Decomposition and Dynamic Cut Generation in Integer Programming

Matthew V. Galati

Ted K. Ralphs

http://sagan.ie.lehigh.edu/mgalati

Department of Industrial and Systems Engineering
Lehigh University, Bethlehem, PA

IBM-Research Yorktown Heights (NY) and Rüschlikon (ZH)

The Swiss Operations Research Society

École Polytechnique Fédérale de Lausanne, Monday July 14, 2003

Outline

- Preliminaries, Traditional Decomposition Methods
 - Dantzig-Wolfe Decomposition
 - Lagrangian Relaxation
 - Cutting Plane Method
- Dynamic Decomposition Methods
 - Price and Cut
 - Relax and Cut
 - Decompose and Cut
- Applications/Examples
- DECOMP Library Framework

Preliminaries

Consider the following pure integer linear program (PILP):

$$z_{IP} = \min_{x \in \mathcal{F}} \{c^{\top} x\} = \min_{x \in \mathcal{P}} \{c^{\top} x\} = \min_{x \in \mathbb{Z}^n} \{c^{\top} x : Ax \ge b\}$$

where

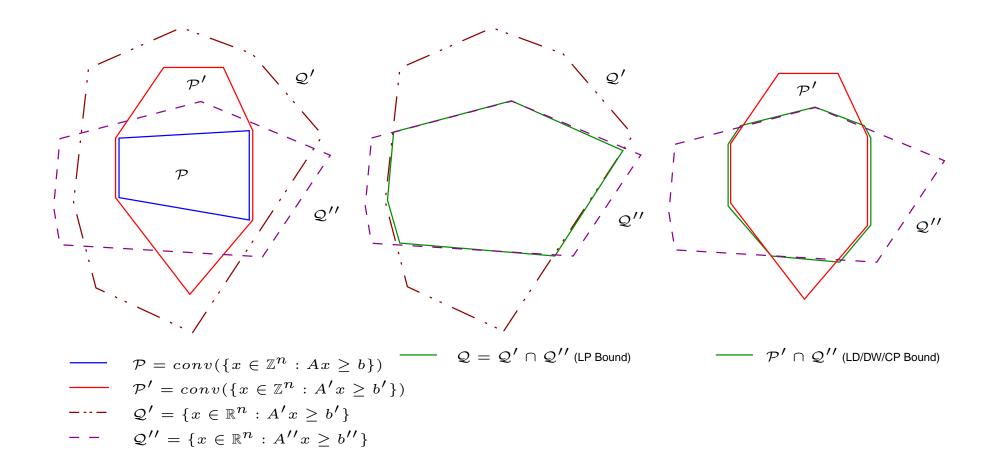
$$\mathcal{F} = \{ x \in \mathbb{Z}^n : A'x \ge b', A''x \ge b'' \} \qquad \mathcal{Q} = \{ x \in \mathbb{R}^n : A'x \ge b', A''x \ge b'' \}$$

$$\mathcal{F}' = \{ x \in \mathbb{Z}^n : A'x \ge b' \} \qquad \qquad \mathcal{Q}' = \{ x \in \mathbb{R}^n : A'x \ge b' \}$$

$$\mathcal{Q}'' = \{ x \in \mathbb{R}^n : A''x \ge b'' \}$$

- Denote $\mathcal{P} = conv(\mathcal{F})$ and $\mathcal{P}' = conv(\mathcal{F}')$.
- Assume that optimization (separation) over P is difficult.
- Assume that optimization (separation) over \mathcal{P}' can be done *effectively*.

Polyhedra, LP Bound, LD/DW/CP Bound



Bounding

- Goal: Compute a lower bound on z_{IP}.
- The most straightforward approach is to solve the initial LP relaxation

$$z_{LP} = \min_{x \in \mathcal{Q}} \{ c^{\top} x \} = \min_{x \in \mathbb{R}^n} \{ c^{\top} x : A' x \ge b', A'' x \ge b'' \}$$

- Decomposition approaches attempt to improve on this bound by utilizing our implicit knowledge of \mathcal{P}' .
- Express the constraints of Q'' explicitly.
- ullet Express the constraints of \mathcal{P}' implicitly through solution of a subproblem.
 - Dantzig-Wolfe Decomposition
 - Lagrangian Relaxation
 - Cutting Plane Method

