Bilevel Programming and Maximally Violated Valid Inequalities

Andrea Lodi*, Ted K. Ralphs°

*DEIS, University of Bologna
Viale Risorgimento 2, 40136, Bologna
andrea.lodi@unibo.it

°Department of Industrial and Systems Engineering
Lehigh University, Bethlehem, PA 18015
ted@lehigh.edu

Key words: Bilevel Programming, Cutting Planes, Separation

1 Introduction

In recent years, *branch-and-cut* algorithms have become firmly established as the most effective method for solving generic mixed integer linear programs (MIPs). Methods for automatically generating inequalities valid for the convex hull of solutions to such MIPs are a critical element of branch-and-cut. This paper examines the nature of the so-called *separation problem*, which is that of generating a valid inequality violated by a given real vector, usually arising as the solution to a relaxation of the original problem. We show that the problem of generating a maximally violated valid inequality often has a natural interpretation as a *bilevel program*. In some cases, this bilevel program can be easily reformulated as a single-level mathematical program, yielding a standard mathematical programming formulation for the separation problem. In other cases, no reformulation exists. We illustrate the principle by considering the separation problem for two well-known classes of valid inequalities.

Formally, we consider a MIP of the form

$$\min\{c^\top x \mid Ax \geq b,\ x \geq 0,\ x \in \mathbb{Z}^I \times \mathbb{R}^C\}, \tag{1}$$

where $A \in \mathbb{Q}^{m\times n}$, $b \in \mathbb{Q}^m$, $c \in \mathbb{Q}^n$, I is the set of indices of components that must take integer values in any feasible solution and C consists of the indices of the remaining components. We assume that other bound constraints on the variables (if any) are included among the problem constraints.

The *continuous* or *linear programming* (LP) relaxation of the above MIP is the mathematical program obtained by dropping the integrality requirement on the variables in I, namely

$$\min_{x \in \mathcal{P}} c^\top x, \tag{2}$$

where $\mathcal{P} = \{x \in \mathbb{R}^n \mid Ax \geq b,\ x \geq 0\}$ is the polyhedron described by the linear constraints of the MIP (1). It is not difficult to see that the convex hull of the set of feasible solutions to (1) is also a polyhedron. This means that in principle, any MIP is equivalent to a linear program over this implicitly defined polyhedron, which we denote as $\mathcal{P}_I$.

A bilevel mixed integer linear program (BMIP) is a generalization of a standard MIP used to model hierarchical decision processes. In a BMIP, the variables are split into a set of *upper-level variables*, denoted by x below, and a set of *lower-level variables*, denoted by y below. Conceptually, the values of the upper-level variables are fixed first, subject to the restrictions of a set of *upper-level constraints*, after which the second-stage variables are fixed by solving a MIP parameterized on the fixed values of the upper-level variables. The canonical integer bilevel MIP is given by

$$\min\Big\{c^1 x + d^1 y \mid x \in \mathcal{P}_U \cap (\mathbb{Z}^{I_1} \times \mathbb{R}^{C_1}),$$
$$y \in \text{argmin}\{d^2 y \mid y \in \mathcal{P}_L(x) \cap (\mathbb{Z}^{I_2} \times \mathbb{R}^{C_2})\}\Big\},$$

where

$$\mathcal{P}_U = \Big\{x \in \mathbb{R}^{n_1} \mid A^1 x \geq b^1, x \geq 0\Big\}$$

is the polyhedron defining the *upper-level feasible region*;

$$\mathcal{P}_L(x) = \Big\{y \in \mathbb{R}^{n_2} \mid G^2 y \geq b^2 - A^2 x, y \geq 0\Big\}$$

is the polyhedron defining the *lower-level feasible region* with respect to a given $x \in \mathbb{R}^{n_1}$; $A^1 \in \mathbb{Q}^{m_1 \times n_1}$; $b^1 \in \mathbb{Q}^{m_1}$; $A^2 \in \mathbb{Q}^{m_2 \times n_1}$, $G^2 \in \mathbb{Q}^{m_2 \times n_2}$; and $b^2 \in \mathbb{Q}^{m_2}$. The index sets I_1, I_2, C_1, and C_2 are the bilevel counterparts of the sets I and C defined previously. For more detailed information, Colson et al. [2005] provide an introduction to and comprehensive survey of the bilevel programming literature, while Moore and Bard [1990] introduce the discrete case. Dempe [2003] provides a detailed bibliography.

