Worst Case and Probabilistic Analysis of the 2-Opt Algorithm for the TSP

Matthias Englert Heiko Röglin Berthold Vöcking

Department of Computer Science
RWTH Aachen

Oberwolfach 2007

Traveling Salesperson Problem

Traveling Salesperson Problem (TSP)

- Input: weighted (complete) graph $G = (V, E, d)$ with $d : E \to \mathbb{R}$.

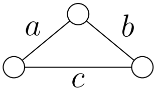
Traveling Salesperson Problem (TSP)

- Input: weighted (complete) graph $G = (V, E, d)$ with $d : E \to \mathbb{R}$.
- Goal: Find Hamiltonian cycle of minimum length.

Theoretical Results

- General TSP
  - Strongly NP-hard.
  - Not approximable within any polynomial factor.

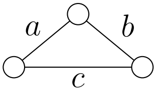
Theoretical Results



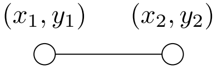
$a + b \geq c$

- General TSP
  - Strongly NP-hard.
  - Not approximable within any polynomial factor.
- Metric TSP
  - Strongly NP-hard.
  - 3/2-approximation [Christofides, (1976)]
  - APX-hard: lower bound of 220/219 [Papadimitriou, Vempala (2000)]

Theoretical Results



$a + b \geq c$



$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$

- General TSP
  - Strongly NP-hard.
  - Not approximable within any polynomial factor.
- Metric TSP
  - Strongly NP-hard.
  - 3/2-approximation [Christofides, (1976)]
  - APX-hard: lower bound of 220/219 [Papadimitriou, Vempala (2000)]
- Euclidean TSP
  - Cities $\subset \mathbb{R}^d$
  - Strongly NP-hard ($\Rightarrow$ no FPTAS) [Papadimitriou (1977)]
  - PTAS exists [Arora (1996), Mitchell (1996)].

Experimental Results

- Numerous experimental studies.
  - TSPLIB contains "real-world" and random (Euclidean) instances.
  - DIMACS Implementation Challenge [Johnson and McGeoch (2002)].

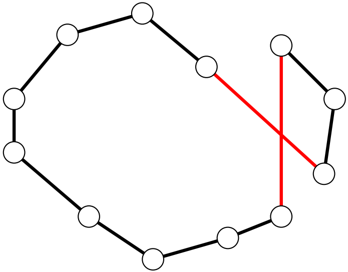
Experimental Results

- Numerous experimental studies.
  - TSPLIB contains “real-world” and random (Euclidean) instances.
  - DIMACS Implementation Challenge [Johnson and McGeoch (2002)].

Some conclusions:

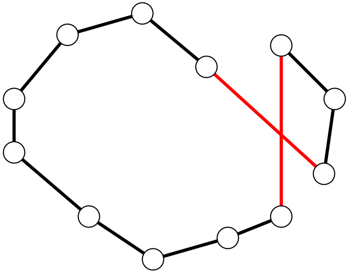
- Worst-case results are often too pessimistic.
- The PTAS is too slow on large scale instances.
- The most successful algorithms (w. r. t. quality and running time) in practice rely on local search.

2-Opt Heuristic



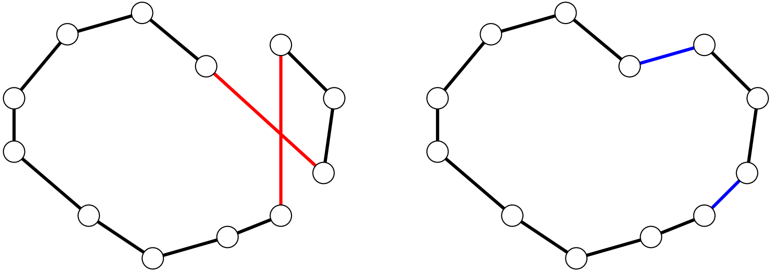
1. Start with an arbitrary tour.

2-Opt Heuristic



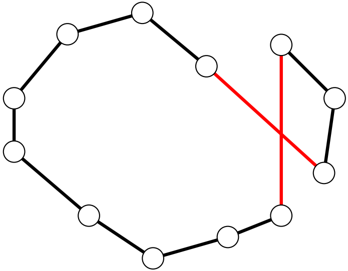
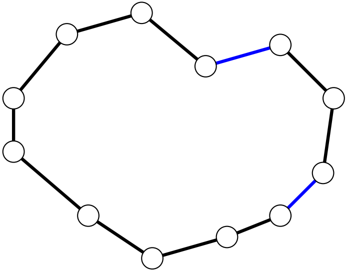
1. Start with an arbitrary tour.
2. Remove two edges from the tour.

