polynomial time algorithm for the optimization problem for FP.*

The theorem says that separation in general is equally hard as optimization but, as we shall see in the next section, when applying the polyhedral approach we develop *specific* families of valid inequalities for a given problem type, such as the odd-set constraints (9) developed for the matching problem.

The *separation problem based on a family FI of valid inequalities.* Given a solution x^*, find an inequality $cx \leq \delta$ belonging to FI, satisfying $cx^* > \delta$, or prove that no such inequality in FI exists.

The separation problem based on a family of valid inequalities may be polynomially solvable even if the underlying optimization problem is NP-hard. Moreover, even if a family of inequalities is NP-hard to separate we may still be able to separate it effectively using a heuristic. Good separation heuristics together with a good implementation of a preprocessing routine and a branch-and-bound scheme, form the

basis for the success of the polyhedral approach.

2 Polyhedral Results for Selected Combinatorial Structures

The results presented in the previous section did provide very important theoretical answers, but no efficient computational tools. In the early seventies there was a renewed interest in developing general purpose integer programming solvers. Instead of Gomory's cutting plane method, which tended to be very time consuming, one developed facet defining inequalities and corresponding separation algorithms for various problem types, and embedded the separation algorithms in a branch-and-bound framework. In the early days one generated violated inequalities only in the root node of the branch-and-bound tree, whereas in modern implementations inequalities may be generated in every node. Since the added inequalities could be proved to be necessary to describe the convex hull of feasible solutions one could expect that they would be more effective than the Gomory cutting planes. Moreover, developing facet defining inequalities and associated separation algorithms for some basic combinatorial structures that occur frequently in more general combinatorial optimization problems, would possibly be very useful when solving a wide range of problems. In the late seventies and in the eighties remarkable computational progress was made. Here we shall describe some classes of facet defining valid inequalities developed for a few basic, important, combinatorial optimization problems. The main purpose is to give an impression of how inequalities and separation algorithms are developed, and how they can be used, not only for the problem for which they are developed, but also for more general structures. Since the space provided here is not enough for a complete survey, we recommend the following literature to the interested reader. The books by Schrijver (1986), and Nemhauser and Wolsey (1988) provide a broad theoretical foundation as well as many examples. The latest developments on solving large traveling salesman problems are reported by Applegate et al. (1994). The article by Jünger et al. (1995) contains a comprehensive survey of computational results obtained by using polyhedral techniques. In Part II of this article we present an extensive list of different problem types for which polyhedral results are known, together with references.

Before we start the more technical exposition we give some basic definitions. An inequality $\pi x \leq \pi_0$ is called *valid* for P if each point in P satisfies the inequality. The set $F = \{x \in P : \pi x = \pi_0\}$ is called a *face* of P, and the valid inequality $\pi x \leq \pi_0$ is said to *define* the face F. A face is said to be *proper* if it is not empty and if it is properly contained in P, i.e. if $\emptyset \neq F \neq P$. The dimension of a proper face F, $\dim(F)$, is strictly smaller than the dimension of P. If $\dim(F) = \dim(P) - 1$, i.e., if the dimension of F is maximal, then F is called a *facet*. The facet defining inequalities are important since they are precisely the inequalities needed to define the convex hull of feasible solution in addition to the set of inequalities that are satisfied with equality by every feasible point.

2.1 The Traveling Salesman Problem

Consider a complete undirected graph $G = (V, E)$ with $n = |V|$. In the traveling salesman problem (TSP) we want to find a minimum length *Hamiltonian cycle*, i.e. a minimum length cycle containing each vertex exactly once. A modern application of the TSP occurs when manufacturing printed circuit boards. It is then necessary to drill numerous small holes for the wiring, which is done using a numerically controlled drilling machine. To speed up production it is desirable to find the drilling sequence that gives the shortest cycle. Another application occurs in vehicle routing. Here we need to simultaneously decide on the number of vehicles needed to serve a set of clients, as well as the tour, from a central depot, to a subset of the clients, and back to the depot, that each vehicle has to make.

Let $x_e = 1$ if edge e is belongs to the Hamiltonian cycle, and let $x_e = 0$ otherwise. Moreover, let d_e denote the length of edge $e \in E$. As before, $E(S)$ denotes the set of edges with both endvertices in S. Often, the vertices of the graph are called *cities*, and the Hamiltonian cycle is called a *tour*.

