Recent Progress in Approximation Algorithms for the Traveling Salesman Problem

Lecture 4: s-t path TSP for graph 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*).

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 $x(E') \equiv \sum_{e \in E'} x_e$.

Lemma

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

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.

From last time

To prove that Best-of-Many Christofides is at most $\frac{5}{3}OPT_{LP}$ for optimal LP solution x^*, show that

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

is feasible for the T_i-join LP:

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

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

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

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

τ-Narrow Cuts

Definition

S is *τ-narrow* if $x^*(\delta(S)) < 1 + \tau$ for fixed $\tau \leq 1$.

Only S such that $|S \cap \{s, t\}| = 1$ are τ-narrow.

Definition

Let $\mathcal{C}_\tau$ be all τ-narrow cuts S with $s \in S$.

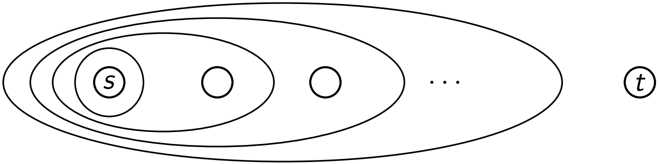
τ-Narrow Cuts

The τ-narrow cuts in $\mathcal{C}_\tau$ have a nice structure.

Theorem (An, Kleinberg, Shmoys (2012))

If $S_1, S_2 \in \mathcal{C}_\tau$, $S_1 \neq S_2$, then either $S_1 \subset S_2$ or $S_2 \subset S_1$.

So the τ-narrow cuts look like $s \in Q_1 \subset Q_2 \subset \cdots \subset Q_k \subset V$.



Correction Factor

Let e_Q be the minimum-cost edge in $\delta(Q)$. Then consider the following (from Gao (2014)):

$$y_i = \alpha\chi_{F_i} + \beta x^* + \sum_{Q \in \mathcal{C}_\tau, |Q \cap T_i| \text{ odd}} (1 - 2\alpha - \beta x^*(\delta(Q)))\, \chi_{e_Q}$$

for $\alpha, \beta, \tau \geq 0$ such that

$$\alpha + 2\beta = 1 \qquad \text{and} \qquad \tau = \frac{1 - 2\alpha}{\beta} - 1.$$

Theorem

y_i is feasible for the T_i-join LP.

Parameters

By choosing

$$\alpha = 1 - \frac{2}{\sqrt{5}}, \quad \beta = \frac{1}{\sqrt{5}}, \quad \tau = 3 - \sqrt{5},$$

then one can show that the total cost is at most

$$\frac{1+\sqrt{5}}{2} OPT_{LP}.$$

Today

Today: Combine the two special cases: Look at s-t path TSP in the case of graph TSP instances (e.g. input is undirected graph, cost $c(i,j)$ is number of edges in shortest i-j path)

Integrality Gap

The performance of Best-of-Many Christofides' cannot do better than the *integrality gap* of the LP relaxation.

Integrality Gap

The performance of Best-of-Many Christofides' cannot do better than the *integrality gap* of the LP relaxation.

The integrality gap is

$$\mu \equiv \sup \frac{OPT}{OPT_{LP}}$$

over all instances of the problem.

Integrality Gap

The performance of Best-of-Many Christofides' cannot do better than the *integrality gap* of the LP relaxation.

The integrality gap is

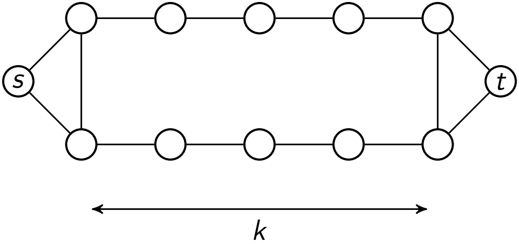
$$\mu \equiv \sup \frac{OPT}{OPT_{LP}}$$

over all instances of the problem.

Note that An, Kleinberg, Shmoys have shown $\mu \leq 1.618$, since Best-of-Many Christofides' algorithm finds a tour of cost at most $1.618 \cdot OPT_{LP}$.

