Recent Progress in Approximation Algorithms for the Traveling Salesman Problem

Lecture 3: The s-t path TSP

David P. Williamson
Cornell University

July 18-22, 2016
São Paulo School of Advanced Science on
Algorithms, Combinatorics, and Optimization

s-t path TSP

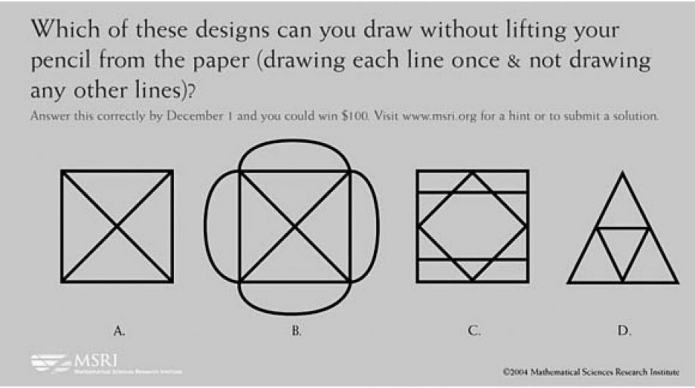
Recall the s-t path TSP:
Usual TSP input plus $s, t \in V$, find a min-cost path from s to t visiting all other nodes in between (an s-t *Hamiltonian path*).

Eulerian path

There is an Eulerian path that starts at s, ends at t, and visits every edge exactly once iff s and t have odd-degree and all other vertices have even degree.

Eulerian path

There is an Eulerian path that starts at s, ends at t, and visits every edge exactly once iff s and t have odd-degree and all other vertices have even degree.

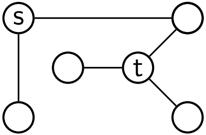


Which of these designs can you draw without lifting your pencil from the paper (drawing each line once & not drawing any other lines)?

Answer this correctly by December 1 and you could win $100. Visit www.msri.org for a hint or to submit a solution.

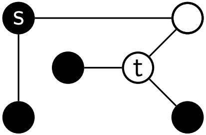
Hoogeveen's algorithm

Let F be the min-cost spanning tree. Let T be the set of vertices whose *parity needs changing*: s iff s has even degree in F, t iff t has even degree in F, and $v \neq s, t$ iff v has odd degree. Then find a minimum-cost perfect matching M on the vertices in T. Find Eulerian path on $F \cup M$; shortcut to an s-t Hamiltonian path.



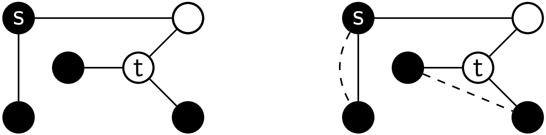
Hoogeveen's algorithm

Let F be the min-cost spanning tree. Let T be the set of vertices whose *parity needs changing*: s iff s has even degree in F, t iff t has even degree in F, and $v \neq s, t$ iff v has odd degree. Then find a minimum-cost perfect matching M on the vertices in T. Find Eulerian path on $F \cup M$; shortcut to an s-t Hamiltonian path.



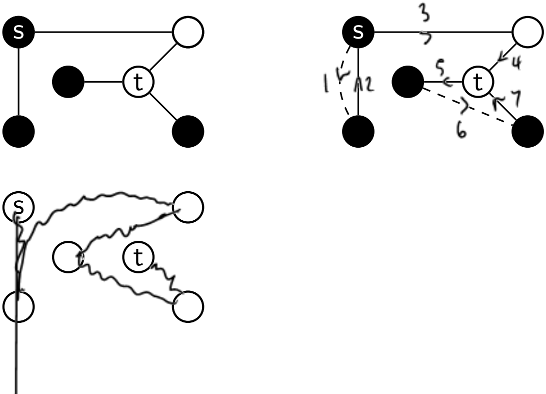
Hoogeveen's algorithm

Let F be the min-cost spanning tree. Let T be the set of vertices whose *parity needs changing*: s iff s has even degree in F, t iff t has even degree in F, and $v \neq s, t$ iff v has odd degree. Then find a minimum-cost perfect matching M on the vertices in T. Find Eulerian path on $F \cup M$; shortcut to an s-t Hamiltonian path.