$$\min \sum_{e \in E} d_e x_e \tag{10}$$

$$\text{s.t.} \qquad \sum_{e : v \in e} x_e = 2 \qquad \text{for all } v \in V, \tag{11}$$

$$\sum_{e \subset E(S)} x_e \leq |S| - 1 \qquad \text{for all } \emptyset \subset S \subset V, \tag{12}$$

$$x_e \in \{0, 1\} \qquad \text{for all } e \in E. \tag{13}$$

The formulation restricted to the constraints (11) and (13) is called the *2-matching relaxation* of TSP and its solutions are referred to as *2-matchings*. Such solutions may constitute disjoint cycles, or *subtours*. Constraints (12), introduced by Dantzig et al. (1954), prevent subtours, and are therefore called *subtour elimination constraints*. Edmonds (1965) studied the polyhedral structure of the 2-matching polytope, and obtained a complete linear description of the convex hull of feasible solutions by adding so-called *2-matching inequalities* to the linear relaxation of the 2-matching polytope, i.e. to constraints (11) and $0 \leq x_e \leq 1$ for all $e \in E$. Since the 2-matching problem is a relaxation of TSP, the 2-matching inequalities are also valid for TSP. We illustrate these inequalities by considering a solution to the linear relaxation of the 2-matching polytope illustrated in Figure 4. The thick lines in the figure correspond to variables that have value 1 and the thin lines correspond to variables with value 0.5. The intuition behind the 2-matching inequality is as follows. From Figure 4 we see that the triangles induced by the vertices $\{1, 2, 3\}$ or $\{4, 5, 6\}$ would cause a violation of the degree constraints (11) if we impose integrality. Therefore, consider one of the triangles, say $H = \{1, 2, 3\}$. Let $E(H)$ be the set of edges with both endvertices in H, and let $E' = \{\{1, 4\}, \{2, 5\}, \{3, 6\}\}$, i.e each edge in E' has exactly one endvertex in H. Furthermore, let $x(F) = \sum_{e \in F} x_e$. From the set of edges $E(H) \cup E'$ at most four can belong to a 2-matching since otherwise at least one of the vertices in H will have degree 3, which violates constraints (11). The cumulative value of the
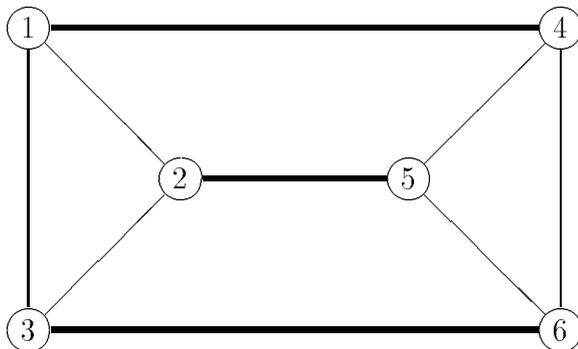
Figure 4: A fractional solution violating a 2-matching constraint.

variables corresponding to edges in $E(H) \cup E'$ is 4.5. Hence, we can conclude that the inequality $x(E(H)) + x(E') \leq 4$ is violated by the solution described above. In general, a 2-matching constraint has the form

$$x(E(H)) + x(E') \leq |H| + \left\lfloor \frac{1}{2}|E'| \right\rfloor, \tag{14}$$

where $H \subset V$ and where the edges in E' have precisely one endvertex in H. Note that only 2-matching constraints with an odd number of edges in E' can be facet-defining, since they are otherwise implied by the degree constraints.

Comb inequalities were introduced by Chvátal (1975) as a generalization of the 2-matching constraints. In the comb inequalities the edges in E' are replaced by an odd number, s, of disjoint vertex sets $T_1, \ldots, T_s$, called *teeth*, each having one vertex in common with the *handle H*. The comb inequality is written as

$$x(E(H)) + \sum_{j=1}^{s} x(E(T_j)) \leq |H| + \sum_{j=1}^{s} (|T_j| - 1) - \frac{1}{2}(s + 1). \tag{15}$$

Chvátal's comb inequalities were generalized by Grötschel and Padberg (1979) who introduced structures where each tooth can have more than one vertex in common with the handle. The *clique tree inequalities*, introduced by Grötschel and Pulleyblank (1986), are further generalization of comb inequalities in the sense that clique trees contain multiple handles, which are connected through the teeth. Many more exotic classes of inequalities have been derived to date, but the search for new classes is still vivid. A good overview of the current state-of-the-art is provided by Applegate et al. (1994). Goemans (1993) considers the quality of various classes of inequalities with respect to their induced relaxations.

The separation problem based on the subtour elimination constraints can be viewed as a minimum cut problem, which is polynomially solvable using max-flow algorithms. Separation of the 2-matching constraints is also polynomial, which was shown by Padberg and Grötschel (1985). Violated 2-matching constraints are however usually identified using a heuristic, since this is still effective and faster in practice. No polynomial time algorithm is known for solving the separation problem based on the comb inequalities, but there are fast heuristic methods available that perform

13

quite well. For clique tree inequalities, not even good heuristics are known that will perform well in general. To illustrate the progress made by using the polyhedral approach to solve the TSP, we present, in Table 1, the sizes of the largest instances that have been solved to optimality since 1954. Note that the values given in the column z_{LP}^{root} have been rounded to the nearest integer. In the tables to follow we use the following notation: z_{LP} denotes the value of the LP-relaxation, and z_{IP} denotes the optimal value of ILP. By % *gap* we mean the percentage duality gap, i.e. $(z_{IP} - z_{LP})/z_{IP}$. The percentage duality gap closed, denoted % *gap closed*, is calculated as $(z_{LP}^{\text{root}} - z_{LP})/(z_{IP} - z_{LP})$, where z_{LP}^{root} is the value of the LP-relaxation after all violated inequalities that have been identified in the root node of the branch-and-bound tree have been added. The number of branch-and-bound nodes needed to verify the optimal solution is given in the column *B&B nodes*.