A *valid inequality* for a set $\mathcal{S} \subseteq \mathbb{R}^n$ is a pair (α, β), where $\alpha \in \mathbb{R}^n$ is the *coefficient vector* and $\beta \in \mathbb{R}$ is the *right-hand side*, such that $\alpha^\top x \geq \beta$ for all $x \in \mathcal{S}$. Associated with any valid inequality (α, β) is the half-space $\{x \in \mathbb{R}^n \mid \alpha x \geq \beta\}$, which must contain $\mathcal{S}$. It is easy to see that any inequality valid for $\mathcal{S}$ is also valid for the convex hull of $\mathcal{S}$.

For a polyhedron $\mathcal{Q} \subseteq \mathbb{R}^n$, the so-called *separation problem* is to generate a valid inequality violated by a given vector. Formally, we define the problem as follows.

Definition 1 *The* separation problem *for a polyhedron $\mathcal{Q}$ is to determine for a given $\hat{x} \in \mathbb{R}^n$ whether or not $\hat{x} \in \mathcal{Q}$ and if not, to produce an inequality $(\bar{\alpha}, \bar{\beta}) \in \mathbb{R}^{n+1}$ valid for $\mathcal{Q}$ and for which $\bar{\alpha}^\top \hat{x} < \bar{\beta}$.*

A closely associated problem that is more relevant in practice is the *maximally violated valid inequality problem* (MVVIP), which is as follows.

Definition 2 *The* maximally violated valid inequality problem *for a polyhedron $\mathcal{Q}$ is to determine for a given $\hat{x} \in \mathbb{R}^n$ whether or not $\hat{x} \in \mathcal{Q}$ and if not, produce an inequality $(\bar{\alpha}, \bar{\beta}) \in \mathbb{R}^{n+1}$ valid for $\mathcal{Q}$ and for which $(\bar{\alpha}, \bar{\beta}) \in \mathrm{argmin}_{(\alpha,\beta)\in\mathbb{R}^{n+1}}\{\alpha^\top \hat{x} - \beta \mid \alpha^\top x \geq \beta \ \forall x \in \mathcal{Q}\}$.*

It is well-known that both the separation problem and the MVVIP for a polyhedron $\mathcal{Q}$ are polynomially equivalent to the associated optimization problem, which is to determine $\min_{x\in\mathcal{Q}} d^\top x$ [Grötschel et al., 1981] for a given $d \in \mathbb{R}^n$. In the present context, this means it is unlikely that the MVVIP for $\mathcal{P}_I$ can be solved easily unless the MIP itself can be solved easily.

Because the general MVVIP is usually too difficult to solve, valid inequalities are generated by solving (either exactly or approximately) the MVVIP for one or more relaxations of the original problem. These relaxations often come from considering valid inequalities in a specific *family* or *class*, i.e., inequalities that share a special structure. Applegate et al. [2007] called this paradigm for generation of valid inequalities the *template paradigm*. Generally speaking, a class of valid inequalities for a given set $\mathcal{S}$ is simply a subset of all valid inequalities for $\mathcal{S}$. Such subsets can be defined in a number of ways and may be either finite or infinite. Associated with any given class $\mathcal{C}$ is its closure $\mathcal{F}_\mathcal{C}$, consisting of the region defined by the intersection of all half-spaces associated with inequalities in the class. If the class is finite, then the closure is a polyhedron. Otherwise, it may or may not be a polyhedron.

Let us consider a given class of valid inequalities $\mathcal{C}$. Assuming the closure $\mathcal{F}_\mathcal{C}$ is a polyhedron, both the separation problem and the MVVIP for $\mathcal{C}$ can be identified with the previously defined separation problem and MVVIP for $\mathcal{F}_\mathcal{C}$. A number of authors have noted that the MVVIP for certain classes of valid inequalities can be formulated as structured mathematical programs in their own right and solved using standard optimization techniques (see, e.g., [Balas, 1979], Caprara and Letchford [2003] and [Fischetti and Lodi, 2007]). We wish to show that the underlying structure of the MVVIP is inherently *bilevel*.

