Recent Progress in Approximation Algorithms for the Traveling Salesman Problem

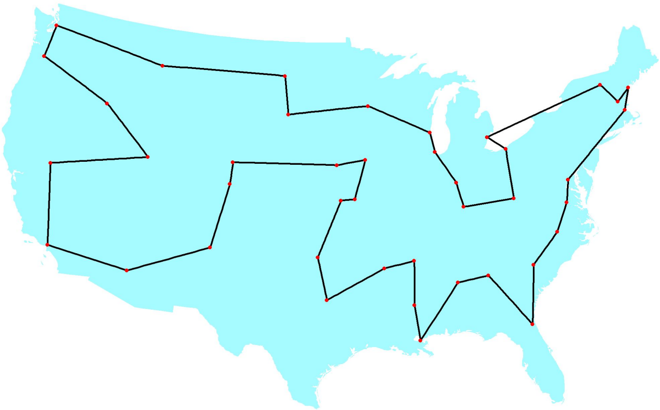
Lecture 2: Graph TSP

David P. Williamson
Cornell University

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

US 42 cities

From Bill Cook's site at the University of Waterloo (www.math.uwaterloo.ca/tsp)

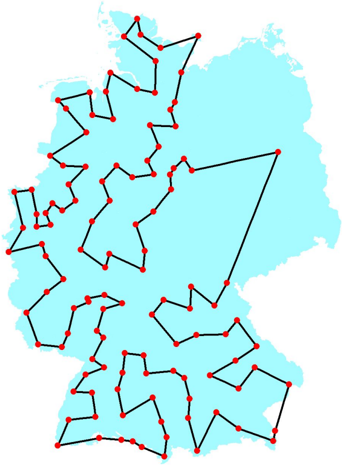


Solved by Dantzig, Fulkerson, and Johnson (1954) using linear programming.

US 33 cities

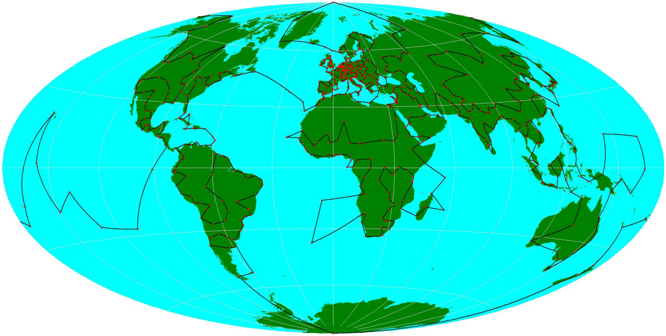
Proctor and Gamble contest in 1962.

Germany 120 cities



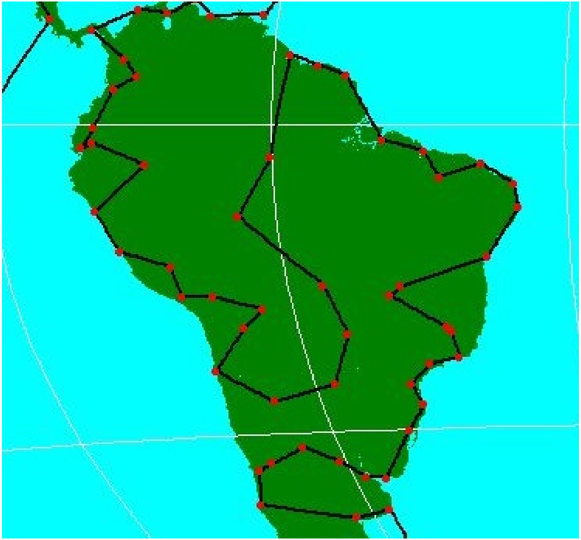
Solved by Grötschel (1977).

World tour 666 cities

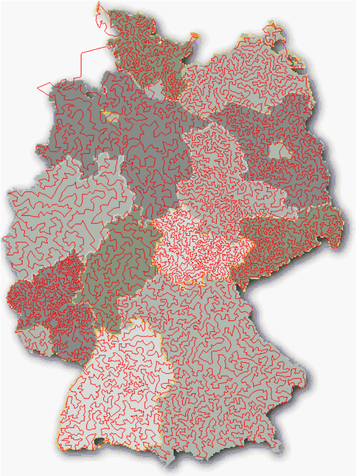


Solved by Grötschel and Holland (1987).

