Structural Analysis of Branch-and-Cut and the Learnability of Gomory Mixed Integer Cuts

Nina Balcan Carnegie Mellon University

Siddharth Prasad Carnegie Mellon University

Tuomas Sandholm Carnegie Mellon University, Optimized Markets, Inc., Strategic Machine, Inc., Strategy Robot, Inc.

Ellen Vitercik Stanford University

NeurIPS 2022

Branch-and-bound

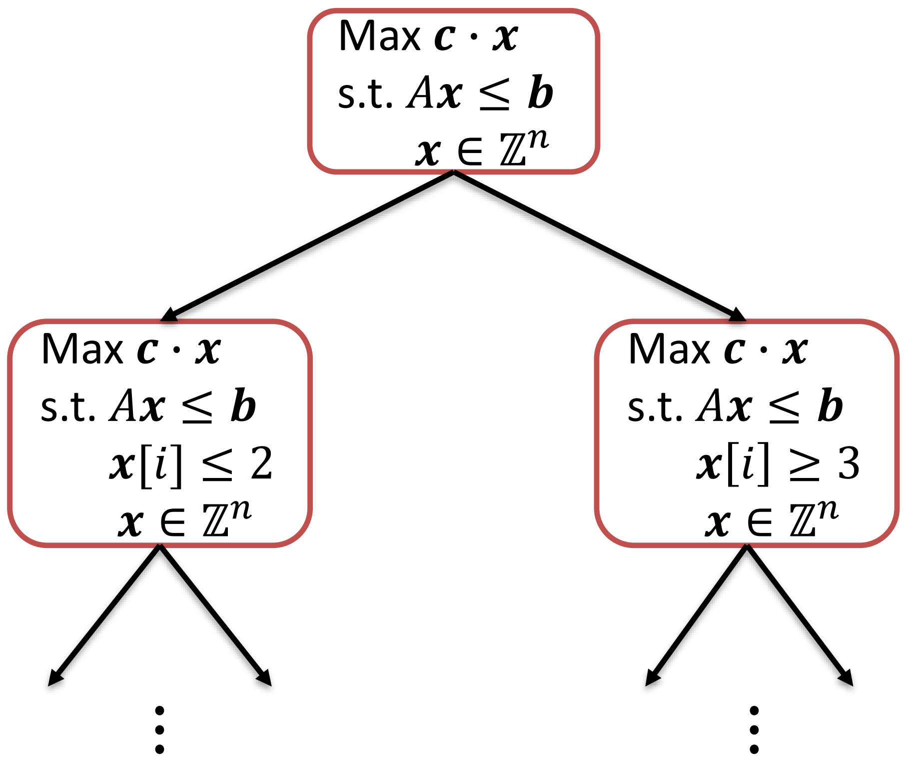
- Powerful tree-search algorithm used to solve IPs in practice

- Uses the linear programming (LP) relaxation to do an informed search through the set of feasible integer solutions

IP		LP relaxation
Max $\boldsymbol{c} \cdot \boldsymbol{x}$ s.t. $A\boldsymbol{x} \leq \boldsymbol{b}$ $\boldsymbol{x} \in \mathbb{Z}^n$	→	Max $\boldsymbol{c} \cdot \boldsymbol{x}$ s.t. $A\boldsymbol{x} \leq \boldsymbol{b}$ $\boldsymbol{x} \in \mathbb{R}^n$

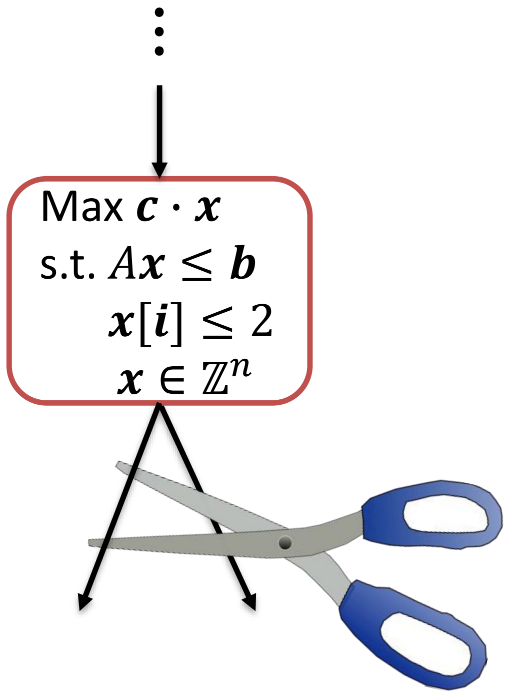
Branch-and-bound: branching

- Choose variable *i* to branch on.
- Generate one subproblem with $\boldsymbol{x}[i] \leq \lfloor \boldsymbol{x}_{\mathrm{LP}}^*[i] \rfloor$ another with $\boldsymbol{x}[i] \geq \lceil \boldsymbol{x}_{\mathrm{LP}}^*[i] \rceil$