2-Opt Heuristic



1. Start with an arbitrary tour.
2. Remove two edges from the tour.
3. Complete the tour by two other edges.

2-Opt Heuristic

1. Start with an arbitrary tour.
2. Remove two edges from the tour.
3. Complete the tour by two other edges.
4. Repeat steps 2 and 3 until no local improvement is possible anymore.

Why 2-Opt?

Experiments on Random Euclidean Instances

[Johnson and McGeoch (2002)]

Approximation Ratio

- Christofides (for $n \leq 10^5$): ≈ 1.1
- 2-Opt (for $n \leq 10^6$): ≈ 1.05

Why 2-Opt?

Experiments on Random Euclidean Instances

[Johnson and McGeoch (2002)]

Approximation Ratio

- Christofides (for $n \leq 10^5$): ≈ 1.1
- 2-Opt (for $n \leq 10^6$): ≈ 1.05

Number of Local Improvements of 2-Opt

- Greedy Starts: Probably $O(n)$
- Random Starts: Probably $O(n \log n)$

Running time of 2-Opt: Known and New Results

	General TSP	Euclidean metric	Manhattan metric
average			
smoothed			
worst-case	$2^{\Omega(n)}$		

Average-case results: [Chandra, Karloff, Tovey (1999)].
Worst-case results: [Lueker (1975)].

Running time of 2-Opt: Known and New Results

	General TSP	Euclidean metric	Manhattan metric
average			
smoothed			
worst-case	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$

Average-case results: [Chandra, Karloff, Tovey (1999)].
Worst-case results: [Lueker (1975)].
Our results.

Running time of 2-Opt: Known and New Results

	General TSP	Euclidean metric	Manhattan metric
average		$\tilde{O}(n^{10})$	$\tilde{O}(n^6)$
smoothed			
worst-case	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$

Average-case results: [Chandra, Karloff, Tovey (1999)].
Worst-case results: [Lueker (1975)].
Our results.

Running time of 2-Opt: Known and New Results

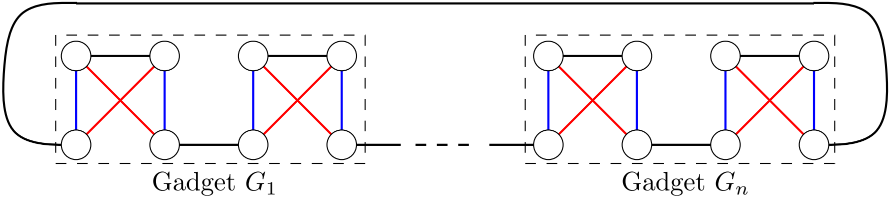
	General TSP	Euclidean metric	Manhattan metric
average	$n^{3+o(1)}$	$\tilde{O}(n^{10})$ $\tilde{O}(n^{4.33})$ [$\tilde{O}(n^{3.83})$]	$\tilde{O}(n^{6})$ $\tilde{O}(n^{4})$ [$\tilde{O}(n^{3.5})$]
smoothed			
worst-case	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$

Average-case results: [Chandra, Karloff, Tovey (1999)].
Worst-case results: [Lueker (1975)].
Our results.

Lower Bound

Theorem

For every $n \in \mathbb{N}$, there is a graph in the Euclidean plane with $8n$ vertices on which 2-Opt can make $2^{n+3} - 14$ steps.



Possible States of a Gadget:

(Long,Long), (Long,Short), (Short,Long), (Short,Short)