Dantzig-Wolfe Decomposition

The bound is obtained by solving the Dantzig-Wolfe LP:

$$z_{DW} = \min_{\lambda \in \mathbb{R}_{+}^{\mathcal{F}'}} \{ c^{\top} (\sum_{s \in \mathcal{F}'} s \lambda_s) : A''(\sum_{s \in \mathcal{F}'} s \lambda_s) \ge b'', \sum_{s \in \mathcal{F}'} \lambda_s = 1 \}, \tag{1}$$

- Solution method: simplex algorithm with dynamic column generation
- Subproblem: optimization over P'
- Suppose $\hat{\lambda}$ is an optimal solution to (1) then $z_{IP} \geq z_{DW} = c^{\top} \hat{x} \geq z_{LP}$, where

$$\hat{x} = \sum_{s \in \mathcal{F}'} s \hat{\lambda}_s \in \mathcal{P}' \tag{2}$$

Lagrangian Relaxation

The bound is obtained by solving the Lagrangian dual.

$$z_{LR}(u) = \min_{x \in \mathcal{P}'} \{ (c^{\top} - u^{\top} A'') x + u^{\top} b'' \}$$
 (3)

$$z_{LD} = \max_{u \in \mathbb{R}_+^{m''}} \{ z_{LR}(u) \} \tag{4}$$

- Solution method: subgradient optimization
- Subproblem: optimization over P'
- Rewriting z_{LD} as an LP we see it is dual to the Dantzig-Wolfe LP

$$z_{LD} = \max_{\alpha \in \mathbb{R}, u \in \mathbb{R}_{+}^{m''}} \{ \alpha + u^{\top}b'' : \alpha \le (c^{\top} - u^{\top}A'')s \ \forall s \in \mathcal{F}' \}$$
 (5)

• So we have $z_{IP} > z_{LD} = z_{DW} > z_{LP}$.

Cutting Plane Methods

- ullet The bound is obtained by augmenting the initial LP relaxation with facets of \mathcal{P}' .
- This approach yields the bound

$$z_{CP} = \min_{x \in \mathcal{P}'} \{ c^{\top} x : A'' x \ge b'' \}$$
 (6)

- Solution method: simplex with dynamic cut generation
- Subproblem: separation from P'
- Note that \hat{x} from (2) is an optimal solution to (6), so $z_{IP} \geq z_{CP} = z_{DW} \geq z_{LP}$.

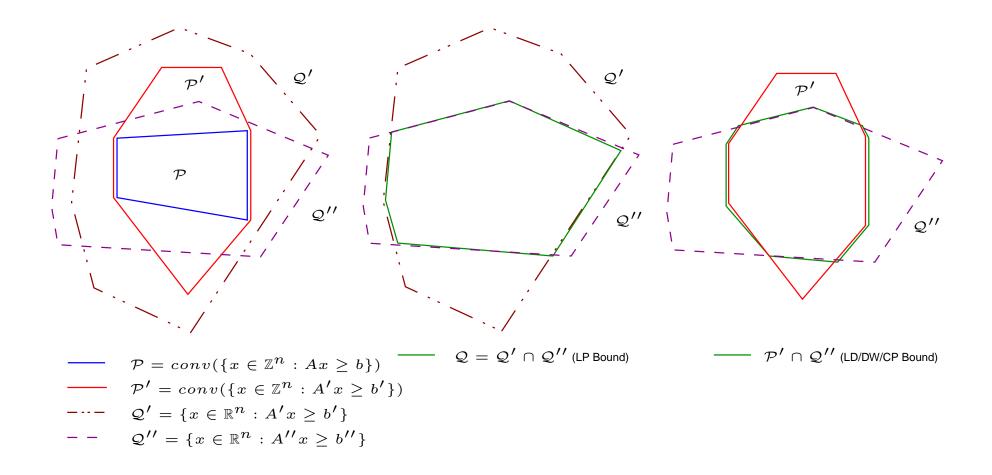
A Common Framework

All three decomposition methods compute the same quantity [Geoffrion74].

$$z_{IP} \ge c^{\top} \hat{x} = z_{LD} = z_{DW} = z_{CP} \ge z_{LP}$$

- The basic ingredients are the same:
 - the original polyhedron (\mathcal{P}) ,
 - an implicit polyhedron (\mathcal{P}') , and
 - an explicit polyhedron (Q'').
- The essential difference is how the implicit polyhedron is represented:
 - CP: as the intersection of half-spaces (the outer representation), or
 - DW/LD : as the convex hull of a finite set (inner representation).