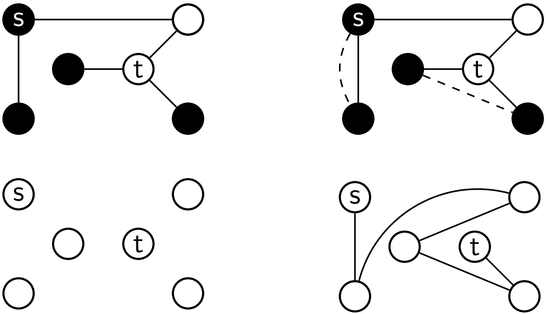
Hoogeveen's algorithm

Let F be the min-cost spanning tree. Let T be the set of vertices whose *parity needs changing*: s iff s has even degree in F, t iff t has even degree in F, and $v \neq s, t$ iff v has odd degree. Then find a minimum-cost perfect matching M on the vertices in T. Find Eulerian path on $F \cup M$; shortcut to an s-t Hamiltonian path.



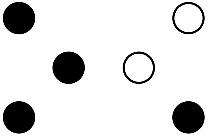
Hoogeveen's algorithm

Let F be the min-cost spanning tree. Let T be the set of vertices whose *parity needs changing*: s iff s has even degree in F, t iff t has even degree in F, and $v \neq s, t$ iff v has odd degree. Then find a minimum-cost perfect matching M on the vertices in T. Find Eulerian path on $F \cup M$; shortcut to an s-t Hamiltonian path.



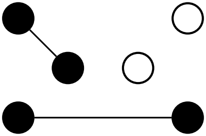
T-joins

Rather than a minimum-cost perfect matching on T, will construct a minimum-cost *T-join*: a set of edges that has odd degree at every vertex in T, even degree at every vertex not in T.



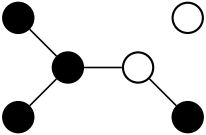
T-joins

Rather than a minimum-cost perfect matching on T, will construct a minimum-cost *T-join*: a set of edges that has odd degree at every vertex in T, even degree at every vertex not in T.



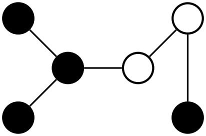
T-joins

Rather than a minimum-cost perfect matching on T, will construct a minimum-cost *T-join*: a set of edges that has odd degree at every vertex in T, even degree at every vertex not in T.



T-joins

Rather than a minimum-cost perfect matching on T, will construct a minimum-cost *T-join*: a set of edges that has odd degree at every vertex in T, even degree at every vertex not in T.



Hoogeveen's algorithm

Let F be the min-cost spanning tree. Let T be the set of vertices whose parity needs changing. Then find a minimum-cost T-join J. Find Eulerian path on $F \cup J$; shortcut to an s-t Hamiltonian path.

Hoogeveen's algorithm

Let F be the min-cost spanning tree. Let T be the set of vertices whose parity needs changing. Then find a minimum-cost T-join J. Find Eulerian path on $F \cup J$; shortcut to an s-t Hamiltonian path.

Theorem

Hoogeveen's algorithm is a $\frac{5}{3}$-approximation algorithm.

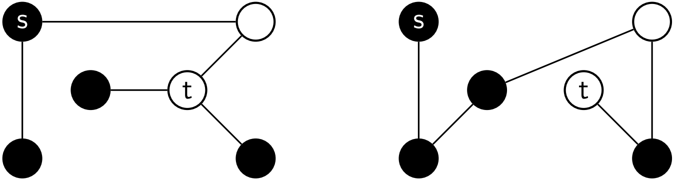
Proof of theorem

Idea: Construct 3 T-joins from the edges of the spanning tree F and the optimal s-t TSP path.

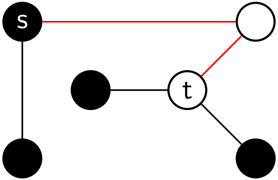
Since $c(F) \leq OPT$, then cost of min-cost T-join is at most $\frac{1}{3}(c(F) + OPT) \leq \frac{2}{3} OPT$.

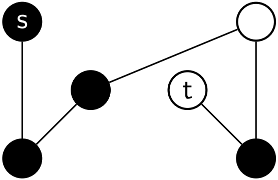
$c(F) +$ cost of T-join $\leq OPT + \frac{2}{3} OPT \leq \frac{5}{3} OPT$.