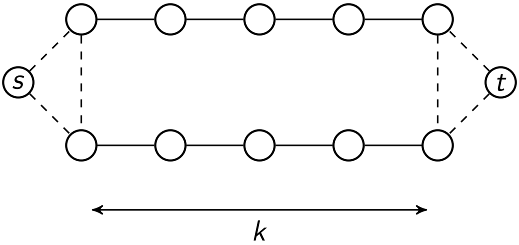
Integrality Gap

We can show a lower bound on the integrality gap using an instance of *graph TSP*: input is a graph $G = (V, E)$, cost c_e for $e = (i, j)$ is number of edges in a shortest i-j path in G.



Integrality Gap

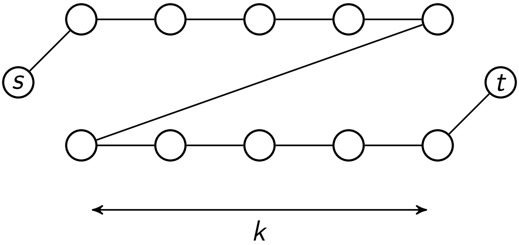
We can show a lower bound on the integrality gap using an instance of *graph TSP*: input is a graph $G = (V, E)$, cost c_e for $e = (i, j)$ is number of edges in a shortest i-j path in G.



$$OPT_{LP} \approx 2k$$

Integrality Gap

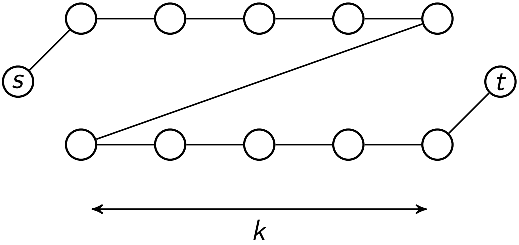
We can show a lower bound on the integrality gap using an instance of *graph TSP*: input is a graph $G = (V, E)$, cost c_e for $e = (i, j)$ is number of edges in a shortest i-j path in G.



$$OPT \approx 3k$$

Integrality Gap

We can show a lower bound on the integrality gap using an instance of *graph TSP*: input is a graph $G = (V, E)$, cost c_e for $e = (i, j)$ is number of edges in a shortest i-j path in G.



$$\frac{OPT}{OPT_{LP}} \to \frac{3}{2} \text{ as } k \to \infty$$

Graph Instances

Sebő and Vygen (2014) show that for graph TSP instances of s-t path TSP, can get a $\frac{3}{2}$-approximation algorithm (i.e. the algorithm produces a solution of cost at most $\frac{3}{2}OPT_{LP}$), so the integrality gap is tight for these instances.

We'll present a simplified version of this result due to Gao (2013).

Graph Instances

Given the input graph $G = (V, E)$ and an optimal solution, can replace any edge (i, j) in the optimal solution with the i-j path in G since these have the same cost.

So finding an optimal solution is equivalent to finding a multiset F of edges such that (V, F) is connected, $deg_F(s)$ and $deg_F(t)$ are odd, $deg_F(v)$ is even for all $v \in V - \{s, t\}$, and $|F|$ is minimum.

LP Relaxation

$$\text{Min} \quad \sum_{e \in E} x(e)$$

subject to:

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

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

Let x^* be an optimal LP solution.

Narrow Cuts

As before, focus on *narrow cuts* S such that $x^*(\delta(S)) < 2$ (i.e. a τ-narrow cut for $\tau = 1$). Recall:

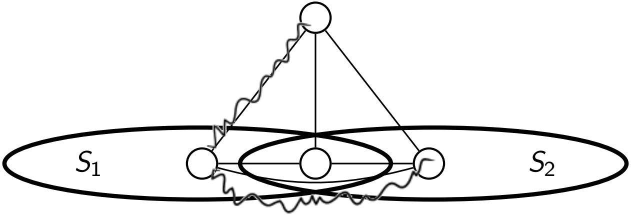
Theorem (An, Kleinberg, Shmoys (2012))

If S_1, S_2 are narrow cuts, $S_1 \neq S_2$, then either $S_1 \subset S_2$ or $S_2 \subset S_1$.

Proof

First need to show that

$$x^*(\delta(S_1)) + x^*(\delta(S_2)) \geq x^*(\delta(S_1 - S_2)) + x^*(\delta(S_2 - S_1)).$$



Proof

Theorem (An, Kleinberg, Shmoys (2012))

If S_1, S_2 are narrow cuts, $S_1 \neq S_2$, then either $S_1 \subset S_2$ or $S_2 \subset S_1$.