cities	z_{LP}^{root}	z_{IP}	B&B nodes	application	year	reported by:
49	12,345	12,345	1	map USA	1954	Dantzig et al.
120	6,942	6,942	1	map Germany	1980	Grötschel
318	38,765	41,345	35	drilling	1980	Crowder & Padberg
532	27,628	27,686	85	map USA	1987	Padberg & Rinaldi
666	294,080	294,358	21	world map	1991	Grötschel & Holland
1,002	258,860	259,045	13	drilling	1990	Padberg & Rinaldi
2,392	378,027	378,032	3	drilling	1990	Padberg & Rinaldi
3,038	137,660	137,694	287	drilling	1992	Applegate et. al
4,461	182,528	182,566	2,092	map Germany	1994	Applegate et. al
7,397	23,253,123	23,260,728	2,247	programmable logic arrays	1994	Applegate et. al

Table 1: Computational results for the traveling salesman problem.

2.2   The Knapsack Problem

Consider a set $N = \{1, \ldots, n\}$ of items, each having a weight a_j, and a value c_j. A "knapsack" is to be filled with a subset of the items, such that the cumulative weight of the items does not exceed a given threshold, and such that the cumulative value is maximum. The knapsack polytope occurs as a substructure in many capacitated combinatorial optimization problems. An example is the *capacitated facility location problem* presented in Section 2.3. The knapsack problem is formulated as

$$\max \quad \sum_{j \in N} c_j x_j \tag{16}$$

$$\text{s.t.} \quad \sum_{j \in N} a_j x_j \leq b, \tag{17}$$

$$x_j \in \{0, 1\} \qquad \text{for all } j \in N. \tag{18}$$

Assume that the vectors c, a, and the right-hand side b are rational, and let X_K denote the set of feasible solutions to the knapsack problem. We call a set C a *cover*, or a *dependent set*, with respect to N if $\sum_{j \in C} a_j > b$. A cover is *minimal* if $\sum_{j \in S} a_j \leq b$

for all $S \subset C$. If we choose all elements from the cover C, it is clear that the right-hand side of (17) is exceeded. Hence, the following *knapsack cover inequality* (Balas (1975), Hammer et al. (1975) and Wolsey (1975))

$$\sum_{j \in C} x_j \leq |C| - 1 \tag{19}$$

is valid. A generalization of (19) is given by the family of $(1, k)$-*configuration inequalities* (Padberg (1980)). Let $\bar{C} \subseteq N$, and $t \in N \setminus \bar{C}$ be such that $\sum_{j \in \bar{C}} a_j \leq b$ and such that $Q \cup \{t\}$ is a minimal cover for all $Q \subseteq \bar{C}$ with $|Q| = k$. Let $T(r) \subseteq \bar{C}$ vary over all subsets of cardinality r of $\bar{C}$, where r is an integer satisfying $k \leq r \leq |\bar{C}|$. The $(1, k)$-configuration inequality

$$(r - k + 1)x_t + \sum_{j \in T(r)} x_j \leq r \tag{20}$$

is valid for $\text{conv}(X_K)$, and if $k = |\bar{C}|$ the cover inequalities (19) are obtained. The $(1, k)$-configuration inequalities are primarily designed to deal with elements j of the knapsack having a large coefficient a_j.

In general inequalities (19) are not facet defining, but they can be made to become facets by applying certain techniques, called *lifting*, to systematically increase the dimension of the face induced by the inequalities. Lifting techniques are described in Part II of this article. A special case of a lifted cover inequality, where all lifting coefficients are equal to zero or one, is obtained by considering the *extension $E(C)$* of a minimal cover C, where $E(C) = \{k \in N \setminus C : a_k \geq a_j, \text{ for all } j \in C\}$. The inequality $\sum_{j \in E(C)} x_j \leq |C| - 1$ is valid for $\text{conv}(X_K)$ and under certain conditions it also defines a facet of $\text{conv}(X_K)$.

The separation problem based on the cover inequalities can again be viewed as a knapsack problem as we show below. Assume we are given the point x^*. To find a cover inequality (19) violated by x^* we need to find a set C such that $\sum_{j \in C} x_j^* > |C| - 1$ and $\sum_{j \in C} a_j > b$. Let $z_j = 1$ if $j \in C$, and let $z_j = 0$ otherwise, and assume without loss of generality that a_j, $j \in N$, and b are integral. For (19) to be violated the z_j-variables have to satisfy the constraints

$$\sum_{j \in N} x_j^* z_j > \left(\sum_{j \in N} z_j \right) - 1 \quad \text{and} \quad \sum_{j \in N} a_j z_j \geq b + 1.$$