Proof of theorem

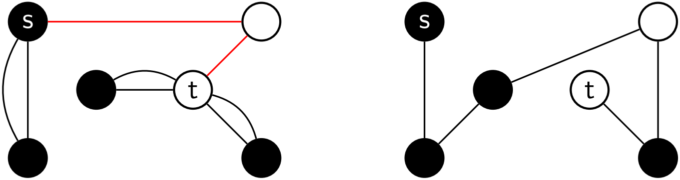


Proof of theorem

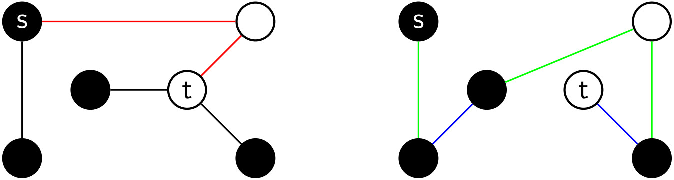




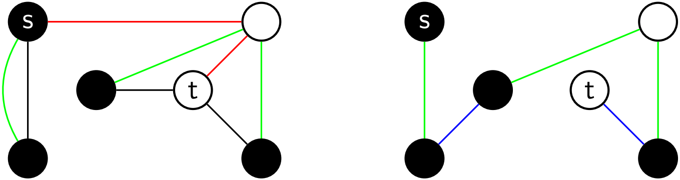
Proof of theorem



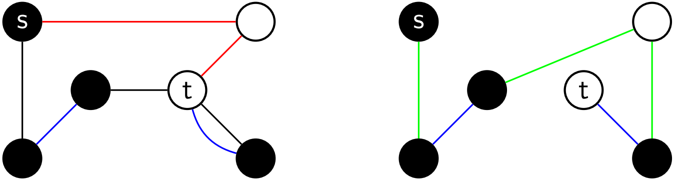
Proof of theorem



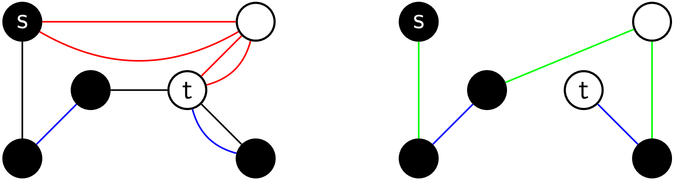
Proof of theorem



Proof of theorem

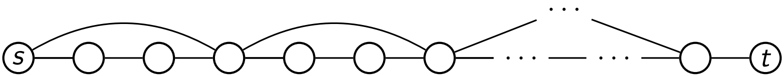


Proof of theorem



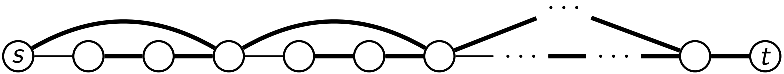
Tight Example

The analysis is tight. Consider the *graph TSP* instance below: cost $c(e)$ for $e = (i, j)$ is number of edges in shortest i-j path in graph.



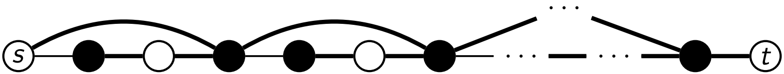
Tight Example

The analysis is tight. Consider the *graph TSP* instance below: cost $c(e)$ for $e = (i, j)$ is number of edges in shortest i-j path in graph.



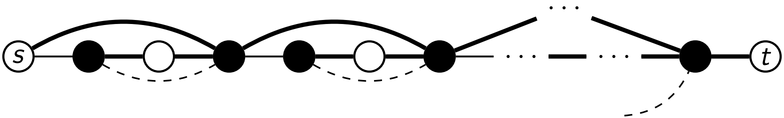
Tight Example

The analysis is tight. Consider the *graph TSP* instance below: cost $c(e)$ for $e = (i, j)$ is number of edges in shortest i-j path in graph.



Tight Example

The analysis is tight. Consider the *graph TSP* instance below: cost $c(e)$ for $e = (i, j)$ is number of edges in shortest i-j path in graph.



Improvements

No improvement on Hoogeveen's algorithm for *s*-*t* path TSP, until just the last few years.

An, Kleinberg, Shmoys	2012	1.618
Sebő	2013	1.6
Vygen	2015	1.599
Gottschalk and Vygen	2015	1.56
Sebő and Van Zuylen	2016	1.52

Goal: Understand the An et al. algorithm and analysis; will sketch some of the ideas of the improvements.