Spse S_1, S_2 and $S_1 \not\subseteq S_2$, $S_2 \not\subseteq S_1$ $\quad S_1 - S_2 \neq \emptyset$, $S_2 - S_1 \neq \emptyset$.

t

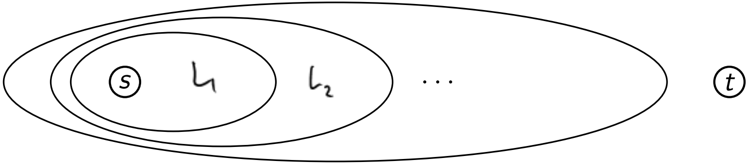
$$
\begin{aligned}
4 = 2+2 &> x^*(\delta(S_1)) + x^*(\delta(S_2)) \\
&\geq x^*(\delta(S_1 - S_2)) + x^*(\delta(S_2 - S_1)) \\
&\geq 2+2 = 4 \qquad \lightning
\end{aligned}
$$

Narrow Cuts

Theorem (An, Kleinberg, Shmoys (2012))

If S_1, S_2 are narrow cuts, $S_1 \neq S_2$, then either $S_1 \subset S_2$ or $S_2 \subset S_1$.

So the narrow cuts look like $s \in S_1 \subset S_2 \subset \cdots \subset S_k \subset V$.

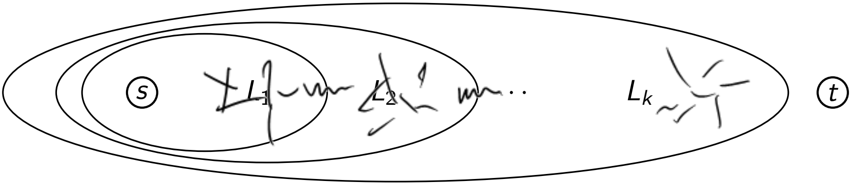


Let $S_0 \equiv \emptyset$, $S_{k+1} \equiv V$, $L_i \equiv S_i - S_{i-1}$.

Key Idea

Find a tree spanning L_i in the support of x^* for each i. Connect each of these via a single edge from L_i to L_{i+1}. Let F be the resulting tree, T the vertices in F whose parity needs changing.

Then $|F| = n - 1$ and $|\delta(S_i) \cap F| = 1$ for each narrow cut S_i.



Key Lemma

Recall:

Lemma

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

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

Lemma

$y = \frac{1}{2}x^*$ is feasible for the the T-join LP.

Proof

Let S be s.t. $|S \cap T|$ odd (S odd)

If $|S \cap \{s,t\}| \neq 1$. Then $x^*(\delta(S)) \geq 2 \Rightarrow y = \frac{1}{2}x^* \Rightarrow y(\delta(S)) \geq 1$.

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

If S not narrow. Then $x^*(\delta(S)) \geq 2 \Rightarrow y(\delta(S)) \geq 1$.

If S is narrow, then $|F \cap \delta(S)| = 1 \Rightarrow |S \cap T|$ is not odd.

Gao (2013)

Theorem (Gao (2013))

For spanning tree F constructed by the algorithm, let J be a minimum-cost T-join. Then $c(F \cup J) \leq \frac{3}{2} OPT_{LP}$.

$$\text{Min} \quad \sum_{e \in E} x(e)$$

subject to:

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

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

Proof

If J is min-cost T-join

$$c(F \cup J) \leq n-1 + \frac{1}{2}x^*(E) \leq \frac{3}{2}x^*(E) = \frac{3}{2}\sum_{e\in E} x^*_e = \frac{3}{2}OPT_{LP}.$$

Since

$$x^*(E) = \frac{1}{2}\sum_{v\in V} x^*(\delta(v)) \geq \frac{1}{2}\left[2(n-2)+1+1\right]$$

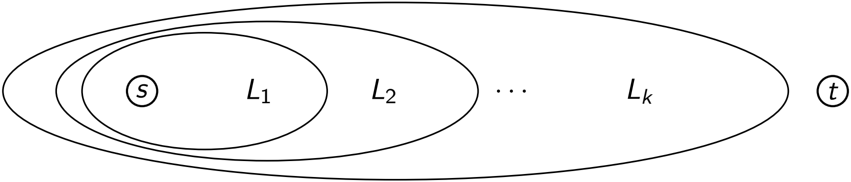
$$\geq n-1$$

Last Lemma

Let $E(x^*) = \{e \in E : x^*(e) > 0\}$ be the *support* of LP solution x^*, $H = (V, E(x^*))$ the support graph of x^*, $H(S)$ the graph induced by a set S of vertices.