Brazil closeup

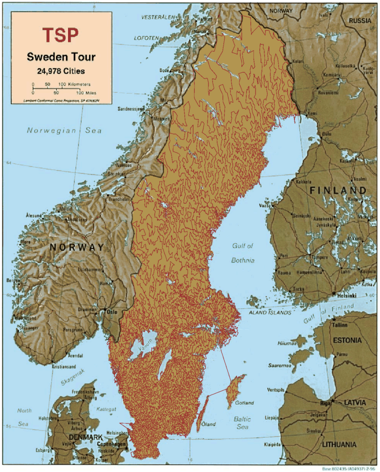


Germany 15112 cities



Solved by Applegate, Bixby, Chvátal, and Cook (2001)).

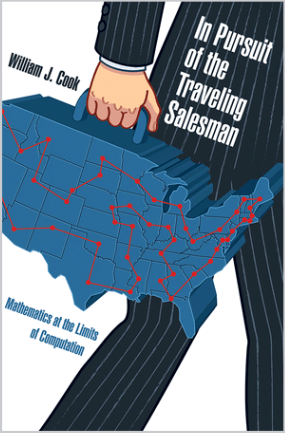
Sweden 24978 cities



Solved by Applegate, Bixby, Chvátal, Cook, and Helsgaun (2004).

Book

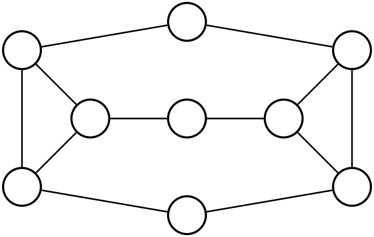
To learn more about the history of the TSP, read



I have three free copies to give away. Send me a paragraph about why you want/deserve a copy to dpw@cs.cornell.edu. Best three responses win.

Graph TSP

Recall *Graph TSP*: Input is connected graph $G = (V, E)$ and cost $c(i, j)$ is number of edges in shortest path from i to j in G.



Graph TSP

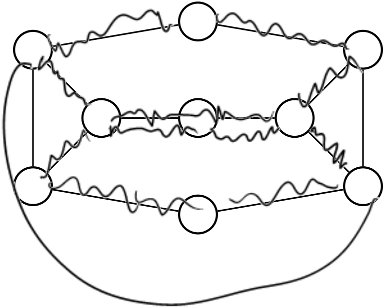
There has been recent progress on the case of graph TSP.

Oveis Gharan, Saberi, Singh	2010	$3/2 - \epsilon$	
Mömke, Svensson	2011	1.461	
Mömke, Svensson	2011	4/3	if graph *subcubic*
Mucha	2011	13/9	
Sebő and Vygen	2012	1.4	

Another perspective

Equivalent problem: Find Eulerian multigraph of (V, E) with the fewest number of edges. Recall *Eulerian* means every vertex has even degree, and the graph is connected.

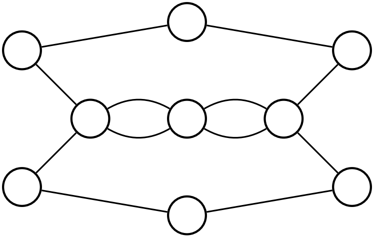
Given any tour, replace any non-edge (i, j) with all edges in the shortest i-j path.



Another perspective

Equivalent problem: Find Eulerian multigraph of (V, E) with the fewest number of edges. Recall *Eulerian* means every vertex has even degree, and the graph is connected.

Given any tour, replace any non-edge (i, j) with all edges in the shortest i-j path.



Today

Theorem (Mömke, Svensson (2011))

If G is cubic and 2-vertex-connected, then there is a $\frac{4}{3}$-approximation algorithm for the Graph TSP.

Today

Theorem (Mömke, Svensson (2011))

If G is cubic and 2-vertex-connected, then there is a $\frac{4}{3}$-approximation algorithm for the Graph TSP.

Cubic means all vertices have degree three. *2-vertex-connected* means that removing any one vertex (and its incident edges) from the graph does not disconnect the graph.

Idea

First idea: Since all vertices are odd-degree, add a matching to G. Then all vertices have even degree and G connected.

Lemma

Lemma (Naddef, Pulleyblank (1981))

Given a 2-edge-connected, cubic graph G with costs $c(e)$ on the edges, there is a perfect matching of cost at most $\frac{1}{3}\sum_{e\in E} c(e)$.

Suppose we set $c(e) = 1$ for all $e \in E$. Then cost of the graph plus matching is at most

$$|E| + \frac{1}{3}|E| = \frac{4}{3}|E| = \frac{4}{3}\left(\frac{3}{2}|V|\right)$$

$$= 2|V|$$

$$\leq 2 \cdot OPT$$

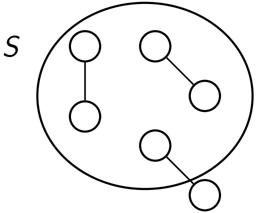
Proof of Naddef-Pulleyblank lemma

Lemma (Naddef, Pulleyblank (1981))

Given a 2-edge-connected, cubic graph G with costs $c(e)$ on the edges, there is a perfect matching of cost at most $\frac{1}{3}\sum_{e\in E} c(e)$.