The bilevel nature of the MVVIP for a class $\mathcal{C}$ arises from the fact that for a given coefficient vector $\alpha \in \mathbb{R}^n$, the calculation of the right-hand side β required to ensure (α, β) is a member of the class (if such a β exists) may itself be an optimization problem that we refer to as the *right-hand side generation problem* (RHSGP). The complexity of the separation problem depends strongly on the complexity of the RHSGP. In cases where the RHSGP is in the complexity class *NP*-hard, it is generally not possible to formulate the separation problem as a traditional mathematical program. In fact, such separation problems may not even be in the complexity class *NP*. Roughly speaking, the reason for this is that the problem of determining whether a given inequality is valid is then itself a hard problem.

Putting these question aside for now, however, let us simply define the set $\mathcal{C}_\alpha \subseteq \mathbb{R}^n$ to be the projection of $\mathcal{C}$ into the space of coefficient vectors. In other words, $\mathcal{C}_\alpha$ is the set of all vectors that are coefficients for some valid inequality in $\mathcal{C}$. Then the MVVIP for $\mathcal{C}$ with respect to a given $\hat{x} \in \mathbb{R}^n$ can in principle be formulated mathematically as

$$\min \quad \alpha^\top \hat{x} - \beta \tag{3}$$

$$\alpha \in \mathcal{C}_\alpha \tag{4}$$

$$\beta = \min \alpha^\top x \tag{5}$$

$$x \in \mathcal{F}_\mathcal{C}. \tag{6}$$

The problem (3)–(6) is a bilevel program in which the *upper-level objective* (3) is to find the maximally violated inequality in the class. The upper-level constraints (4) require that the inequality is a member of the class. The lower-level problem (5)–(6) is to generate the strongest possible right-hand side associated with a given coefficient vector.

It is easy to see that the above separation problem may be very difficult to solve in some cases. In fact, the complexity depends strongly on the complexity of the RHSGP and whether the sense of the optimization "agrees" with that of the MVVIP itself. Most of the separation algorithms appearing in the literature define the set $\mathcal{F}_\mathcal{C}$ in such a way that the bilevel program (3)–(6) collapses into a single-level program, generally linear or mixed integer linear.

In the remainder of the paper, we describe two well-known classes of valid inequalities and give a bilevel interpretation of their associated separation problems. In Section 2, we consider the well-known class of *disjunctive valid inequalities* for general MIPs. For such a class, we show that it is quite straightforward to convert the BMIP (3)–(6) into a single-level mathematical program, though the MVVIP might nevertheless remain difficult from a practical standpoint. In Section 3, we focus on the so-called *capacity constraints* for the classical *Capacitated Vehicle Routing Problem* (CVRP). There are several closely-related variants of this class of valid inequalities and we show that for the strongest of these, there is no straightforward way to convert the BMIP into a single-level program. That is the main contribution of the present paper, and to the best of our knowledge, it is a new result. Finally, some conclusions are drawn in Section 4.

2 Disjunctive Valid Inequalities for general MIPs

Given a MIP in the form (1), Balas [1979] showed how to derive a valid inequality by exploiting any disjunction of the form

$$\pi^\top x \leq \pi_0 \quad \text{OR} \quad \pi^\top x \geq \pi_0 + 1 \; \forall x \in \mathbb{R}^n, \tag{7}$$

where $\pi \in \mathbb{Z}^I \times \mathbf{0}^C$ and $\pi_0 \in \mathbb{Z}$. More precisely, the family of disjunctive inequalities (also called *split cuts*) are all those valid for the union of the two polyhedra, denoted by $\mathcal{P}_1$ and $\mathcal{P}_2$, obtained from $\mathcal{P}$ by adding inequalities $(-\pi, -\pi_0)$ and $(\pi, \pi_0 + 1)$, respectively.

For a given disjunction of the form (7), the separation problem for the associated family of disjunctive inequalities with respect to a given vector $\hat{x} \in \mathcal{P}$ can be written as a the following bilevel LP:

$$\begin{aligned}
&\min \alpha^\top \hat{x} - \beta && (8)\\
&\quad \alpha_j \geq u^\top A_j - u_o \pi_j \quad j \in I \cup U && (9)\\
&\quad \alpha_j \geq v^\top A_j + v_o \pi_j \quad j \in I \cup U && (10)\\
&\quad u, v, u_0, v_0 \geq 0 && (11)\\
&\quad u_0 + v_0 = 1 && (12)\\
&\quad \beta = \min \alpha^\top x && (13)\\
&\qquad\qquad x \in \mathcal{P}_1 \cup \mathcal{P}_2. && (14)
\end{aligned}$$