A Linear Programming Relaxation

$$\text{Min} \quad \sum_{e \in E} c_e x_e$$

subject to:

$$x(\delta(v)) = \begin{cases} 1 & \text{if } v \in s,t \\ 2 & \text{if } v \notin s,t \end{cases}$$

$$x(\delta(S)) \geq \begin{cases} 1 & \text{if } |S \cap \{s,t\}| = 1 \\ 2 & \text{if } |S \cap \{s,t\}| \neq 1, \quad S \neq \emptyset, V \end{cases}$$

$$0 \leq x_e \leq 1, \qquad \forall e \in E,$$

where $\delta(S)$ is the set of edges with exactly one endpoint in S, and $x(E') \equiv \sum_{e \in E'} x_e$.

A Linear Programming Relaxation

$$\text{Min} \quad \sum_{e \in E} c_e x_e$$

subject to:

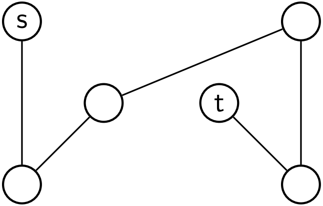
$$x(\delta(v)) = \begin{cases} 1, & v = s, t, \\ 2, & v \neq s, t, \end{cases}$$

$$x(\delta(S)) \geq \begin{cases} 1, & |S \cap \{s, t\}| = 1, \\ 2, & |S \cap \{s, t\}| \neq 1, \end{cases}$$

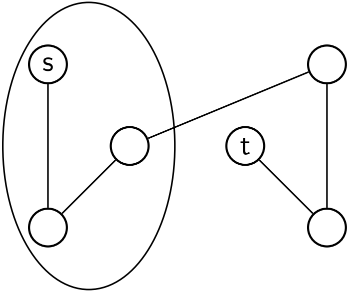
$$0 \leq x_e \leq 1, \qquad \forall e \in E,$$

where $\delta(S)$ is the set of edges with exactly one endpoint in S, and $x(E') \equiv \sum_{e \in E'} x_e$.

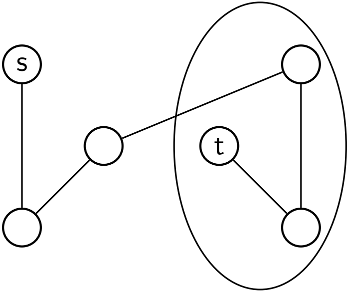
LP relaxation



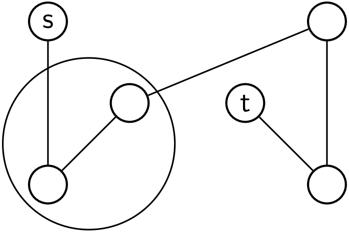
LP relaxation



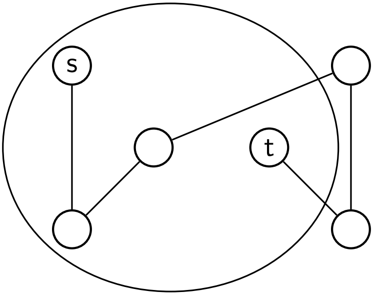
LP relaxation



LP relaxation



LP relaxation

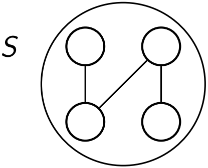


The spanning tree polytope

The spanning tree polytope (convex hull of all spanning trees) is defined by the following inequalities:

$$\begin{aligned} & x(E) = |V| - 1, & \\ & x(E(S)) \leq |S| - 1, & \forall |S| \subseteq V, |S| \geq 2, \\ & x(e) \geq 0, & \forall e \in E, \end{aligned}$$

where $E(S)$ is the set of all edges with both endpoints in S.



The LP relaxation and spanning trees

Lemma

Any solution x feasible for the s-t path TSP LP relaxation is in the spanning tree polytope.