The first of the above constraints can be rewritten as $\sum_{j \in N}(1 - x_j^*)z_j < 1$, leading to the following formulation of the problem of finding the most violated cover inequality (19)

$$\min \quad \eta = \sum_{j \in N}(1 - x_j^*)z_j \tag{21}$$

$$\text{s.t.} \quad \sum_{j \in N} a_j z_j \geq b + 1, \tag{22}$$

$$z_j \in \{0, 1\} \quad \text{for all } j \in N. \tag{23}$$

15

A violated cover inequality is identified if and only if $\eta < 1$. To see that the separation problem (21)-(23) is equivalent to a knapsack problem we only need to complement the z_j-variables, i.e. substitute z_j by $1 - z_j$. Problem (21)-(23), however, is often easier to solve than the original knapsack problem since, at a typical fractional solution x^*, many variables take value zero or one. If $x_j^* = 1$, the coefficient of z_j in (21) is equal to zero, and we can set z_j equal to one. Analogously, if $x_j^* = 0$ we set z_j is equal to zero. Therefore, typically few variables remain in the separation problem. Crowder et al. (1983) developed a heuristic for the separation problem and for lifting the inequalities to become facets. Once a minimal cover C is generated it is also used in a heuristic for finding a violated $(1, k)$-configuration inequality. They implemented the algorithms and solved large, real-life, 0-1 integer programming problems without any apparent structure by automatically generating knapsack cover inequalities. This was one of the early computational breakthroughs in combinatorial optimization, as most of the problems were considered not amenable to exact solution within reasonable time. Table 2 gives a summary of the computational results. The valid inequalities were generated and added in the root node of the branch-and-bound tree only. The results include some initial preprocessing to delete some variables and constraints, and to reduce the size of some coefficients. For more details about preprocessing we refer to Part II of this article. In Table 2 *vars*, *constr.*, and *ineq.* denote the number of variables, constraints, and added valid inequalities respectively.

original problem			preprocessing			cutting plane		B&B	
vars	constr.	z_{LP}	vars.	constr.	z_{LP}	ineq.	z_{LP}	nodes	z_{IP}
33	16	2,520.6	33	16	2,819.4	36	3,065.3	113	3,089.0
40	24	61,796.5	40	24	61,829.1	29	61,862.8	11	62,027.0
201	134	6,875.0	195	134	7,125.0	139	7,125.0	1,116	7,615.0
282	242	176,867.5	282	222	176,867.5	462	255,033.1	1,862	258,411.0
291	253	1,705.1	290	206	1,749.9	278	5,022.7	87	5,223.8
548	177	315.3	527	157	3,125.9	296	8,643.5	36	8,691.0
1,550	94	1,706.5	1,550	94	1,706.5	94	1,706.5	10	1,708.0
1,939	109	2,051.1	1,939	109	2,051.1	110	2,051.1	334	2,066.0
2,655	147	6,532.1	2,655	147	6,532.1	149	6,535.0	214	6,548.0
2,756	756	2,688.7	2,734	739	2,701.1	1,065	3,115.3	2,392	3,124.0

Table 2: Results for general zero-one problems.

2.3 The Capacitated Facility Location Problem

In the *capacitated facility location problem* (CFL) we are given a set $M = \{1, \ldots, m\}$ of possible location sites for facilities and a set $N = \{1, \ldots, n\}$ of clients. The capacity, m_j, at each location site is known, as well as the demand, d_k, of each client. The total demand of the clients in the set $S \subseteq N$ is denoted by $d(S)$. We also know the fixed cost of setting up each facility, f_j, and the per unit transportation cost, c_{jk}, between every facility – client pair. We want to determine at which site a facility should be opened and how the flow should be distributed between the open facilities and the clients such that the sum of the fixed costs and the transportation costs is minimized, and such that all clients are served, and all capacity restrictions are satisfied.

Let $y_j = 1$ if facility j is open, and let $y_j = 0$ otherwise. The flow from facility j to client k is denoted by v_{jk}. The mathematical formulation of CFL is given below.

$$\min \quad \sum_{j \in M} f_j y_j + \sum_{j \in M} \sum_{k \in N} c_{jk} v_{jk} \tag{24}$$

$$\text{s.t.} \qquad \sum_{j \in M} v_{jk} = d_k \qquad\qquad \text{for all } k \in N, \tag{25}$$

$$\sum_{k \in N} v_{jk} \leq m_j y_j \qquad\qquad \text{for all } j \in M, \tag{26}$$

$$0 \leq v_{jk} \leq d_k y_j \qquad\qquad \text{for all } j \in M, \ k \in N, \tag{27}$$

$$y_j \in \{0,1\} \qquad\qquad \text{for all } j \in M. \tag{28}$$

Inequalities (27) are redundant for CFL, but they do strengthen the LP-relaxation of CFL and are therefore included here.