Constraints (9) and (10) together with the non-negativity requirements on the dual multipliers (11) ensure the coefficients constitute those of a disjunctive inequality. (Constraint (12) is one of the possible normalizations to make the mathematical program above bounded, see, e.g., Fischetti et al. [2008].) Once the coefficient vector and the corresponding dual multipliers are known, the RHSGP is easy to solve. To obtain a valid inequality, one has only to set β to $\min\{u^\top b - u_0\pi_0, v^\top b + v_0(\pi_0 + 1)\}$, which is the smallest of the right-hand sides obtained by the sets of multipliers (u, u_0) and (v, v_0) corresponding to the constraints of $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. It is easy to reformulate the bilevel LP above into the following (single level) linear program by a well-known modeling trick:

$$
\begin{aligned}
\min\ & \alpha^\top \hat{x} - \beta && (15)\\
& \alpha_j \geq u^\top A_j - u_o \pi_j \quad j \in I \cup U \\
& \alpha_j \geq v^\top A_j + v_o \pi_j \quad j \in I \cup U \\
& \beta \leq u^\top b - u_0 \pi_0 && (16)\\
& \beta \leq v^\top b + v_0(\pi_0 + 1) && (17)\\
& u_0 + v_0 = 1 \\
& u, v, u_0, v_0 \geq 0.
\end{aligned}
$$

Indeed, note that for given values of the remaining variables, any value of β satisfying the two inequalities (16) and (17) above yields a valid disjunctive constraint. Furthermore, these two inequalities ensure that $\beta \leq \min\{u^\top b - u_0\pi_0, v^\top b + v_0(\pi_0+1)\}$, while the objective function (15) ensures that the largest possible value of β is indeed selected, i.e., $\beta = \min\{u^\top b - u_0\pi_0, v^\top b + v_0(\pi_0+1)\}$. In other words, the objective function (15) gives for free the best value of the right-hand side, thus finding the strongest cut.

If the disjunction is not given a priori, i.e., one is searching among the set of possible disjunctions for the one yielding the most violated constraint, the above program can still be used, but π and π_0 become integer variables. The same trick can be applied to transform the bilevel separation problem into a single-level one, but the problem remains difficult because (i) some of the constraints contain bilinear terms, and (ii) the program involves the integer variables π and π_0. The solution of such a formulation has been addressed by Balas and Saxena [2008] and Dash et al. [2007].

3 Capacity Constraints for the CVRP

Here, we consider the classical *Capacitated Vehicle Routing Problem* (CVRP), as introduced by Dantzig and Ramser [1959], in which a quantity d_i of a single commodity is to be delivered to each customer $i \in N = \{1, \ldots, n\}$ from a central depot $\{0\}$ using a homogeneous fleet of k vehicles, each with capacity K. The objective is to minimize total cost, with $c_{ij} \geq 0$ denoting the fixed cost of transportation from location i to location j, for $0 \leq i, j \leq n$. The costs are assumed to be *symmetric*, i.e., $c_{ij} = c_{ji}$ and $c_{ii} = 0$.

This problem is naturally associated with the complete undirected graph consisting of nodes $N \cup \{0\}$, edge set $E = N \times N$, and edge costs $c_{ij}, \{i, j\} \in E$. In this graph, a solution is the union of k cycles whose only intersection is the depot node and whose union covers all customers. By associating an integer variable with each edge in the graph, we obtain the following integer programming formulation:

$$\min \sum_{e \in E} c_e x_e$$

$$\sum_{e=\{0,j\} \in E} x_e = 2k \tag{18}$$

$$\sum_{e=\{i,j\} \in E} x_e = 2 \quad \forall i \in N \tag{19}$$

$$\sum_{\substack{e=\{i,j\} \in E \\ i \in S, j \notin S}} x_e \geq 2b(S) \quad \forall S \subset N, \ |S| > 1 \tag{20}$$

$$0 \ \leq \ x_e \leq 1 \quad \forall e = \{i,j\} \in E, \ i,j \neq 0 \tag{21}$$

$$0 \ \leq \ x_e \leq 2 \quad \forall e = \{0,j\} \in E \tag{22}$$

$$x_e \quad integral \quad \forall e \in E. \tag{23}$$

Constraints (18) and (19) are the *degree constraints*. In constraints (20), referred to as the *capacity constraints*, $b(S)$ is any of several lower bounds on the number of trucks required to service the customers in set S. These constraints can be viewed as a generalization of the subtour elimination constraints from the *Traveling Salesman Problem* and serve both to enforce the connectivity of the solution and to ensure that no route has total demand exceeding the capacity K. The easily calculated lower bound $\sum_{i \in S} d_i/K$ on the number of trucks is enough to ensure the formulation (18)–(23) is correct, but increasing this bound through the solution of a more sophisticated RHSGP will yield a stronger version of the constraints.