Lemma (Gao (2013))

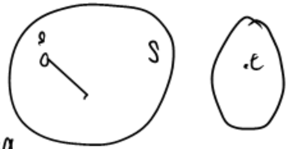
For $1 \leq p \leq q \leq k+1$, $H\left(\bigcup_{p \leq i \leq q} L_i\right)$ *is connected.*



Proof

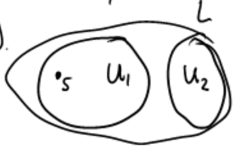
Case 1: $p=1$, $q=k+1$

H connected since $x^*(\delta(S))\ge 1$
for any $S\subset V$, $S\neq\emptyset$.



Case 2: $p=1$, $q<k+1$ Suppose not connected. $\exists$ partition U_1, U_2
s.t. $\delta_H(U_1)\cap\delta_H(U_2)=\emptyset$, $\delta_H(U_1)\cup\delta_H(U_2)=\delta_H(L)$.
$s\in U_1$, $t\notin U_2$. $x^*(\delta(U_1))\geq 1$, $x^*(\delta(U_2))\geq 2$.
$x^*(\delta(L))<2$.

$$2 > x^*(\delta(L)) = x^*(\delta(U_1)) + x^*(\delta(U_2)) \geq 1+2=3.$$



Case 3. $p>1$, $q=k+1$.

Case 4: $p>1$, $q<k+1$.

One Big Question

Is there a $\frac{3}{2}$-approx. alg. for s-t path TSP for general costs?

One Idea

Idea: Construct a spanning tree F just as in Gao's algorithm for the graph case. Then again $y = \frac{1}{2}x^*$ is feasible for the T-join LP, and the overall cost of F plus the T-join is at most $c(F) + \frac{1}{2}\sum_{e\in E} c(e)x^*(e)$.

One Idea

Idea: Construct a spanning tree F just as in Gao's algorithm for the graph case. Then again $y = \frac{1}{2}x^*$ is feasible for the T-join LP, and the overall cost of F plus the T-join is at most $c(F) + \frac{1}{2}\sum_{e\in E} c(e)x^*(e)$.

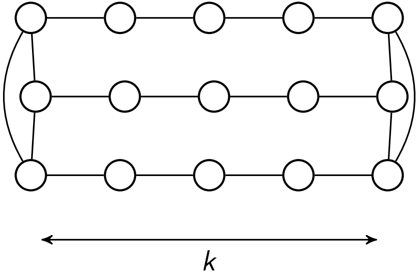
Problem: Not clear how to bound the cost of F. Gao (2014) has an example showing that F can have cost greater than OPT_{LP}.

Further directions

Best-of-Many Christofides from An et al. works with any possible decomposition of the LP solution into spanning trees. Recent improvements of Vygen (2015), Gottschalk and Vygen (2015), and Sebő and Van Zuylen (2016) all use decompositions that have particular properties.

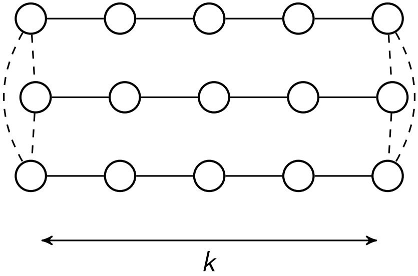
Integrality Gap for standard TSP

As with *s*-*t* path TSP, we can show a lower bound on the integrality gap using an instance of *graph TSP*.