Polyhedra, LP Bound, LD/DW/CP Bound



Outline

- Preliminaries, Traditional Decomposition Methods
 - Dantzig-Wolfe Decomposition
 - Lagrangian Relaxation
 - Cutting Plane Method
- Dynamic Decomposition Methods
 - Price and Cut
 - Relax and Cut
 - Decompose and Cut
- Applications/Examples
- DECOMP Library Framework

Cutting Plane Method (CPM)

• Goal: Improve the bound $\min_{x \in \mathcal{P}'} \{cx : A''x \ge b''\}$ by dynamic tightening of the explicit polyhedron (\mathcal{Q}'') .

Cutting Plane Method

- 1. Construct the initial LP relaxation LP⁰ and set $i \leftarrow 0$. $z_{LP} = \min_{x \in \mathbb{R}^n} \{ c^\top x : A'x \ge b', A''x \ge b'' \}$
- 2. Solve LPⁱ to obtain an optimal solution \hat{x}^i and lower bound $z^i \leftarrow c^{\top} \hat{x}^i$.
- 3. Attempt to separate \hat{x}^i from \mathcal{P} , generating violated inequalities $[D^i, d^i]$.
- 4. If $[D^i, d^i] \neq \emptyset$, set $[A'', b''] \leftarrow \begin{bmatrix} A'' & b'' \\ D^i & d^i \end{bmatrix}, i \leftarrow i + 1$ and go to Step 2.
- 5. If $[D^i, d^i] = \emptyset$, then output z^i .
- Step 3 may generate facets of any number of polyhedra $\bar{\mathcal{P}} \subseteq \mathcal{P}$.
- In principle, there are analogs of this for DW and LR.

Dynamic Decomposition Methods

Dynamic Decomposition Method

1. Construct the initial bounding subproblem P^0 and set $i \leftarrow 0$.

$$z_{DW} = \min_{\lambda \in \mathbb{R}_{+}^{\mathcal{F}'}} \{ c^{\top} (\sum_{s \in \mathcal{F}'} s \lambda_{s}) : A'' (\sum_{s \in \mathcal{F}'} s \lambda_{s}) \ge b'', \sum_{s \in \mathcal{F}'} \lambda_{s} = 1 \}$$

$$z_{LD} = \max_{u \in \mathbb{R}_{+}^{n}} \min_{x \in \mathcal{P}'} \{ (c^{\top} - u^{\top} A'') x + u^{\top} b'' \}$$

$$z_{CP} = \min_{x \in \mathcal{P}'} \{ c^{\top} x : A'' x \ge b'' \}$$

- 2. Solve P^i to obtain a lower bound z^i .
- 3. Attempt to generate a set of improving inequalities $[D^i, d^i]$.
- 4. If $[D^i, d^i] \neq \emptyset$, set $[A'', b''] \leftarrow \begin{bmatrix} A'' & b'' \\ D^i & d^i \end{bmatrix}, i \leftarrow i + 1$ and go to Step 2.
- 5. If $[D^i, d^i] = \emptyset$, then output z^i .

Price and Cut (PC)

• Price and Cut: use DW as the bounding subproblem and attempt to separate \hat{x}

$$z_{DW} = \min_{\lambda \in \mathbb{R}_{+}^{\mathcal{F}'}} \{ c^{\top} (\sum_{s \in \mathcal{F}'} s \lambda_s) : A''(\sum_{s \in \mathcal{F}'} s \lambda_s) \ge b'', \sum_{s \in \mathcal{F}'} \lambda_s = 1 \}$$

Theorem 1 Let F be the face of optimal solutions to the cutting plane LP. Then $(a, \beta) \in \mathbb{R}^{n+1}$ is an improving inequality if and only if $a^\top y < \beta$ for all $y \in F$.

Corollary 1 If $(a, \beta) \in \mathbb{R}^{n+1}$ is an improving inequality and \hat{x} is an optimal solution to the current LP relaxation, then $a^{\top}\hat{x} < \beta$.