Lower Bound

(Long,Long) $= 0$

↓

(Long,Short) $= 1$

↓

(Short,Short) $= 2$

Lower Bound

0 2

(Long,Long) = 0

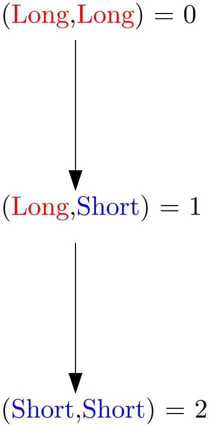
↓

(Long,Short) = 1

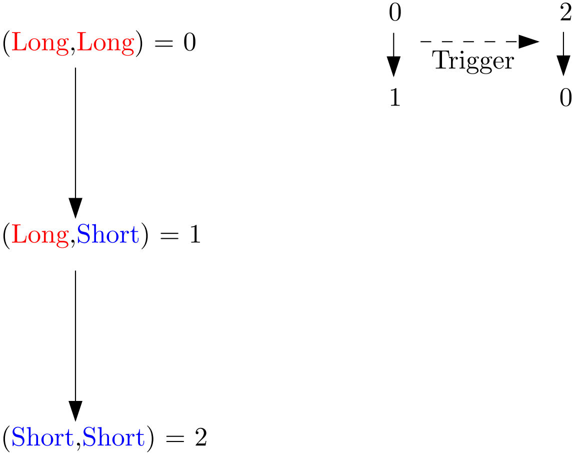
↓

(Short,Short) = 2

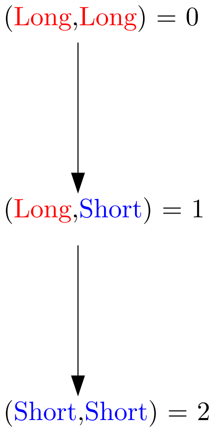
Lower Bound

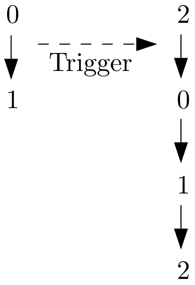


Lower Bound



Lower Bound





Lower Bound

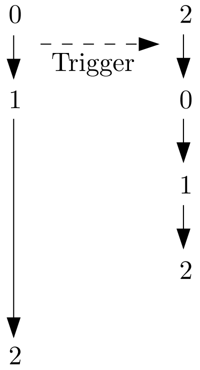
(Long,Long) $= 0$

↓

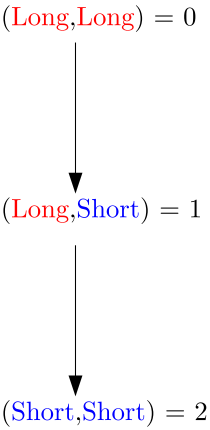
(Long,Short) $= 1$

↓

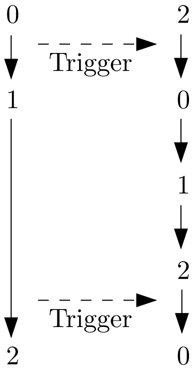
(Short,Short) $= 2$



Lower Bound

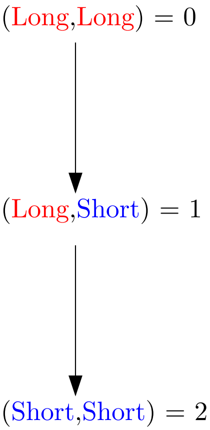


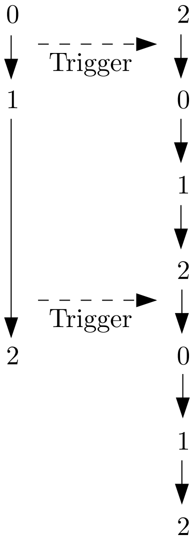
(Long,Long) = 0
(Long,Short) = 1
(Short,Short) = 2