Branch-and-bound: pruning

- Prune subtrees if
  - LP relaxation at a node is integral, infeasible, or
  - (Bounding) LP optimal *worse* than best feasible integer solution found so far

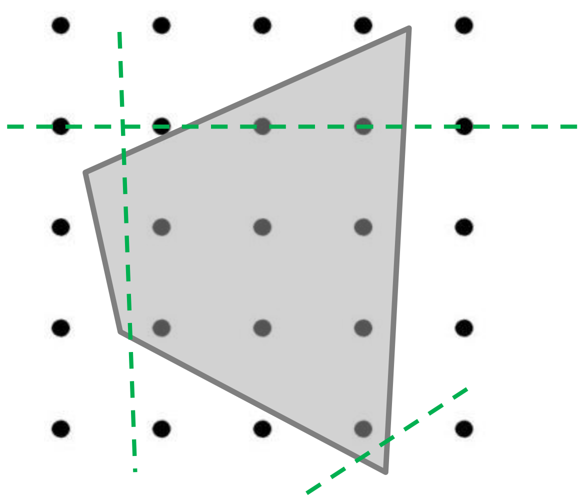


Branch-and-cut

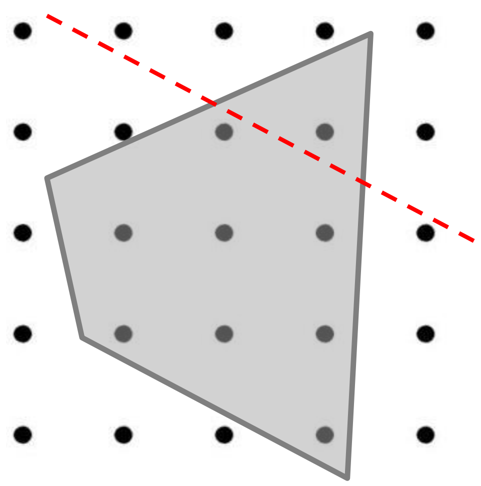
- Branch-and-bound, but at each node may add *cutting planes*
- Method of getting tighter LP relaxation bounds, and thus pruning subtrees sooner

Cutting planes

- Constraint $\boldsymbol{\alpha x} \leq \beta$ is a *valid cutting plane* if it does not cut off any integer feasible points