By aggregating the flow from each depot we can easily see that a version of the knapsack polytope forms a relaxation of CFL. By using the aggregate flow variables $v_j = \sum_{k \in N} v_{jk}$, $j \in M$, we can obtain the aggregate capacity and demand constraints

$$0 \leq v_j \leq m_j y_j \qquad\qquad \text{for all } j \in M, \tag{29}$$

$$\sum_{j \in M} v_j = d(N). \tag{30}$$

If we combine constraints (29) and (30) with constraint (28) we obtain a so-called *surrogate knapsack polytope* $X_{SK} = \{y \in \{0,1\} : \sum_{j \in M} m_j y_j \geq d(N)\}$. Complementing the y_j-variables, i.e. letting $y'_j = 1 - y_j$ for all $j \in M$, gives the knapsack polytope $\{y' \in \{0,1\} : \sum_{j \in M} m_j y'_j \leq \sum_{j \in M} m_j - d(N)\}$. Hence we can use the knapsack cover inequalities (19) when solving CFL. Note that these inequalities can also be derived for subsets $K \subseteq N$ of the clients. The cover inequalities have proved very useful computationally, as is illustrated in Table 3. In the table we report on the number of branch-and-bound nodes and the time needed to verify optimum if we use the linear relaxation of CFL only, compared to if we use the linear relaxation and added violated lifted knapsack cover inequalities. All instances have 33 facilities and 50 clients.

problem	% gap	B&B nodes	time (s)	cover ineq.	% gap closed	B&B nodes	time (s)
50331	1.5	399	686	13	86.0	31	125
50332	1.2	691	1,560	58	54.3	51	450
50333	1.5	259	556	122	54.1	89	769
50334	0.7	239	493	42	76.6	23	213
50335	1.3	685	1,232	25	78.3	49	248

Table 3: Result of adding knapsack cover inequalities to CFL.

The knapsack polytope is a quite drastic relaxation of CFL since it disregards all flows. Aardal et al. (1993) considered the general family of *submodular inequalities*

$$\sum_{j \in J} \sum_{k \in K_j} v_{jk} \leq f(J) - \sum_{j \in J} (f(J) - f(J \setminus \{j\}))(1 - y_j), \tag{31}$$

where $K_j \subseteq K$ for all $j \in J$, and where $f(J)$ for $J \subseteq M$ is the maximum feasible flow from the facilities in J to the clients in K given the arc set $\{(j,k) : j \in J, \ k \in K_j\}$. The function $f(J)$ is a submodular set function. Here, if $y_j = 1$ for all $j \in J$, the resulting inequality becomes $\sum_{j \in J} \sum_{k \in K_j} v_{jk} \leq f(J)$, which is valid since $f(J)$ is the maximum flow. If $y_k = 0$ for $k \in J$, and $y_j = 1$ for all $j \in J \setminus k$, then the value $f(J) - f(J \setminus \{k\})$ is precisely the amount with which the maximum flow decreases if we close facility k. Aardal et al. completely characterize some large subclasses of the family of facet-defining submodular inequalities. Separation algorithms, and computational results of applying polyhedral techniques to solve CFL, are reported by Aardal (1994).

2.4 The Economic Lot Sizing Problem

In the *economic lot-sizing problem* (ELS) we have a planning horizon consisting of T periods. In each period t, a demand d_t must be satisfied by production in one or more of the periods in $\{1, \ldots, t\}$. We have unit production costs, c_t, and setup costs, f_t, which are incurred whenever production takes place in t. Let x_t denote the production level, and y_t indicate whether a setup is made in period t. Moreover, let $d_{s,t}$, $(1 \leq s \leq t \leq T)$, denote the cumulative demand of the periods $\{s, \ldots, t\}$, i.e., $d_{s,t} = \sum_{\tau=s}^{t} d_\tau$. The mathematical programming formulation of ELS is:

$$\min \ \sum_{t=1}^{T} (f_t y_t + c_t x_t) \tag{32}$$

$$\text{s.t.} \quad \sum_{t=1}^{T} x_t = d_{1,T}, \tag{33}$$

$$\sum_{\tau=1}^{t} x_\tau \geq d_{1,t} \qquad \text{for all } 1 \leq t \leq T-1, \tag{34}$$

$$x_t \leq d_{t,T} y_t \qquad \text{for all } 1 \leq t \leq T, \tag{35}$$

$$x_t \geq 0 \qquad \text{for all } 1 \leq t \leq T, \tag{36}$$

$$y_t \in \{0,1\} \qquad \text{for all } 1 \leq t \leq T. \tag{37}$$

ELS is one of the relatively few combinatorial optimization problems that are polynomially solvable (see Wagelmans et al. (1993) for the description of an $\mathcal{O}(T \log T)$ algorithm). For such problems we can expect to be able to give a compact characterization of the convex hull of feasible solutions, c.f. the matching problem (5)-(7). Barany et al. (1984) showed that the constraints $0 \leq y_t \leq 1$ for all $1 \leq t \leq T$, $y_1 = 1$, (33), (36), together with the exponential class of (l,S)-*inequalities* (38) presented below, completely describe the convex hull of solutions.