Proof

$$x(\delta(v)) = \begin{cases} 1, & v = s,t, \\ 2, & v \neq s,t, \end{cases}$$

$$x(\delta(S)) \geq \begin{cases} 1, & |S \cap \{s,t\}| = 1, \\ 2, & |S \cap \{s,t\}| \neq 1, \end{cases}$$

$0 \leq x(e) \leq 1, \quad \forall e \in E.$

$$\xrightarrow{\text{feasible}}$$

$x(E) = |V| - 1,$ ✓

$x(E(S)) \leq |S| - 1, \quad \forall |S| \subseteq V, |S| \geq 2,$ ✓

$x(e) \geq 0, \quad \forall e \in E.$ ✓

$$x(E) = \sum_{e \in E} x_e = \frac{1}{2} \sum_{v \in V} x(\delta(v)) = \frac{1}{2}\left(2(|V|-2)+2\right) = \frac{1}{2}(2|V|-2) = |V|-1.$$

Spse S s.t. $S \cap \{s,t\} = \emptyset$, $S \neq \emptyset, V$.

$$x(E(S)) = \frac{1}{2}\left(\sum_{v \in V} x(\delta(v)) - x(\delta(S))\right) \leq \frac{1}{2}(2|S|-2) = |S|-1.$$ ✓

A warmup to the improvements

Let OPT_{LP} be the value of an optimal solution x^* to the LP relaxation.

Theorem (An, Kleinberg, Shmoys (2012))

Hoogeveen's algorithm returns a solution of cost at most $\frac{5}{3}OPT_{LP}$.

An extremely useful lemma

Let F be a spanning tree, and let T be the vertices whose parity needs fixing in F.

Definition

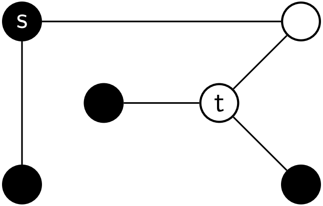
S is an *odd set* if $|S \cap T|$ is odd.

Lemma

Let S be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.

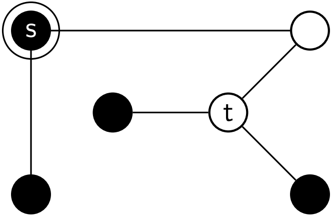
Lemma

Let S be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.



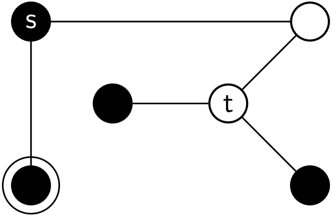
Lemma

Let S be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.



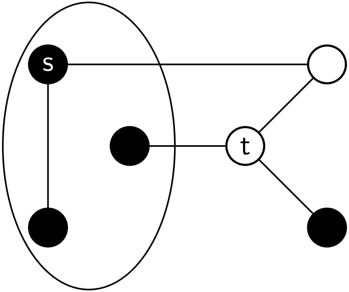
Lemma

Let S be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.



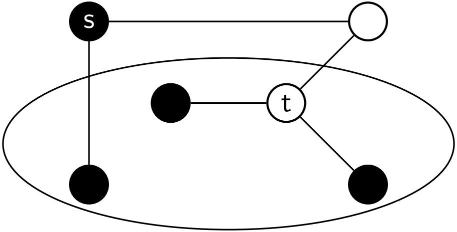
Lemma

Let S be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.



Lemma

Let S be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.



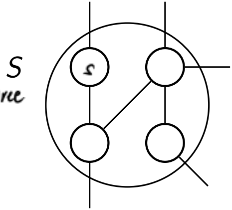
Proof of lemma

Case 1: $|S \cap \{s,t\}| = 1$ Spse $s \in S$. $s \in T$ iff $\deg_F(s)$ is even.

degree of s in spanning tree F

t

$\therefore$ S odd set $\Rightarrow$ even # of odd-degree vertices in S.



$\therefore |\delta(S) \cap F|$ is even.

$$\underbrace{\sum_{v \in S} deg_F(v)}_{\text{odd}} = \underbrace{2|E(S) \cap F|}_{\text{even}} + \underbrace{|\delta(S) \cap F|}_{\text{odd.}}$$

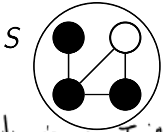
Case 2: $|S \cap \{s,t\}| \neq 1$. & S odd. $\Rightarrow$ odd # of odd degree vertices in S.

$\Rightarrow |\delta(S) \cap F|$ is odd.