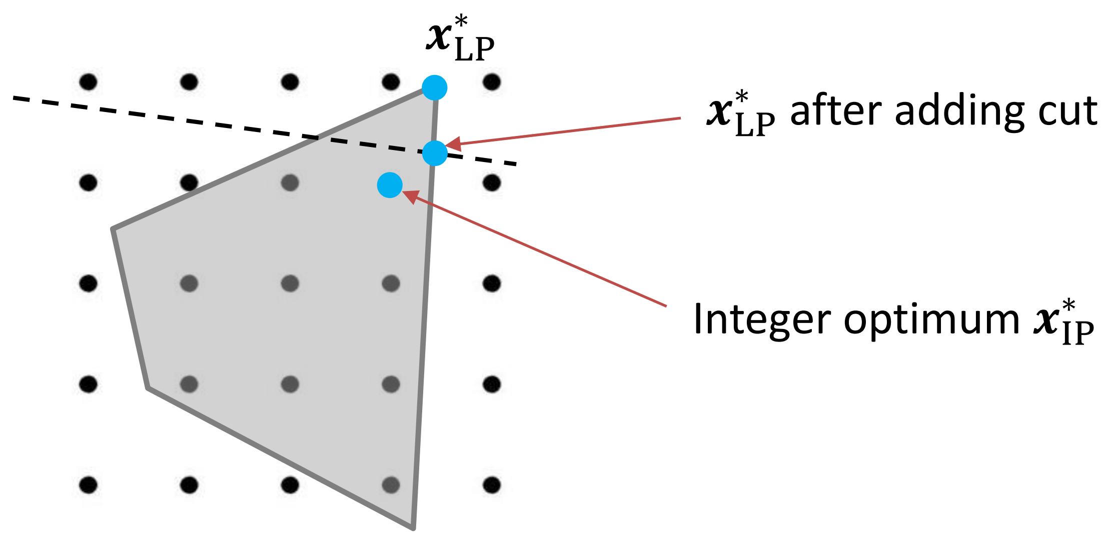
Valid cutting planes



An invalid cutting plane

Cutting planes

- If $\boldsymbol{\alpha x} \leq \beta$ is valid and separates the LP optimum, can speed up B&C by pruning nodes sooner



Gomory Mixed Integer Cuts

$$\sum_{i:f_i\le f_0} f_i x_i + \frac{f_0}{1-f_0}\sum_{i:f_i>f_0}(1-f_i)x_i \ge f_0$$

- *Gomory Cuts Revisited*, 1996. Balas, Ceria, Cornuéjols, Natraj
  - achieved remarkable speedups by effectively integrating these cuts into branch-and-bound (believed to be practically useless before)

- Today: Gomory mixed integer cuts a crucial component of commercial solvers like CPLEX and Gurobi

G. Cornuéjols. Revival of the Gomory cuts in the 1990’s. *Annals of OR*. 2007.

Main contributions

- The first generalization guarantees for using machine learning to add Gomory mixed integer cuts
- A novel structural analysis of the branch-and-cut algorithm that pins down its possible behaviors

Generalization guarantees for cutting planes

Distribution-dependent cut selection

Learning to cut

If a cut yields small average branch-and-cut tree size over IP samples…

$$\begin{array}{c}\text{Max } \boldsymbol{c}_{\mathbf{1}} \cdot \boldsymbol{x} \\ \text{s.t. } A_1 \boldsymbol{x} \leq \boldsymbol{b}_{\mathbf{1}} \\ \boldsymbol{x} \in \mathbb{Z}^n\end{array} \quad \bullet\bullet\bullet \quad \begin{array}{c}\text{Max } \boldsymbol{c}_N \cdot \boldsymbol{x} \\ \text{s.t. } A_N \boldsymbol{x} \leq \boldsymbol{b}_N \\ \boldsymbol{x} \in \mathbb{Z}^n\end{array} \quad \sim D$$

IP 1 IP N

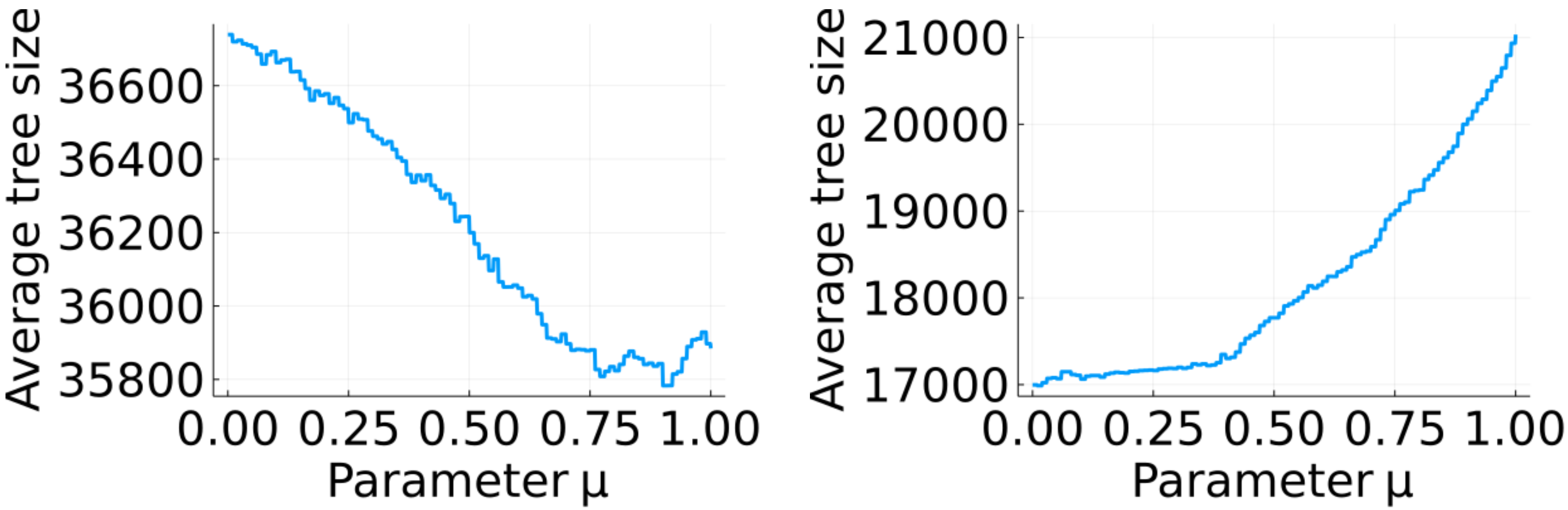
…is it likely to yield a small branch-and-cut tree on a fresh IP?