Take any $1 \leq l \leq T$ and $S \subset L = \{1, \ldots, l\}$. The (l,S)-inequalities are written as

$$\sum_{t \in L \setminus S} x_t + \sum_{t \in S} d_{t,l} y_t \geq d_{1,l}. \tag{38}$$

The intuition behind the (l, S)-inequalities is as follows. Assume that no production takes place in the periods in S. Then the full demand $d_{1,l}$ has to be produced in the periods in $L \setminus S$, giving $\sum_{t \in L \setminus S} x_t \geq d_{1,l}$. Now, suppose we do produce in some of the periods in S, and let period k be the first such period. The production for demand in periods $\{1, ..., k-1\}$ then has to be done in periods in $L \setminus S$. It is however possible that the remaining demand, $d_{k,l}$, is produced in a single period in S, which explains the coefficients of the y_t-variables. Although the class of (l, S)-inequalities is exponential, we can still solve ELS efficiently by the polyhedral approach since the separation problem based on these inequalities is polynomially solvable (Barany et al. (1984)).

We can generalize ELS by introducing *startup costs*, i.e. a payment for the first period in a set of consecutive periods in which production takes place. This new problem is referred to as ELSS. Below we demonstrate that the (l, S)-inequalities can be generalized to incorporate the variables representing the startups such that the resulting inequalities are valid for ELSS. A typical situation where startups are relevant is when painting items. If we want to start painting after a break of a couple of periods, we need to clean the residue from the old paint, and fill new paint in the machine, which incur a cost. Let the variables z_t, $1 \leq t \leq T$, indicate whether a startup takes place, and let g_t denote the startup cost in period t. The startups are introduced by adding inequalities

$$z_t \geq y_t - y_{t-1}, \quad \text{for all } 1 \leq t \leq T, \quad \text{with } y_0 = 0 \tag{39}$$

to the constraints, and the terms $g_t z_t$, for all t, to the objective function of the formulation of ELS. Let l, L, and S be defined as above, and let R be a subset of S such that the first element of S belongs to R as well. Furthermore, let $p(t) = \max\{j \in S : j < t\}$. If $S \cap \{1, ..., t-1\} = \emptyset$, then $p(t) = 0$. The following (l, R, S)-*inequalities*,

$$\sum_{t \in L \setminus S} x_t + \sum_{t \in R} d_{t,l} y_t + \sum_{t \in S \setminus R} d_{t,l}(z_{p(t)+1} + \ldots + z_t) \geq d_{1,l}, \tag{40}$$

introduced by Van Hoesel, Wagelmans, and Wolsey (1994), describe together with constraints (33), (36), (39), $y_t \geq 0$, and $z_t \leq 1$, $1 \leq t \leq T$ the convex hull of feasible solutions to ELSS. To see that inequalities (40) are valid we again distinguish the cases whether production occurs in periods in S. The case where S contains no production period is analogous to the same case for the (l, S)-inequalities. Now, assume that period k is the first period in S where production takes place, and choose τ as the smallest period such that a setup occurs in all periods $\{\tau, ..., k\}$. Since at least one of the variables $z_\tau, y_\tau, ..., y_k$ appears in the left-hand side of the (l, R, S)-inequality with a coefficient at least equal to $d_{k,l}$, the inequality is valid. The separation problem based on inequalities (40) can be solved as a set of T shortest path problems.

3   Concluding Remarks

We have attempted to present a broad introduction to the theory of polyhedral combinatorics. There are of course several theoretical issues that we have not discussed,

and we have given little attention to the many computational aspects of polyhedral techniques that are necessary to address to solve problems successfully. These issues will however be treated in Part II of this article, where we also present a list, with references, of problems for which polyhedral results are known, and a brief discussion of some alternative techniques for solving combinatorial optimization problems.

Acknowledgment

References

K. AARDAL (1994) "Capacitated facility location: separation algorithms and computational experience", CentER Discussion Paper 9480, Tilburg (submitted for publication).

K. AARDAL, A. HIPOLITO, C.P.M. VAN HOESEL AND B. JANSEN (1995) "A branch-and-cut algorithm for the frequency assignment problem", Working Paper, Limburg University, Maastricht.

K. AARDAL AND C.P.M. VAN HOESEL (1995) "Uncapacitated facility location: lifting, separation and computations" (in preparation).

K. AARDAL, M. LABBÉ, J. LEUNG AND M. QUEYRANNE (1994) "On the two-level uncapacitated facility location problem", CentER Discussion Paper 9486, Tilburg (submitted for publication).

K. AARDAL, Y. POCHET AND L.A. WOLSEY (1993) "Capacitated facility location: valid inequalities and facets", CORE Discussion Paper 9323, Louvain-la-Neuve, (to appear in *Mathematics of Operations Research*).

J.M. VAN DEN AKKER, C.P.M. VAN HOESEL AND M.W.P. SAVELSBERGH (1993) "A time-indexed formulation of single-machine scheduling problems part I: facets and characterizations", Memorandum COSOR 93-27, Eindhoven University of Technology, Eindhoven (submitted for publication).