The MVVIP for capacity constraints with a generic lower bound $b(S)$ can be formulated as a BMIP of the form (3)–(6) as follows. Because we are looking for a set $\bar{S} \subset N$ for which an inequality (20) is maximally violated, we define the binary variables

$$y_i = \begin{cases} 1 \text{ if customer } i \text{ belong to } \bar{S} \\ 0 \text{ otherwise} \end{cases} \qquad i \in N, \tag{24}$$

and

$$z_e = \begin{cases} 1 \text{ if edge } e \text{ belong to } \delta(\bar{S}) \\ 0 \text{ otherwise} \end{cases} \qquad e \in E, \tag{25}$$

where $\delta(\bar{S})$ denotes the set of edges in E with one endpoint in $\bar{S}$, to model selection of the members of the set $\bar{S}$ and the coefficients of the corresponding inequality. Thus, the formulation is

$$\min \sum_{e \in E} \hat{x}_e z_e - 2b(\bar{S}) \tag{26}$$

$$z_e \geq y_i - y_j \qquad \forall e = \{i, j\} \tag{27}$$

$$z_e \geq y_j - y_i \qquad \forall e = \{i, j\} \tag{28}$$

$$\max b(\bar{S}) \tag{29}$$

$$b(\bar{S}) \text{ is a valid lower bound.} \tag{30}$$

For improved tractability, the RHSGP (29)–(30) can be replaced by the calculation of a specific bound. One of the strongest possible lower bounds is obtained by solving to optimality the (strongly *NP*-hard) *Bin Packing Problem* (BPP) with the customer demands in set $\bar{S}$ being packed into the minimum number of bins of size K (Cornuéjols and Harche [1993] describe a further strengthening of the right-hand side, but we we do not consider this bound here). The RHSGP based on the BPP can be modeled by using the binary variables

$$w_i^\ell = \begin{cases} 1 \text{ if customer } i \text{ is served by vehicle } \ell \\ 0 \text{ otherwise} \end{cases} \qquad (i \in N, \ell = 1, \ldots, k), \tag{31}$$

and

$$h_\ell = \begin{cases} 1 \text{ if vehicle } \ell \text{ is used} \\ 0 \text{ otherwise} \end{cases} \qquad (\ell = 1, \ldots, k). \tag{32}$$

Then the full separation problem reads as follows:

$$\min \sum_{e \in E} \hat{x}_e z_e - 2b(\bar{S}) \tag{33}$$

$$z_e \geq y_i - y_j \qquad \forall e = \{i, j\} \tag{34}$$

$$z_e \geq y_j - y_i \qquad \forall e = \{i, j\} \tag{35}$$

$$b(\bar{S}) = \min \sum_{\ell=1}^{n} h_\ell \tag{36}$$

$$\sum_{\ell=1}^{n} w_i^\ell = y_i \qquad \forall i \in N \tag{37}$$

$$\sum_{i \in N} d_i w_i^\ell \leq K h_\ell \qquad \ell = 1, \ldots, n, \tag{38}$$

where of course all variables y, z, w and h are binary.

It is clear that the BMIP (33)–(38) cannot be straightforwardly reduced to a single-level program because the sense of the optimization of the RHSGP is opposed to that of the MVVIP. In other words, because of the upper-level objective function (33), the absence of the lower-level objective would result in a BPP solution using the largest number of bins instead of the smallest.

We can simplify the RHSGP by relaxing the integrality requirement on w and h to obtain

$$b(\bar{S}) = \min \sum_{\ell=1}^{n} h_\ell \tag{39}$$

$$\sum_{\ell=1}^{n} w_i^\ell = y_i \qquad \forall i \in N \tag{40}$$

$$\sum_{i \in N} d_i w_i^\ell \leq K h_\ell \qquad \ell = 1, \ldots, n \tag{41}$$

$$w_i^\ell \in [0,1],\ h_\ell \in [0,1] \qquad i \in N, \ell = 1, \ldots, n, \tag{42}$$

which is also a lower bound for the BPP. In this case, the RHSGP can be solved in closed form, with an optimal solution being

$$b(\bar{S}) = \frac{\sum_{i \in \bar{S}} d_i}{K} = \frac{\sum_{i \in N} d_i y_i}{K}. \tag{43}$$

Hence, the MVVIP reduces to a single-level MIP that can in turn be solved in polynomial time by transforming it into a network flow problem as proven by McCormick et al. [2003].