$$\begin{array}{c}\text{Max } \boldsymbol{c} \cdot \boldsymbol{x} \\ \text{s.t. } A \boldsymbol{x} \leq \boldsymbol{b} \\ \boldsymbol{x} \in \mathbb{Z}^n\end{array} \quad \sim D$$

Tuning a GMI cut selection parameter

- E.g. mixture of $d = 2$ scores

$$\mu \cdot \text{parallelism} + (1 - \mu) \cdot \text{efficacy}$$



Two different distributions over facility location IPs.

Generalization for Gomory mixed integer cuts

Theorem [Balcan, Prasad, Sandholm, Vitercik NeurIPS'22]: for all GMI cuts $\boldsymbol{u} \in [-U, U]^m$, difference between average training performance over N samples and expected performance is (whp)

$$\tilde{O}\left(H\sqrt{\frac{mn^3 \log\left(mn\tau U \|A\|_{1,1} \|\boldsymbol{b}\|_1\right)}{N}}\right)$$

Proof uses our structural analysis of branch-and-cut

A structural analysis of branch-and-cut

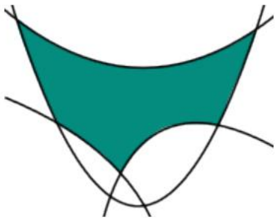
- Given two valid cutting planes

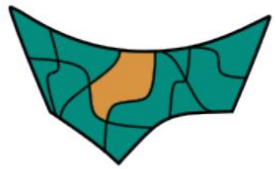
$$\boldsymbol{\alpha_1 x} \leq \beta_1 \text{ and } \boldsymbol{\alpha_2 x} \leq \beta_2$$

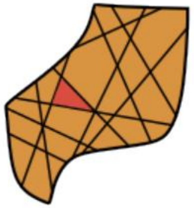
- When does B&C behave identically on the following IPs?

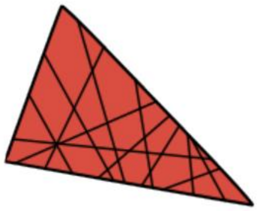
Max $\boldsymbol{c} \cdot \boldsymbol{x}$ s.t. $A\boldsymbol{x} \leq \boldsymbol{b}$ $\boldsymbol{\alpha_1 x} \leq \beta_1$ $\boldsymbol{x} \in \mathbb{Z}^n$	Max $\boldsymbol{c} \cdot \boldsymbol{x}$ s.t. $A\boldsymbol{x} \leq \boldsymbol{b}$ $\boldsymbol{\alpha_2 x} \leq \beta_2$ $\boldsymbol{x} \in \mathbb{Z}^n$

Branch-and-cut piecewise invariance


(a) LP optimum closed form


(b) Invariant branching


(c) Invariant LP integrality


(d) Invariant B&C execution

Theorem [BPSV NeurIPS'22] : Given IP $(\boldsymbol{c}, A, \boldsymbol{b})$, there are $O\left(14^n(m+2n)^{3n^2}\tau^{5n^2}\right)$ polynomial hypersurfaces of degree ≤ 5 that partition $\mathbb{R}^{n+1}$ into connected components such that the branch-and-cut tree built after adding the cut $\boldsymbol{\alpha x} \leq \beta$ is invariant over all $(\boldsymbol{\alpha}, \beta)$ within a given component.

$$\tau := \left\lceil \max_{\boldsymbol{x} \in P} \|\boldsymbol{x}\|_\infty \right\rceil \leq \|A\|_{\infty,\infty}^n n^{n/2}$$

If A has a row with all positive entries, then $\tau \leq \|\boldsymbol{b}\|_\infty$