J.M. VAN DEN AKKER, C.A.J. HURKENS AND M.W.P. SAVELSBERGH (1995) "A time-indexed formulation of single-machine scheduling problems II: branch-and-cut", Memorandum, Eindhoven University of Technology, Eindhoven.

D. APPLEGATE, R.E. BIXBY, V. CHVÁTAL AND W. COOK (1994) "Finding cuts in the TSP", Preliminary report, available through anonymous ftp from netlib.att.com in the directory netlib/att/math/applegate/TSP.

E. BALAS (1975) "Facets of the knapsack polytope", *Mathematical Programming* **8** 146–164.

E. BALAS, S. CERIA, G. CORNUÉJOLS AND NATRAJ (1994) "Gomory cuts revisited", Working Paper No. 1994–16, Carnegie Mellon University, Graduate School of Industrial Administration, Pittsburgh.

I. Barany, T.J. Van Roy and L.A. Wolsey (1984) "Uncapacitated lot-sizing: the convex hull of solutions", *Mathematical Programming Study* **22** 32-43.

V. Chvátal (1973) "Edmonds polytopes and a hierarchy of combinatorial problems", *Discrete Mathematics* **4** 185–224.

V. Chvátal (1975) "On certain polytopes associated with graphs" *Journal of Combinatorial Theory B* **18** 138–154.

W. Cook, C. Coullard and Gy. Turán (1987) "On the complexity of cutting-plane proofs", *Discrete Applied Mathematics* **18** 25–38.

W. Cook, A.M.H. Gerards, A. Schrijver, and É. Tardos (1986) "Sensitivity theorems in integer linear programming", *Mathematical Programming* **34** 251–264.

Y. Crama and F. Spieksma (1995) "Scheduling jobs of equal length: complexity, facets and computational results", in: *Integer Programming and Combinatorial Optimization*, Lecture Notes in Computer Science **920** (E. Balas and J. Clausen, eds.), Springer, Berlin, pp 227–291.

H.P. Crowder, E.L. Johnson and M.W. Padberg (1983) "Solving large-scale zero-one linear programming problems", *Operations Research* **31** 803–934.

H. Crowder and M.W. Padberg (1980) "Solving large-scale symmetric traveling salesman problems to optimality", *Management Science* **26** 459–509.

G.B. Dantzig, D.R. Fulkerson and S.M. Johnson (1954) "Solution of a large-scale traveling-salesman problem", *Operations Research* **2**, 393–410.

J. Edmonds (1965) "Maximum matching and a polyhedron with 0,1-vertices", *Journal of Research of the National Bureau of Standards* (B) **69** 67–72.

J. Edmonds and R. Giles (1977) "A min-max relation for submodular functions on graphs", [in: *Studies in Integer Programming* (P.L. Hammer, E.L. Johnson, B.H. Korte and G.L. Nemhauser, eds.)] *Annals of Discrete Mathematics* **1** 185–204.

C. van Eijl and C.P.M. van Hoesel (1995) "Linear description of the discrete lot sizing problem with Wagner-Whitin costs", Research Report, Eindhoven University of Technology.

M.R. Garey and D.S. Johnson (1979) *Computers and Intractability: A Guide to the Theory of NP-completeness*, Freeman, San Francisco.

A.M.H. Gerards (1989) "An extension of Könings's theorem to graphs with no odd-K_4", *Journal of Combinatorial Theory, Series B* **47** 330–348.

A.M.H. Gerards (1995) "Matching", in: *Handbooks in Operations Research and Management Science: Networks* (M. Ball, T.L. Magnanti, C.L. Monma and G.L. Nemhauser, eds.) (to appear) , North-Holland, Amsterdam.

A.M.H. Gerards and A. Schrijver (1986) "Matrices with the Edmonds-Johnson property", *Combinatorica* **6** 365–389.

A.M.H. Gerards and F.B. Shepherd (1995) "Characterization of the graphs with all subgraphs t-perfect" (in preparation).

R. Giles and W.R. Pulleyblank (1979) "Total dual integrality and integer polyhedra", *Linear Algebra and Its Applications* **25** 191–196.

M.X. Goemans (1993) "Worst-case comparison of valid inequalities for the TSP", (to appear in *Mathematical Programming*).

R.E. Gomory (1958) "Outline of an algorithm for integer solutions to linear programs", *Bulletin of the American Mathematical Society* **64**, 275–278.

R.E. Gomory (1960) "Solving linear programming problems in integers", in: *Combinatorial Analysis* (R. Bellman and M. Hall, Jr., eds.), Proceedings of Symposia in Applied Mathematics X, American Mathematical Society, Providence, pp. 211–215.

R.E. GOMORY (1963) "An algorithm for integer solutions to linear programs", in: *Recent Advances in Mathematical Programming* (R.L. Graves and P. Wolfe, eds.), McGraw-Hill, New York, pp. 269–302.

M. GRÖTSCHEL (1980) "On the symmetric traveling salesman problem: solution of a 120 city problem", *Mathematical Programming Studies* **12** 61–77.