T-join LP

The solution to the following linear program is the minimum-cost T-join for costs $c \geq 0$:

$$\begin{array}{lll} \text{Min} & \displaystyle\sum_{e\in E} c_e x_e & \\ \text{subject to:} & x(\delta(S)) \underline{\geq 1}, & \forall S \subseteq V, |\underline{S \cap T}| \text{ odd} \\ & x_e \geq 0, & \forall e \in E. \end{array}$$



If J is a soln, is a T-join

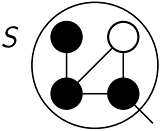
$S: |S\cap T|$ odd

$$\underset{\uparrow\ \text{odd}}{\sum_{v\in S}} deg_J(v) = \underset{\text{even}}{2|E(S)\cap J|} + \underset{\text{odd}}{|\delta(S)\cap J|}$$

T-join LP

The solution to the following linear program is the minimum-cost T-join for costs $c \geq 0$:

$$\begin{array}{lll} \text{Min} & \displaystyle\sum_{e \in E} c_e x_e & \\ \text{subject to:} & x(\delta(S)) \geq 1, & \forall S \subseteq V, |S \cap T| \text{ odd} \\ & x_e \geq 0, & \forall e \in E. \end{array}$$



$$\sum_{v \in S} deg_J(v) = 2|E(S) \cap J| + |\delta(S) \cap J|$$

Proof of theorem

Theorem (An, Kleinberg, Shmoys (2012))

Hoogeveen's algorithm returns a solution of cost at most $\frac{5}{3} OPT_{LP}$.

Lemma

Let S *be an odd set. If* $|S \cap \{s,t\}| = 1$, *then* $|F \cap \delta(S)|$ *is even. If* $|S \cap \{s,t\}| \neq 1$, *then* $|F \cap \delta(S)|$ *is odd.*

$$\begin{array}{lll} \text{Min} & \displaystyle\sum_{e \in E} c(e) x(e) & \\ & x(\delta(S)) \geq 1, & \forall S \subseteq V, |S \cap T| \text{ odd} \\ & x(e) \geq 0, & \forall e \in E. \end{array}$$

Pf Let x^* be an opt soln to LP. Feas. for spanning tree LP.

Let F be MST. Let χ_F be vector s.t. $\chi_F(e) = \begin{cases} 1 & \text{if } e \in F \\ 0 & \text{o.w.} \end{cases}$

Claim $y = \frac{1}{3}\chi_F + \frac{1}{3}x^*$ is feasible for T-join LP.

Then if J is min-cost T-join,

$$c(F) + c(J) \le c(F) + \frac{1}{3}c(F) + \frac{1}{3}\sum_{e\in E} c_e x_e^* \le \frac{5}{3}\sum_{e \in E} c_e x_e^* = \frac{5}{3} OPT_{LP}.$$

Pf of claim: If $|S \cap \{s,t\}| \neq 1$, S odd

$$y(\delta(S)) = \frac{1}{3}|F \cap \delta(S)| + \frac{1}{3}x^*(\delta(S)) \geq \frac{1}{3}\cdot 1 + \frac{1}{3}\cdot 2 = 1. \quad ✓$$

If $|S \cap \{s,t\}| = 1$, S odd

$$y(\delta(S)) = \frac{1}{3}|F \cap \delta(S)| + \frac{1}{3}x^*(\delta(S)) \geq \frac{1}{3}\cdot 2 + \frac{1}{3}\cdot 1 = 1. \quad ✓$$

Convex combination

Let x^* be an optimal LP solution. Let χ_F be the *characteristic vector* of a set of edges F, so that

$$\chi_F(e) = \begin{cases} 1 & e \in F \\ 0 & e \notin F \end{cases}$$

Since x^* is in the spanning tree polytope, can write x^* as a convex combination of spanning trees $F_1, \ldots, F_k$:

$$x^* = \sum_{i=1}^{k} \lambda_i \chi_{F_i},$$

such that $\sum_{i=1}^{k} \lambda_i = 1$, $\lambda_i \geq 0$.

Best-of-Many Christofides' Algorithm

An, Kleinberg, Shmoys (2012) propose the *Best-of-Many Christofides'* algorithm: given optimal LP solution x^*, compute convex combination of spanning trees

$$x^* = \sum_{i=1}^{k} \lambda_i \chi_{F_i}.$$

For each spanning tree F_i, let T_i be the set of vertices whose parity needs fixing, let J_i be the minimum-cost T_i-join. Find s-t Hamiltonian path by shortcutting $F_i \cup J_i$. Return the shortest path found over all i.