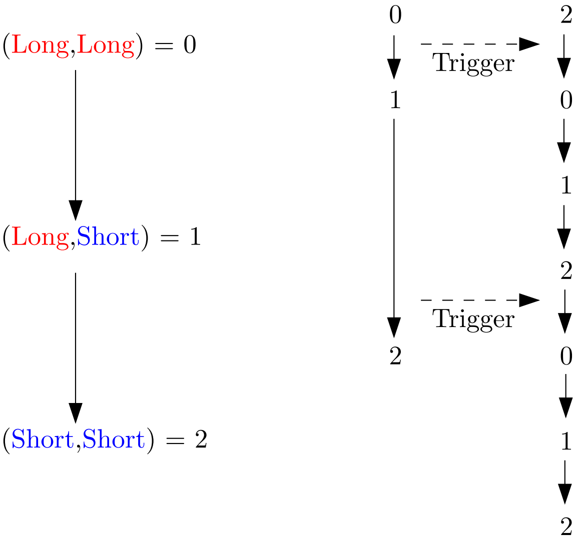
0
2
Trigger
1
0
1
2
Trigger
2
0

Lower Bound


(Long,Long) = 0
(Long,Short) = 1
(Short,Short) = 2

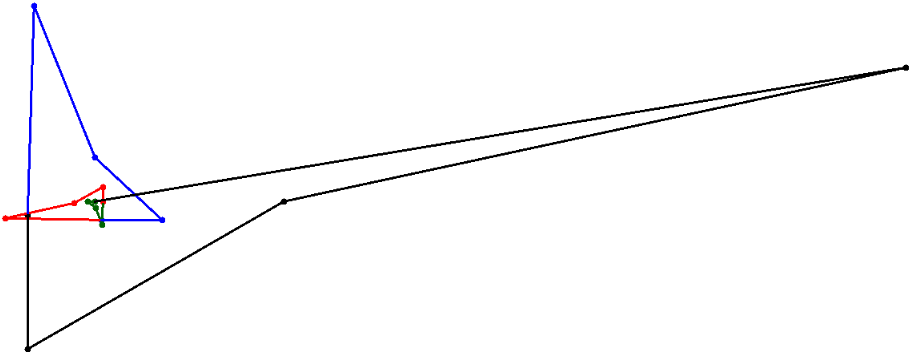

0
1
2
Trigger
Trigger
2
0
1
2
0
1
2

Lower Bound



Gadget G_i is reset 2^{i-1} times to (Long,Long) $= 0$.

Euclidean Embedding of the Gadgets



Upper Bound

Upper Bound

Theorem

Assume that n points are placed independently, uniformly at random in the unit square $[0,1]^2$. The expected number of 2-Opt steps is bounded by $O(n^{4+1/3} \cdot \log n)$ (for every initial tour and every pivot rule).

Simple Polynomial Bound

Theorem

The expected number of 2-Opt steps is bounded by $O(n^7 \log^2 n)$.

Simple Polynomial Bound

Theorem

The expected number of 2-Opt steps is bounded by $O(n^7 \log^2 n)$.

Proof.

- Consider a 2-Opt step $(e_1, e_2) \to (e_3, e_4)$.
- $\Delta(e_1, e_2, e_3, e_4) = l(e_1) + l(e_2) - l(e_3) - l(e_4)$.

□

Simple Polynomial Bound

Theorem

The expected number of 2-Opt steps is bounded by $O(n^7 \log^2 n)$.

Proof.

- Consider a 2-Opt step $(e_1, e_2) \to (e_3, e_4)$.
- $\Delta(e_1, e_2, e_3, e_4) = l(e_1) + l(e_2) - l(e_3) - l(e_4)$.
- $$\Delta = \min_{e_1, e_2, e_3, e_4} |\Delta(e_1, e_2, e_3, e_4)|.$$

□

Simple Polynomial Bound

Theorem

The expected number of 2-Opt steps is bounded by $O(n^7 \log^2 n)$.

Proof.

- Consider a 2-Opt step $(e_1, e_2) \to (e_3, e_4)$.
- $\Delta(e_1, e_2, e_3, e_4) = l(e_1) + l(e_2) - l(e_3) - l(e_4)$.
- $$\Delta = \min_{e_1,e_2,e_3,e_4} |\Delta(e_1, e_2, e_3, e_4)|.$$
- $$\# \text{ 2-Opt Steps} \le \frac{\sqrt{2}n}{\Delta}.$$

□

Simple Polynomial Bound

Theorem

The expected number of 2-Opt steps is bounded by $O(n^7 \log^2 n)$.

Proof.

- Consider a 2-Opt step $(e_1, e_2) \to (e_3, e_4)$.
- $\Delta(e_1, e_2, e_3, e_4) = l(e_1) + l(e_2) - l(e_3) - l(e_4)$.
- $$\Delta = \min_{e_1,e_2,e_3,e_4} |\Delta(e_1, e_2, e_3, e_4)|.$$
- $$\#\text{ 2-Opt Steps} \leq \frac{\sqrt{2}n}{\Delta}.$$
- Bound Δ by a union bound: There are $O(n^4)$ different 2-Opt steps, analyze $\Delta(e_1, e_2, e_3, e_4)$ for one of them. $\Rightarrow \Delta \approx 1/(n^4 \log n)$.

□

Idea for Improvement

- The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$.