- Generation of the cuts takes place in original space which maintains the structure of the column generation subproblem.
- **PC** and CPM: Corollary 1 means if we cut off \hat{x} we will probably improve the bound.
- **PC** vs CPM: empirical, optimization over \mathcal{P}' vs separation over \mathcal{P}'

Relax and Cut (RC)

• Relax and Cut: use LD as the bounding subproblem and attempt to separate \hat{s} .

$$z_{LD} = \max_{u \in \mathbb{R}^n_+} \min_{x \in \mathcal{P}'} \{ (c^{\top} - u^{\top} A'') x + u^{\top} b'' \}$$

- Solving LD with subgradient optimization no access to original primal solution \hat{x} .
- ullet Limited information from optimal primal solution to LD: $\hat{s} \in \mathcal{F}'$.
- Advantage: It is often much easier to separate a member of \mathcal{F}' from \mathcal{P} than an arbitrary real vector.
- Disadvantage: There is no way to verify the condition in Corollary 1.
- Questions:
 - What are the chances of generating an improving inequality?
 - Can we characterize the relationship between \hat{s} and \hat{x} ?

Some Useful Results

■ The set of alternative optimal primal solutions to LD is $S \cap \mathbb{Z}^n$, where S is the face of \mathcal{P}' defined as

$$S = \{ x \in \mathcal{P}' : (c^{\top} - \hat{u}^{\top} A'') x = (c^{\top} - \hat{u}^{\top} A'') \hat{s} \}$$

and \hat{s} is any optimal primal solution to the Lagrangian dual.

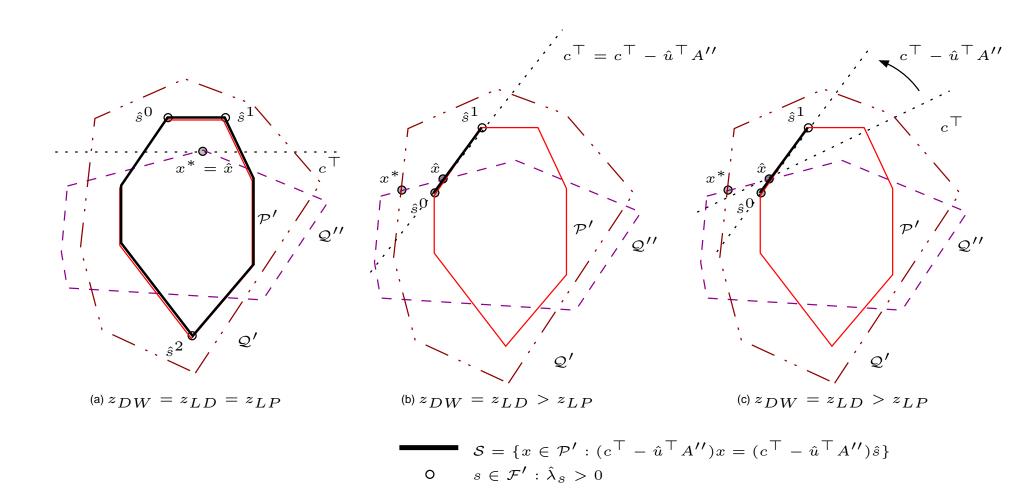
Theorem 2
$$D = \{s \in \mathcal{F}' : \hat{\lambda}_s > 0\} \subseteq \mathcal{S} \cap \mathbb{Z}^n$$

• If $\hat{\lambda}$ is an optimal solution the DW-LP, any $s \in \mathcal{F}'$ such that $\hat{\lambda}_s > 0$ is an optimal primal solution for the Lagrangian dual. Also $\hat{x} \in \mathcal{S}$.

Theorem 3 If \hat{x} is an inner point of \mathcal{P} , then $\mathcal{S} = \mathcal{P}'$.

• If \hat{x} is an inner point of \mathcal{P}' , then $\hat{\alpha} = 0$ (dual of DW-LP convexity constraint) and all members of \mathcal{F}' are optimal for LD.

Illustration of Results



Price and Cut (revisited)

Theorem 4 If $(a, \beta) \in \mathbb{R}^{(n+1)}$ is an improving inequality, then there must exist an $s \in \mathcal{F}'$ with $\hat{\lambda}_s > 0$ such that $a^{\top} s < \beta$.

- PC vs CPM: Theorem 4 tells us that knowledge of the optimal decomposition D should help us generate improving inequalities.
- Idea: Rather than (or in addition to) separating \hat{x} , we separate each $s \in D$.
- Recall: It is often much easier to separate a member of \mathcal{F}' from \mathcal{P} than an arbitrary real vector.
- **PC** vs RC: RC only gives us one member S, while PC gives us $D \subseteq S$.