Recall that the minimum-cost perfect matching can be found as the solution to the following LP:

$$\begin{array}{lll} \text{Minimize} & \displaystyle\sum_{e\in E} c(e)z(e) & \\ & z(\delta(i)) = 1 & \forall i \in V \\ & z(\delta(S)) \geq 1 & \forall S \subset V, |S| \text{ odd.} \end{array}$$



$\delta(S)$ = all edges w/ one endpt inside S, one outside

Proof of the Naddef-Pulleyblank lemma

Claim: $z(e) = \frac{1}{3}$ for all $e \in E$ is a feasible soln to matching LP.
Clearly cost of min-cost matching $\leq \frac{1}{3} \sum_{e \in E} c(e)$.

Lemma (Naddef, Pulleyblank (1981))

Given a 2-edge-connected, cubic graph G with costs $c(e)$ on the edges, there is a perfect matching of cost at most $\frac{1}{3} \sum_{e \in E} c(e)$.

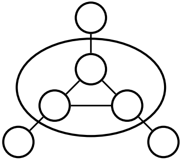
$$\text{Min} \sum_{e \in E} c(e) z(e)$$

$$z(\delta(i)) = 1 \quad \forall i \in V$$

$$z(\delta(S)) \geq 1 \quad \forall S \subset V, |S| \text{ odd}$$

$$z(e) \geq 0 \quad \forall e \in E.$$

S



$E(S)$: edges w/ both endpts in S

$\delta(S)$: edges w/ exactly one endpt in S

$z(e) = \frac{1}{3}$ for all $e \in E$.

Clearly $z(\delta(i)) = 1$ because G cubic.

Pick any S, $|S|$ odd. $\underbrace{3|S|}_{\text{odd}} = \underbrace{2|E(S)|}_{\text{even}} + \underbrace{|\delta(S)|}_{\text{odd}}$

If $|\delta(S)| \geq 3$ then clearly $z(\delta(S)) \geq 1$. But $|\delta(S)| \neq 1$ since G 2-edge connected. So $|\delta(S)| \geq 3$ since $|\delta(S)|$ odd.

New idea

Key idea: Use matchings to figure out which edges to **remove** and to add.

Removable pairing

Definition (Mömke, Svensson (2011))

Given G 2-vertex connected, $R \subseteq E$ *removable edges*, $P \subseteq R \times R$ is a *removable pairing* if:

- Any edge is in at most one pair of P;
- Edges in a pair have a common endpoint of degree at least 3;
- If we remove edges in R from G with at most one edge per pair in P removed, the resulting graph is still connected.

Theorem

Theorem (Mömke, Svensson (2011))

Given a removable pairing (R, P) and G 2-vertex-connected and cubic, there is an Eulerian multigraph with at most $\frac{4}{3}|E| - \frac{2}{3}|R|$ edges.

Theorem

Theorem (Mömke, Svensson (2011))

Given a removable pairing (R, P) and G 2-vertex-connected and cubic, there is an Eulerian multigraph with at most $\frac{4}{3}|E| - \frac{2}{3}|R|$ edges.

Idea: Take the graph, compute a perfect matching. If matching edge is in R, remove it, otherwise add it.

Theorem

Theorem (Mömke, Svensson (2011))

Given a removable pairing (R, P) and G 2-vertex-connected and cubic, there is an Eulerian multigraph with at most $\frac{4}{3}|E| - \frac{2}{3}|R|$ edges.

Proof of second claim: At most $\frac{4}{3}|E| - \frac{2}{3}|R|$ edges.

Apply Naddef-Pulleyblank with $c(e) = 1$ if $e \in E-R$, $c(e) = -1$ if $e \in R$.

For matching M, add e to G if $e \in E-R$, remove e from G if $e \in R$.

$\therefore$ edges is $|E| + |M-R| - |M \cap R| = |E| + \sum_{e \in M} c(e)$

$$\leq |E| + \frac{1}{3}|E-R| - \frac{1}{3}|R|$$

$$= \frac{4}{3}|E| - \frac{2}{3}|R|.$$

Theorem

Theorem (Mömke, Svensson (2011))

Given a removable pairing (R, P) and G 2-vertex-connected and cubic, there is an Eulerian multigraph with at most $\frac{4}{3}|E| - \frac{2}{3}|R|$ edges.

Proof of first claim: Result is connected and has even degree at all nodes.

Connected: since we remove ≤ 1 edge per pair (because it's a matching) and properties of a removable pairing.