Best-of-Many Christofides' Algorithm

$$x^* = \sum_{i=1}^{k} \lambda_i \chi_{F_i}.$$

For each spanning tree F_i, let T_i be the set of vertices whose parity needs fixing, J_i be the minimum-cost T_i-join. Find s-t Hamiltonian path by shortcutting $F_i \cup J_i$. Return the shortest path found over all i.

Theorem

The Best-of-Many Christofides' algorithm returns a solution of cost at most $\frac{5}{3} OPT_{LP}$.

Proof

Pf Best s-t TSP path $\le \sum_{i=1}^{k} \lambda_i \left[c(F_i) + c(\overset{T_i\text{-join}}{J_i}) \right]$

$\frac{1}{3}\chi_{F_i} + \frac{1}{3}x^*$ is feasible soln. for T_i-join LP. So

$$\sum_{i=1}^{k} \lambda_i \left[c(F_i) + c(J_i) \right] \le \sum_{i=1}^{k} \lambda_i \left[c(F_i) + \frac{1}{3} c(F_i) + \frac{1}{3} \sum_{e \in E} c_e x_e^* \right]$$

$$= \sum_{i=1}^{k} \lambda_i \left(\frac{4}{3} \sum_{e \in E} c_e \chi_{F_i}(e) + \frac{1}{3} \sum_{e \in E} c_e x_e^* \right)$$

$$= \sum_{e \in E} c_e \left(\frac{4}{3} \underbrace{\sum_{i=1}^{k} \lambda_i \chi_{F_i}(e)}_{x_e^*} + \frac{1}{3} x_e^* \right)$$

$$= \frac{5}{3} \sum_{e \in E} c_e x_e^* \le \frac{5}{3} \cdot OPT_{LP}.$$

Improvement?

To do better, we need to improve the analysis for the costs of the T_i-joins; recall that we use that

$$y_i = \frac{1}{3}\chi_{F_i} + \frac{1}{3}x^*$$

is feasible for the T_i-join LP.

Consider

$$y_i = \alpha\chi_{F_i} + \beta x^*.$$

Then the cost of the best s-t Hamiltonian path is at most

$$(1 + \alpha + \beta)OPT_{LP}.$$

Improvement?

Proof that y_i feasible for T_i-join LP had two cases. Assume S odd ($|S \cap T_i|$ odd).

If $|S \cap \{s, t\}| \neq 1$, then

$$y_i(\delta(S)) = \alpha|F_i \cap \delta(S)| + \beta x^*(\delta(S)) \geq \alpha + 2\beta.$$

We will want $\alpha + 2\beta \geq 1$, so the T_i-join LP constraint is satisfied.

Improvement?

If $|S \cap \{s,t\}| = 1$, then

$$y_i(\delta(S)) = \alpha|F_i \cap \delta(S)| + \beta x^*(\delta(S)) \geq 2\alpha + \beta x^*(\delta(S)).$$

Improvement?

If $|S \cap \{s,t\}| = 1$, then

$$y_i(\delta(S)) = \alpha |F_i \cap \delta(S)| + \beta x^*(\delta(S)) \geq 2\alpha + \beta x^*(\delta(S)).$$

Since we assume $\alpha + 2\beta \geq 1$, we only run into problems if

$$x^*(\delta(S)) < \frac{1-2\alpha}{\beta}.$$

Note that $\alpha = 0$, $\beta = \frac{1}{2}$ works if $x^*(\delta(S)) \geq 2$ for all $S \subset V$, and gives a tour of cost at most $\frac{3}{2} OPT_{LP}$.

Improvement?

If $|S \cap \{s,t\}| = 1$, then

$$y_i(\delta(S)) = \alpha|F_i \cap \delta(S)| + \beta x^*(\delta(S)) \geq 2\alpha + \beta x^*(\delta(S)).$$

Since we assume $\alpha + 2\beta \geq 1$, we only run into problems if

$$x^*(\delta(S)) < \frac{1-2\alpha}{\beta}.$$

Note that $\alpha = 0$, $\beta = \frac{1}{2}$ works if $x^*(\delta(S)) \geq 2$ for all $S \subset V$, and gives a tour of cost at most $\frac{3}{2}OPT_{LP}$.

So focus on cuts for which $x^*(\delta(S)) < 2$, and add an extra "correction" term to y_i to handle these cuts.