Decompose and Cut (DC)

● Idea: Using a standard CPM framework - given a fractional point \hat{x} compute the decomposition $\hat{\lambda}$, then separate each $s \in D$ as in PC (*inverse DW*).

$$z_{CP} = \min_{x \in \mathcal{P}'} \{ c^{\top} x : A'' x \ge b'' \}$$

- **PC** and DC: Both allow us to take advantage of the information we gain from D and the fact that separation of members of \mathcal{F}' is easy.
- PC vs DC: DC can be more efficient than PC since we only compute the decomposition when standard CPM separation fails.

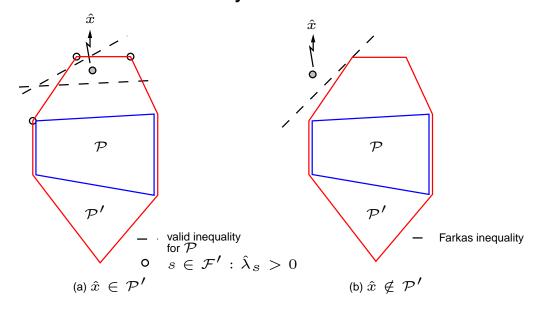
Decompose and Cut

Separation in Decompose and Cut

1. Attempt to decompose \hat{x} into a convex combination of members of \mathcal{F}' by solving the LP:

$$\max_{\lambda \in \mathbb{R}_{+}^{\mathcal{F}'}} \{ \mathbf{0}^{\top} \lambda : \sum_{s \in \mathcal{F}'} s \lambda_{s} = \hat{x}, \sum_{s \in \mathcal{F}'} \lambda_{s} = 1 \}, \tag{7}$$

- **2.1** If (7) is feasible, set $D = \{s \in \mathcal{F}' : \hat{\lambda}_s > 0\}$
- **2.2** Else, return a *Farkas Cut* (a, β) valid for $\mathcal{P}' \subseteq \mathcal{P}$ which violates \hat{x} .
- **3.** Separate each $s \in D$ and return any cuts that also violate \hat{x} .



Decompose and Cut

Column Generation in Decompose and Cut

- **1.0** Generate an initial subset \mathcal{G} of \mathcal{F}' .
- 1.1 Solve (7) over \mathcal{G} using the dual simplex algorithm.
- **1.2a** If (7) is feasible, return $D = \{s \in \mathcal{F}' : \hat{\lambda}_s > 0\}$.
- **1.2b** Else, optimize over \mathcal{P}' using the resulting Farkas inequality (row of B^{-1}). If the result has negative reduced cost, add it to \mathcal{G} and go to Step 1.1, else return the Farkas inequality.

Outline

- Preliminaries, Traditional Decomposition Methods
 - Dantzig-Wolfe Decomposition
 - Lagrangian Relaxation
 - Cutting Plane Method
- Dynamic Decomposition Methods
 - Price and Cut
 - Relax and Cut
 - Decompose and Cut
- Applications/Examples
- DECOMP Library Framework

Vehicle Routing Problem

ILP Formulation:

$$\sum_{e \in \delta(0)} x_e = 2k$$

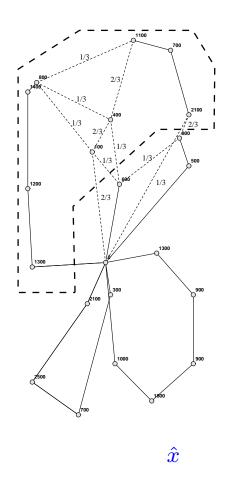
$$\sum_{e \in \delta(i)} x_e = 2 \qquad \forall i \in V \setminus \{0\}$$

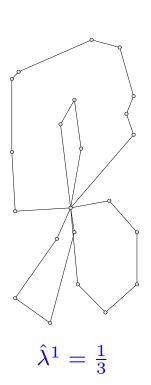
$$\sum_{e \in \delta(S)} x_e \geq 2b(S) \quad \forall S \subset V \setminus \{0\}, |S| > 1$$
(3)