Even degree: G is cubic and for each $i \in V$, we either
add edge incident on i (degree 4) or
remove " " " (degree 2).

Final Lemma

Lemma (Mömke, Svensson (2011))

In any cubic, 2-vertex-connected graph G, there is a removable pairing (R, P) with $|R| \geq |V|$.

Therefore, we can find an Eulerian graph with total number of edges at most

$$\frac{4}{3}|E| - \frac{2}{3}|R| \leq \frac{4}{3}|E| - \frac{2}{3}|V|$$

$$= \frac{4}{3}\left(\frac{3}{2}|V|\right) - \frac{2}{3}|V| = 2|V| - \frac{2}{3}|V|$$

$$= \frac{4}{3}|V| \leq \frac{4}{3}\mathrm{OPT}.$$

Proof of final lemma

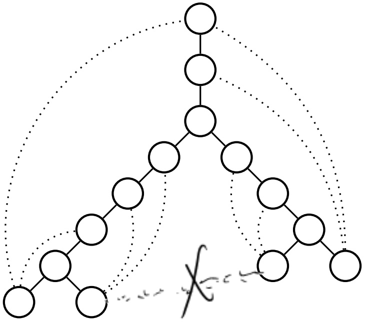
Lemma (Mömke, Svensson (2011))

In any cubic, 2-vertex-connected graph G, there is a removable pairing (R, P) with $|R| \geq |V|$.

Start by considering a depth-first search tree T of the graph G.

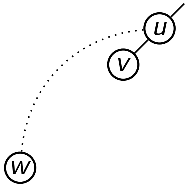
In DFS tree

---- back edges



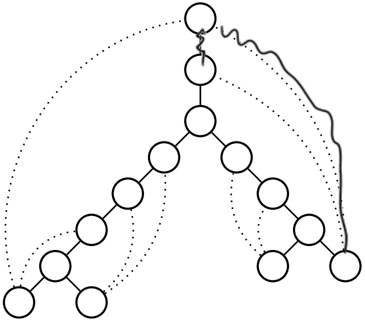
Proof of final lemma

For each back edge (u, w) with u the ancestor of w, tree edge (u, v), make (u, w) and (u, v) a removable pair. For root put only one back edge in the pair.



Proof of final lemma

For each back edge (u, w) with u the ancestor of w, tree edge (u, v), make (u, w) and (u, v) a removable pair. For root put only one back edge in the pair.



Definition again

Definition (Mömke, Svensson (2011))

Given G 2-vertex connected, $R \subseteq E$ *removable edges*, $P \subseteq R \times R$ is a *removable pairing* if:

- Any edge is in at most one pair of P;
- Edges in a pair have a common endpoint of degree at least 3;
- If we remove edges in R from G with at most one edge per pair in P removed, the resulting graph is still connected.

Proof of final lemma

Number of back edges is $|E| - |T| = |E| - (|V| - 1) = \frac{3}{2}|V| - |V| + 1$

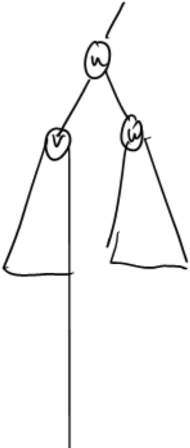
$= \frac{1}{2}|V| + 1$

$|R| = 2(\#\text{back edge} - 1) = 2\left(\frac{1}{2}|V|\right) = |V|$

Proof of final lemma

Need to show that G stays connected if we remove at most one edge per pair. Prove by induction bottom up on subtrees; let T_u be subtree rooted at vertex u.

Case 1: u has two children v and w, one parent in T.



By induction, T_v, T_w stay connected

$\Rightarrow$ T_u stay connected

Proof of final lemma

Need to show that G stays connected if we remove at most one edge per pair. Prove by induction bottom up on subtrees; let T_u be subtree rooted at vertex u.

Case 2: u has one child v, one parent in T.

By induction T_v stays connected.
and therefore T_u stays connected if
either (u,v) removed or
(u,w) removed (but not both!)

Extensions

Mömke and Svensson (2011) show that the same ideas can be extended to:

- a $\frac{4}{3}$-approximation algorithm for subcubic graphs (all degrees at most 3)
- a 1.461-approximation algorithm for all graphs (improved by Mucha (2012) to $\frac{13}{9}$)

Extensions

Mömke and Svensson (2011) show that the same ideas can be extended to:

- a $\frac{4}{3}$-approximation algorithm for subcubic graphs (all degrees at most 3)
- a 1.461-approximation algorithm for all graphs (improved by Mucha (2012) to $\frac{13}{9}$)

Sebő and Vygen (2012) add some new ideas and get a 1.4-approximation algorithm.