Integrality Gap for standard TSP

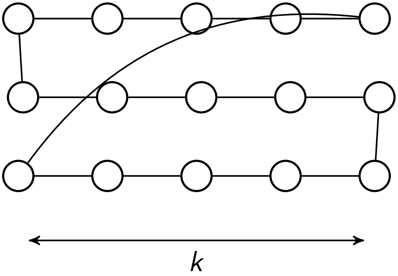
As with *s-t* path TSP, we can show a lower bound on the integrality gap using an instance of *graph TSP*.



$$OPT_{LP} \approx 3k$$

Integrality Gap for standard TSP

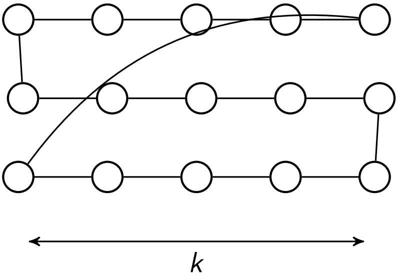
As with *s*-*t* path TSP, we can show a lower bound on the integrality gap using an instance of *graph TSP*.



$$OPT \approx 4k$$

Integrality Gap for standard TSP

As with *s*-*t* path TSP, we can show a lower bound on the integrality gap using an instance of *graph TSP*.



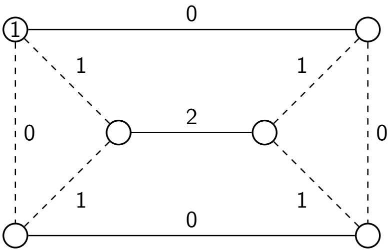
$$\frac{OPT}{OPT_{LP}} \rightarrow \frac{4}{3} \text{ as } k \rightarrow \infty$$

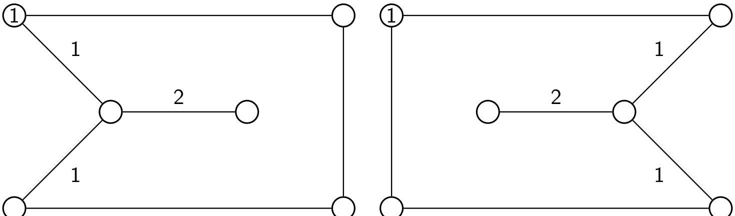
What about the standard TSP?

One can define Best-of-Many Christofides' for the standard TSP: solve the subtour LP, get LP solution x^*. Then since $\frac{n-1}{n}x^*$ is in the spanning tree polytope, find a decomposition of x^* into a convex combination of spanning trees. Run Christofides' algorithm on each one, return the best solution found.

Best-of-Many Christofides'

Unfortunately, the following example (due to Schalekamp and Van Zuylen) shows that an arbitrary decomposition into spanning trees will not improve on Christofides' $\frac{3}{2}$-approximation algorithm.





Experiments

Together with an undergraduate (Kyle Genova), we tried several different algorithms for decomposing the subtour LP into spanning trees.

We ran these algorithms on several types of instances:

- 59 Euclidean TSPLIB (Reinelt 1991) instances up to 2103 vertices (avg. 524);
- 5 non-Euclidean TSPLIB instances (gr120, si175, si535, pa561, si1032);
- 39 Euclidean VLSI instances (Rohe) up to 3694 vertices (avg. 1473);
- 9 graph TSP instances (Kunegis 2013) up to 1615 vertices (avg. 363).

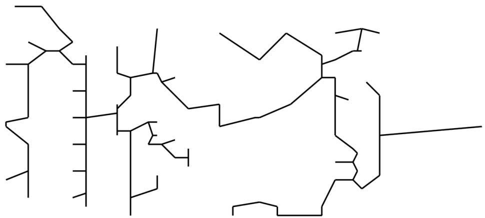
The results

	Std	ColGen		ColGen+SR	
		Best	Ave	Best	Ave
TSPLIB (E)	9.56%	4.03%	6.44%	3.45%	6.24%
VLSI	9.73%	7.00%	8.51%	6.40%	8.33%
TSPLIB (N)	5.40%	2.73%	4.41%	2.22%	4.08%
Graph	12.43%	0.57%	1.37%	0.39%	1.29%

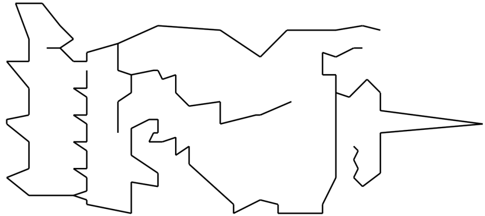
	MaxEnt		Split		Split+SR	
	Best	Ave	Best	Ave	Best	Ave
TSPLIB (E)	3.19%	6.12%	5.23%	6.27%	3.60%	6.02%
VLSI	5.47%	7.61%	6.60%	7.64%	5.48%	7.52%
TSPLIB (N)	2.12%	3.99%	2.92%	3.77%	1.99%	3.82%
Graph	0.31%	1.23%	0.88%	1.77%	0.33%	1.20%

Costs given as percentages in excess of the cost of the optimal tour.