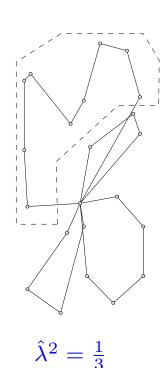
- b(S) = lower bound on the number of trucks required to service S = $\left[\left(\sum_{i \in S} d_i\right)/C\right]$ (normally)
- Relaxations:
 - Multiple Traveling Salesman Problem: Set $C = \sum_{i \in S} d_i$.
 - k-Tree: Set $C = \sum_{i \in S} d_i$. Relax (2) but leave $\sum_{e \in E} x_e = n + k$.
- Facets of VRP (under certain conditions): GSECs (3), Combs, Multistars
- Decompose and Cut VRP/kTSP for GSECs [Ralphs, et al. On the Capacitated Vehicle Routing Problem, Mathematical Programming 03]

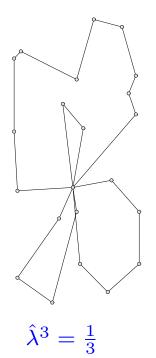
Example of Decomposition VRP/k-TSP

- ullet Optimization over kTSP can be done efficiently TSP
- Separation of \hat{x} for GSECs \mathcal{NP} -Complete
- Separation of a $kTSP \in \mathcal{F}'$ for GSECs in O(n)



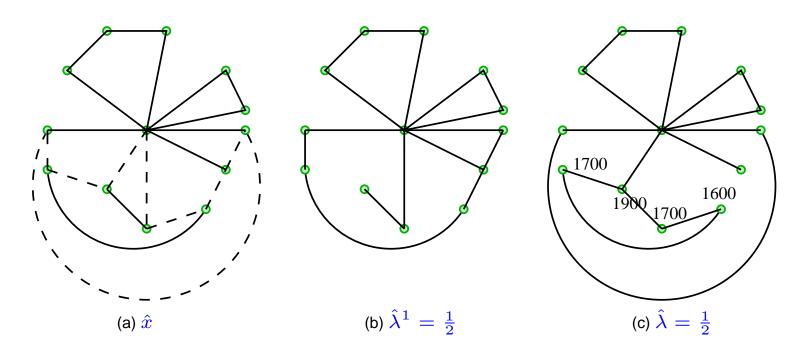






Example of Decomposition VRP/k-Tree

- Optimization over kTree in $O(n^2 \log n)$ [Wei and Yu]
- Separation of \hat{x}
 - for GSECs NP-Complete
 - for Combs and Multistars is difficult
- Separation of a $kTree \in \mathcal{F}'$
 - for GSECs in O(n)
 - for Combs and Multistars can be done efficiently [Martinhon, et al.]



Axial Assignment Problem

PILP Formulation:

$$\min \sum_{(i,j,k)\in T} c_{ijk} x_{ijk}
\sum_{(j,k)\in J\times K} x_{ijk} = 1 \quad \forall i\in I
\sum_{(i,k)\in I\times K} x_{ijk} = 1 \quad \forall j\in J
\sum_{(i,j)\in I\times J} x_{ijk} = 1 \quad \forall k\in K
x_{ijk} \in \{0,1\} \qquad \forall (i,j,k)\in T = I\times J\times K$$
(4)

- Relaxation: Assignment Problem relax (1)
- Facets of AAP: $Q_1(u)$ and $P_1(u,v)$ cliques of the intersection graph of $K_{n,n,n}$

▶ Let
$$C(u) = \{w \in T : |u \cap w| = 2\}, C(u, v) = \{w \in T : |u \cap w| = 1, |w \cap v| = 2\}$$

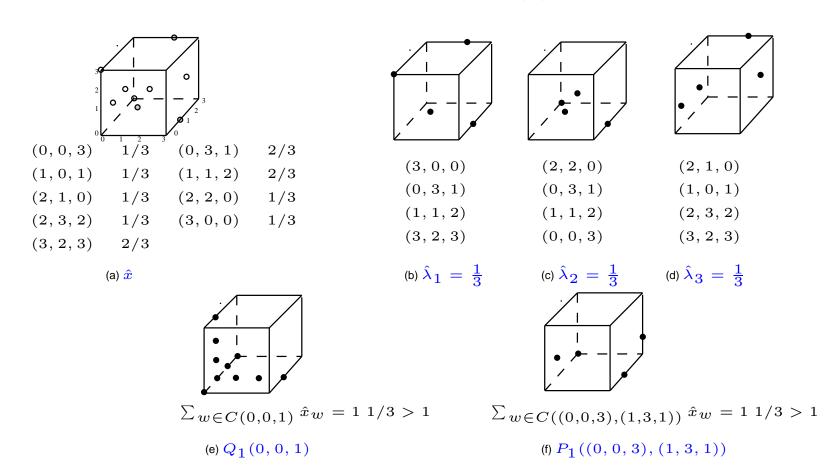
$$x_u + \sum_{w \in C(u)} x_w \leq 1 \quad \forall u \in T \qquad (5)$$

$$x_u + \sum_{w \in C(u, v)} x_w \leq 1 \quad \forall u, v \in T, u \cap v = \emptyset \qquad (6)$$