An intermediate valid lower bound is obtained by rounding the bound (43), i.e., using $b(S) = \left\lceil \frac{\sum_{i \in N} d_i y_i}{K} \right\rceil$. Although such rounding can be done after the fact, relaxing the integrality on w, but not h, i.e., replacing conditions (42) by

$$w_i^\ell \in [0,1],\ h_\ell \in \{0,1\} \quad i \in N, \ell = 1, \ldots, n,$$

results in reduction of the MVVIP to the single-level MIP

$$\begin{aligned} \min \sum_{e \in E} \hat{x}_e z_e - 2b & \\ z_e \geq y_i - y_j & \qquad \forall e = \{i,j\} \\ z_e \geq y_j - y_i & \qquad \forall e = \{i,j\} \\ b \geq \frac{\sum_{i \in N} d_i y_i}{K} & \\ b \ integral & \\ y_i \in \{0,1\},\ z_e \in \{0,1\} & \qquad \forall i \in N, \forall e \in E, \end{aligned}$$

which was shown by Cornuéjols and Harche [1993] to be *NP*-hard.

4 Conclusions

We have presented a conceptual framework for the formulation of general separation problems as bilevel programs. This framework reflects the inherent bilevel nature of the separation problem arising from the fact that calculation of a valid right-hand side for a given coefficient vector is itself an optimization problem. In cases where this optimization problem is difficult in a complexity sense, it is generally not possible to formulate the separation problem as a traditional mathematical program. We conjecture that the MVVIP for most classes of valid inequalities can be thought of as having this hierarchical structure, but that certain of them can nonetheless be reformulated effectively. This is either because the RHSGP is easy to solve or because it goes "in the right direction" with respect to the MVVIP itself. We believe that the paradigm presented here may be useful for the analysis of other intractable classes of valid inequalities. In a future study, we plan to further formalize the conceptual framework presented here with a further investigation of the complexity issues, additional examples of this phenomena, and an assessment whether these ideas may be useful from a computational perspective.

Acknowledgements

We warmly thank Leo Liberti for useful comments about the paper.

References

D. L. Applegate, R. E. Bixby, V. Chvátal, and W. J. Cook. *The Traveling Salesman Problem: A Computational Study.* Princeton University Press, 2007.

E. Balas. Disjunctive programming. *Annals of Discrete Mathematics*, 5:3–51, 1979.

E. Balas and A. Saxena. Optimizing over the split closure. *Mathematical Programming*, 113:219–240, 2008.

A. Caprara and A. Letchford. On the separation of split cuts and related inequalities. *Mathematical Programming*, 94:279–294, 2003.

B. Colson, P. Marcotte, and G. Savard. Bilevel programming: A survey. *4OR: A Quarterly Journal of Operations Research*, 3:87–107, 2005.

G. Cornuéjols and F. Harche. Polyhedral Study of the Capacitated Vehicle Routing Problem. *Mathematical Programming*, 60:21–52, 1993.

G. B. Dantzig and R. H. Ramser. The Truck Dispatching Problem. *Management Science*, 6:80–91, 1959.

S. Dash, O. Günlük, and A. Lodi. On the MIR closure of polyhedra. In M. Fischetti and D. P. Williamson, editors, *Integer Programming and Combinatorial Optimization - IPCO 2007*, volume 4513 of *Lecture Notes in Computer Science*, pages 337–351. Springer-Verlag, 2007.

S. Dempe. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization*, 52:333–359, 2003.

M. Fischetti and A. Lodi. Optimizing over the first Chvátal closure. *Mathematical Programming*, 110:3–20, 2007.

M. Fischetti, A. Lodi, and A. Tramontani. On the separation of disjunctive cuts. Technical Report OR-08-2, DEIS, University of Bologna, 2008.

M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1:169–197, 1981.

S. T. McCormick, M. R. Rao, and G. Rinaldi. Easy and difficult objective functions for max cut. *Mathematical Programming*, 94:459–466, 2003.

J. Moore and J. Bard. The mixed integer linear bilevel programming problem. *Operations Research*, 38:911–921, 1990.