The results



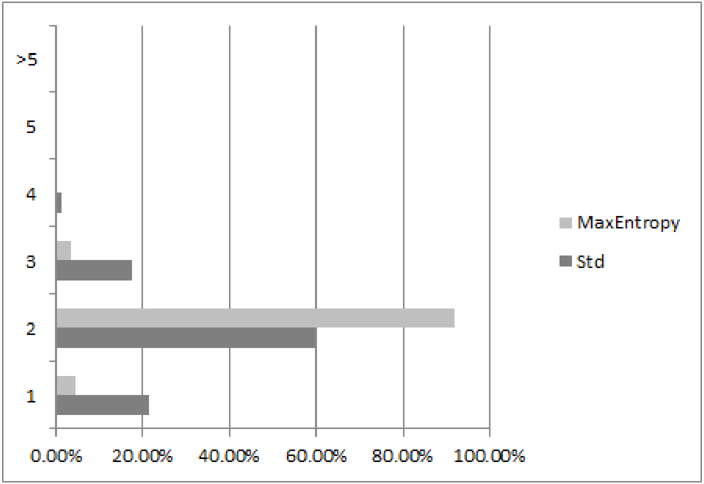
Standard Christofides MST (Rohe VLSI instance XQF131)



Splitting off + SwapRound

The results

BoMC yields more vertices in the tree of degree two.



The results

So while the tree costs more (as percentage of optimal tour)...

	Std	BOM
TSPLIB (E)	87.47%	98.57%
VLSI	89.85%	98.84%
TSPLIB (N)	92.97%	99.36%
Graph	79.10%	98.23%

The results

...the matching costs much less.

	Std	CG	CG+SR	MaxE	Split	Sp+SR
TSPLIB (E)	31.25%	11.43%	11.03%	10.75%	10.65%	10.41%
VLSI	29.98%	14.30%	14.11%	12.76%	12.78%	12.70%
TSPLIB (N)	24.15%	9.67%	9.36%	8.75%	8.77%	8.56%
Graph	39.31%	5.20%	4.84%	4.66%	4.34%	4.49%

Conclusion

Q: Are there empirical reasons to think BoMC might be provably better than Christofides' algorithm?

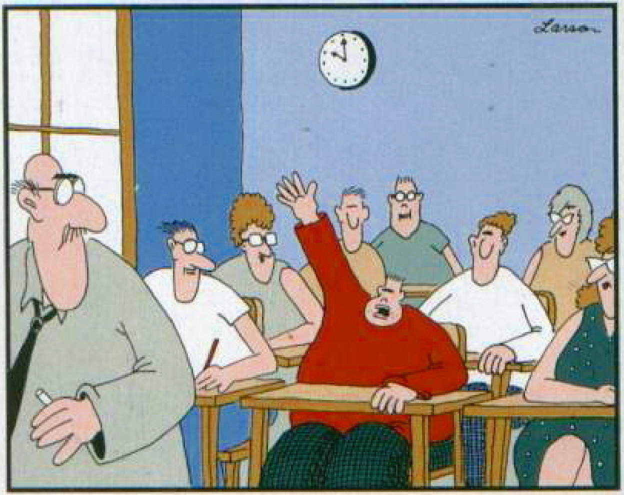
Conclusion

Q: Are there empirical reasons to think BoMC might be provably better than Christofides' algorithm?
A: Yes.

The Big Open Questions

- Beat $\frac{3}{2}$ for TSP
- Achieve $\frac{3}{2}$ for s-t path TSP
- Achieve $\frac{4}{3}$ for graph TSP

The End



"Mr. Osborne, may I be excused?
My brain is full."