■ Relax and Cut - AP3/AP for Q_1 [Balas and Saltzman, An Algorithm for the Three-Index Assignment Problem Operations Research 91]

Example of Decomposition AAP/AP

- Optimization over AP in $O(n^{5/2} \log(nC))$
- Separation of \hat{x} for Clique Facets in $O(n^3)$
- Separation of an $AP \in \mathcal{F}'$ for Clique Facets in O(n)



Outline

- Preliminaries, Traditional Decomposition Methods
 - Dantzig-Wolfe Decomposition
 - Lagrangian Relaxation
 - Cutting Plane Method
- Dynamic Decomposition Methods
 - Price and Cut
 - Relax and Cut
 - Decompose and Cut
- Applications/Examples
- DECOMP Library Framework

DECOMP Library Framework

- Goal: Framework to allow for direct comparison of all three dynamic decomposition methods.
- COIN-or: Computational Infrastructure for Operations Research
- BCP: Parallel Branch, Price and Cut (LP-based Bounding) [Ladányi, Ralphs]
- ALPs: Abstract Library for Parallel Search [Ladányi, Ralphs, Saltzman]
 - BiCePS: Branch, Constrain and Price Software (Generic Bounding)
 - BLIS: BiCePS Linear Integer Solver = BCP
- DECOMP provides
 - CGL-like full implementation of Decompose and Cut
 - BiCePS plug-and-play for Price and Cut and Relax and Cut
- DECOMP user simply derives two methods:
 - solve_relaxed_problem (includes several built-in solvers)
 - separate_relaxed_point

Decompose and Cut Implementation Details

- Initialization of \mathcal{G} : solve over \mathcal{P}' with $c=-\hat{x}^{\epsilon}$.
- Active LP column management reduced cost fixing.
- Lifting the Farkas inequality $(\hat{x} \notin \mathcal{P}')$.
- Consistency Condition restriction of column generation search
 - $\hat{x}_i = 0 \Rightarrow s_i = 0, \forall s \in D$
 - $\hat{x}_i = 1 \Rightarrow s_i = 1, \forall s \in D$
- Is it necessary to be exact in solving the column generation subproblem?
 - Try optimizing over \mathcal{P}' heuristically first need negative reduced cost.
 - Do we necessarily want extreme points of \mathcal{P}' ?
- Decomposition into members of F [Kopman 99]
 - Column generation subproblem is an optimization problem over P!!
 - Applegate, Bixby, Chvátal, and Cook, TSP Cuts Which Do Not Conform to the Template Paradigm, Computational Combinatorial Optimization 2001

Applications Under Development

- Vehicle Routing Problem
 - k-Traveling Salesman Problem : GSECs
 - k-Tree : GSECs, Combs, Multistars
- Axial Assignment Problem
 - Assignment Problem : Clique-Facets
- Steiner Problem in Graphs
 - Minimum Spanning Tree : Lifted SECs
- Knapsack Constrained Circuit Problem
 - Knapsack Problem : Maximal-Set Inequalities
- Edge-Weighted Clique Problem
 - Tree Relaxation : Trees, Cliques
- Traveling Salesman Problem [Labonte/Boyd]
 - Fractional 2-Factor Problem : SECs

Conclusions

- Provided some insight into the relationship between: the optimal LP face F, the optimal DW solution \hat{x} , the optimal LD solution \hat{s} and the knowledge gained from the optimal decomposition $\hat{\lambda}$.
- Alternative (and often much easier) methods for separation: over \mathcal{F}' vs \mathcal{Q} .
 - Incorporated this idea into traditional Price and Cut.
 - Introduced a promising new paradigm for separation Decompose and Cut.
- Presented a unifying framework for dynamic cut generation in traditional decomposition methods.
 - We are currently in the process of developing a software framework DECOMP to implement and directly compare each of these methods.