M. GRÖTSCHEL AND O. HOLLAND (1991) "Solution of large-scale symmetric traveling salesman problems", *Mathematical Programming* **51** 141–202.

M. GRÖTSCHEL, L. LOVÁSZ AND A. SCHRIJVER (1981) "The ellipsoid method and its consequences in combinatorial optimization", *Combinatorica* **1** 169–197 [corrigendum: **4** 291–295].

M. GRÖTSCHEL AND M.W. PADBERG (1979) "On the symmetric traveling salesman problem I: inequalities", *Mathematical Programming* **16** 265–280.

M. GRÖTSCHEL AND W.R. PULLEYBLANK (1986) "Clique tree inequalities and the symmetric traveling salesman problem", *Mathematics of Operations Research* **11** 537–569.

Z. GU, G.L. NEMHAUSER AND M.W.P. SAVELSBERGH (1995) "Sequence independent lifting of cover inequalities", in: *Integer Programming and Combinatorial Optimization*, Lecture Notes in Computer Science **920** (E. Balas and J. Clausen, eds.), Springer, Berlin, pp 452–462.

P.L. HAMMER, E.L JOHNSON AND U.N. PELED (1975) "Facets of regular 0-1 polytopes", *Mathematical Programming* **8** 179–206.

C.P.M. VAN HOESEL, A.P.M. WAGELMANS, AND L.A. WOLSEY (1994) "Economic lot sizing with start-up costs: the convex hull", *SIAM Journal of Discrete Mathematics* **7** pages 141–151.

S. VAN HOESEL AND A. KOLEN (1994) "A linear description of the discrete lot sizing and scheduling problem", *European Journal of Operational Research* **75** pages 342–353.

S. VAN HOESEL, A. KOLEN AND VAN EIJL (1995) "A class of valid inequalities for the discrete lot sizing and scheduling problem", (in preparation).

M. JÜNGER, G. REINELT AND S. THIENEL (1995) "Practical problem solving with cutting plane algorithms in combinatorial optimization", in: *Discrete Mathematics and Theoretical Computer Science* **20** (W. Cook, L. Lovász and P. Seymour, eds.)

R.M. KARP (1972) "Reducibility among combinatorial problems", in: *Complexity of Computer Computations* (R.E. Miller and J.W. Thatcher, eds.), Plenum Press, New York, pp 85–103.

R.M. KARP AND C.H. PAPADIMITRIOU (1980) "On linear characterizations of combinatorial optimization problems", *Proceedings of the 21st Annual Symposium on Foundation of Computer Science*, IEEE, New York, pp. 1–9.

H.W. LENSTRA, JR. (1983) "Integer programming with a fixed number of variables", *Mathematics of Operations Research* **8** 538–548.

G.L. NEMHAUSER AND M.W.P. SAVELSBERGH (1992) "A cutting plane algorithm for the single-machine scheduling problem with release times", in: *Combinatorial Optimization: New Frontiers in the Theory and Practice* (M. Akgül, H. Hamacher, S. Tufecki, eds.) NATO ASI Series F: Computer and System Sciences 82, Springer, Berlin, 63–84.

G.L. NEMHAUSER AND L.A. WOLSEY (1988) *Integer and Combinatorial Optimization*, Wiley, New York.

M.W. PADBERG (1980) "$(1, k)$-configurations and facets for packing problems", *Mathematical Programming* **18** 94–99.

M.W. PADBERG AND M. GRÖTSCHEL (1985) "Polyhedral computations" in: *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization* (E. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan end D.B. Shmoys, eds.), John Wiley and Sons.

M.W. Padberg and G. Rinaldi (1987) "Optimization of a 532-city symmetric traveling salesman polytope by branch and cut", *Operations Research Letters* **6** 1–7. [Addendum **9** (1990) 353.]

M.W. Padberg and G. Rinaldi (1990) "Facet identification for the symmetric traveling salesman problem", *Mathematical Programming* **47** 219–257.

M.W.P. Savelsbergh, G.C. Sigismondi and G.L. Nemhauser (1994) "Functional description of MINTO, a Mixed INTeger Optimizer", *Operations Research Letters* **15** 47-58.

A. Schrijver (1980) "On cutting planes", [in: *Combinatorics 79 Part II* (M. Deza and I.G. Rosenberg, eds.)] *Annals of Discrete Mathematics* **9** 291–296.

A. Schrijver (1981) "On total dual integrality", *Linear Algebra and Its Applications* **38** 27–32.

A. Schrijver (1986) *Theory of Linear and Integer Programming*, Wiley, Chichester.

P.D. Seymour (1980) "Decomposition of regular matroids", *Journal of Combinatorial Theory (B)* **28** 305–359.

A.P.M. Wagelmans, C.P.M. van Hoesel and A.W.J. Kolen (1992) "Economic lot-sizing: an $\mathcal{O}(n \log n)$ algorithm that runs in linear time in the Wagner-Whitin case", *Operations Research* **40S** S145-S156.

L.A. Wolsey (1975) "Faces for a linear inequality in 0-1 variables", *Mathematical Programming* **8** 165–178.