Idea for Improvement

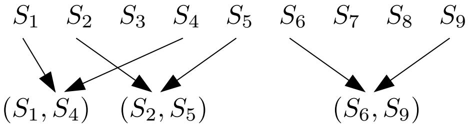
- The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$.
- Consider two consecutive steps: They yield $\Delta + \Delta_2 > 2\Delta$.

Idea for Improvement

- The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$.
- Consider two consecutive steps: They yield $\Delta + \Delta_2 > 2\Delta$.
- Consider linked pair: $(e_1, e_2) \to (e_3, e_4)$ and $(e_3, e_5) \to (e_6, e_7)$.

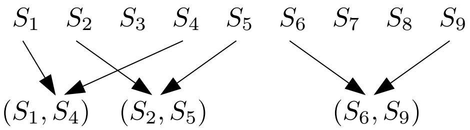
Idea for Improvement

- The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$.
- Consider two consecutive steps: They yield $\Delta + \Delta_2 > 2\Delta$.
- Consider linked pair: $(e_1, e_2) \to (e_3, e_4)$ and $(e_3, e_5) \to (e_6, e_7)$.
- Sequence of t consecutive steps, contains $\Omega(t)$ linked pairs:



Idea for Improvement

- The bound is too pessimistic: Not every step yields the smallest possible improvement $\Delta \approx 1/(n^4 \log n)$.
- Consider two consecutive steps: They yield $\Delta + \Delta_2 > 2\Delta$.
- Consider linked pair: $(e_1, e_2) \to (e_3, e_4)$ and $(e_3, e_5) \to (e_6, e_7)$.
- Sequence of t consecutive steps, contains $\Omega(t)$ linked pairs:



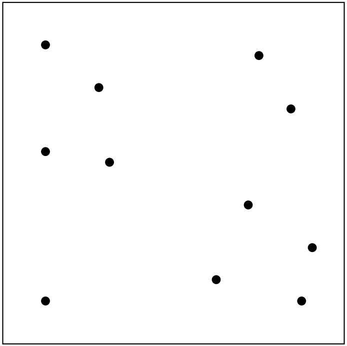
- $\Delta_{\text{Linked}} \approx 1/(n^{3+1/3} \log^{2/3} n)$.

Extensions and Open Problems

Smoothed Analysis

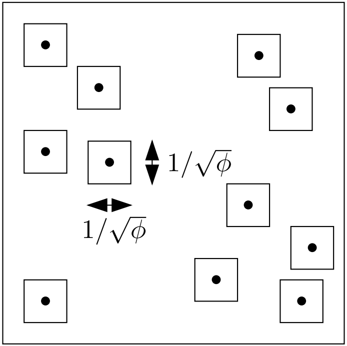
Smoothed Analysis

Each point $i \in \{1, \ldots, n\}$ is chosen independently according to a probability density $f_i : [0,1]^2 \rightarrow [0, \phi]$.



Smoothed Analysis

Each point $i \in \{1, \ldots, n\}$ is chosen independently according to a probability density $f_i : [0,1]^2 \to [0, \phi]$.



Smoothed Analysis

	General TSP	Euclidean metric	Manhattan metric
average	$n^{3+o(1)}$	$\tilde{O}(n^{4.33})$	$\tilde{O}(n^{4})$
smoothed	$m \cdot n^{1+o(1)} \cdot \phi$	$\tilde{O}(n^{4.33} \cdot \phi^{2.67})$	$\tilde{O}(n^{4} \cdot \phi)$
worst-case	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$

Approximation Ratio

- Worst Case: $O(\log n)$
- Worst Case: $\Omega(\log n / \log\log n)$

Approximation Ratio

- Worst Case: $O(\log n)$
- Worst Case: $\Omega(\log n / \log \log n)$
- Average Case: $O(1)$

Approximation Ratio

- Worst Case: $O(\log n)$
- Worst Case: $\Omega(\log n / \log \log n)$
- Average Case: $O(1)$
- Smoothed: $O(\sqrt{\phi})$

Open Problems

Worst-Case Analysis

- Analyze the diameter of the 2-Opt state graph.
- Analyze particular pivot rules like “largest improvement”.

Probabilistic Analysis

- Show exact bounds on the running time of 2-Opt and k-Opt.
- Show small constant approximation ratio for 2-Opt on random Euclidean instances.

Thanks!

Questions?