Computability and Complexity COSC 4200

Approximating TSP

Approximation Algorithms

Let X be a minimization problem.

An algorithm \mathcal{A} is an $\alpha(n)$ -approximation algorithm if for all n, for all instances of size n, \mathcal{A} returns a solution with value at most $\alpha(n)$ times the value of the optimal solution.

For example:

- a log *n*-approximation algorithm
- a 2-approximation algorithm

Approximation Algorithms

Let X be a minimization problem.

An algorithm \mathcal{A} is an $\alpha(n)$ -approximation algorithm if for all n, for all instances of size n, \mathcal{A} returns a solution with value at most $\alpha(n)$ times the value of the optimal solution.

For example:

- a log *n*-approximation algorithm
- a 2-approximation algorithm

Let X be a maximization problem.

An algorithm \mathcal{A} is an $\alpha(n)$ -approximation algorithm if for all n, for all instances of size n, \mathcal{A} returns a solution with value at least $\alpha(n)$ times the value of the optimal solution.

For example:

- \bullet a $\frac{7}{8}$ -approximation algorithm
- a $\frac{\log n}{n}$ -approximation algorithm

Approximation of TSP

Theorem

If $P \neq NP$, then for any polynomial-time computable function $\alpha(n)$, there is no $\alpha(n)$ -approximation algorithm for TSP.

Proof. We modify our reduction of HAM-CYCLE to TSP. Let G = (V, E) be a graph.

Approximation of TSP

Theorem

If $P \neq NP$, then for any polynomial-time computable function $\alpha(n)$, there is no $\alpha(n)$ -approximation algorithm for TSP.

Proof. We modify our reduction of HAM-CYCLE to TSP. Let G = (V, E) be a graph.

Define n = |V|, k = n, and

$$c(i,j) = \begin{cases} 1 & \text{if } (i,j) \in E \\ \alpha(n) \cdot n & \text{if } (i,j) \notin E \end{cases}$$

for all $i, j \in \{1, ..., n\}$.

Notice that:

- if $G \in HAM\text{-}CYCLE$, then there is a TSP tour with cost n.
- if $G \notin HAM\text{-CYCLE}$, then any TSP tour has cost $> \alpha(n) \cdot n$.

Suppose A is a polynomial-time $\alpha(n)$ -approximation algorithm for TSP. Consider the following Algorithm B.

```
input: graph G, instance of HAM-CYCLE
construct the TSP instance from G as described
run \mathcal{A} on the TSP instance
let c = \cos t of tour output by A
if c \leq \alpha(n) \cdot n
    output 'yes'
else
    output 'no'
```

We claim that Algorithm \mathcal{B} decides HAM-CYCLE.

Suppose $G \in HAM\text{-}CYCLE$.

- Then there is a TSP tour with cost *n*.
- The approximation algorithm \mathcal{A} will find a tour with cost at most $\alpha(n) \cdot n$.
- ullet Therefore algorithm ${\cal B}$ will output 'yes.'

Suppose $G \in HAM\text{-}CYCLE$.

- Then there is a TSP tour with cost n.
- The approximation algorithm \mathcal{A} will find a tour with cost at most $\alpha(n) \cdot n$.
- Therefore algorithm B will output 'yes.'

Now suppose $G \not\in HAM\text{-}CYCLE$.

- Then any TSP tour has cost $> \alpha(n) \cdot n$.
- The approximation algorithm \mathcal{A} will return a tour with cost $> \alpha(\mathbf{n}) \cdot \mathbf{n}$.
- Therefore algorithm B will output 'no.'

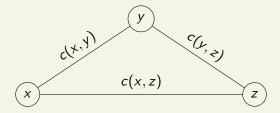
Therefore \mathcal{B} decides HAM-CYCLE and HAM-CYCLE \in P. Since HAM-CYCLE is NP-complete, P = NP. \square

Metric TSP

We showed there is no hope of approximating general TSP (unless P=NP). However, the cost function in the proof is wild: some costs are low and some are very high.

In the Metric TSP problem, the costs must satisfy the triangle inequality: for all x, y, and z,

$$c(x,z) \le c(x,y) + c(y,z).$$



Consider the following algorithm to compute TSP tours:

On input a TSP instance as a weighted graph G:

1 Find a minimum spanning tree, T, of G.

Consider the following algorithm to compute TSP tours:

On input a TSP instance as a weighted graph G:

- Find a minimum spanning tree, T, of G.
- ② Double every edge of T to obtain an Eulerian graph E.

Consider the following algorithm to compute TSP tours:

On input a TSP instance as a weighted graph G:

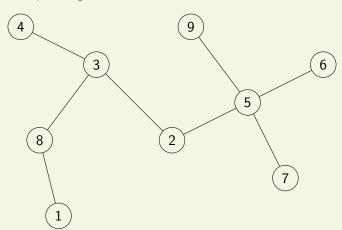
- **1** Find a minimum spanning tree, T, of G.
- ② Double every edge of T to obtain an Eulerian graph E.
- **3** Find an Eulerian tour, \mathcal{T} , of E.

Consider the following algorithm to compute TSP tours:

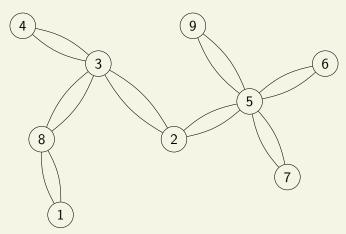
On input a TSP instance as a weighted graph G:

- Find a minimum spanning tree, T, of G.
- ② Double every edge of T to obtain an Eulerian graph E.
- 3 Find an Eulerian tour, \mathcal{T} , of E.
- **①** Output the tour C that visits vertices of G in the order of their first appearing in T.

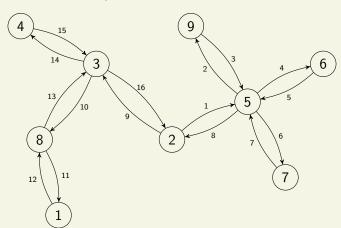
Minimum spanning tree T

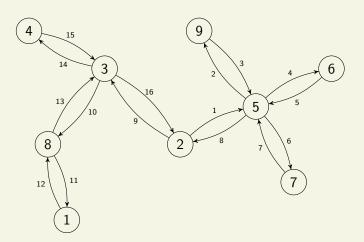


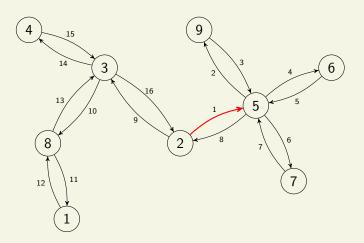
Double every edge to obtain E

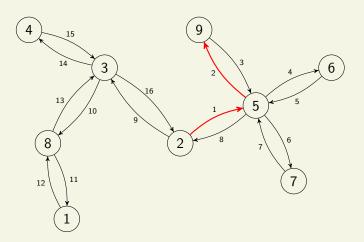


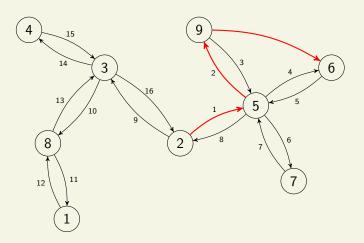
Find an Eulerian tour ${\mathcal T}$

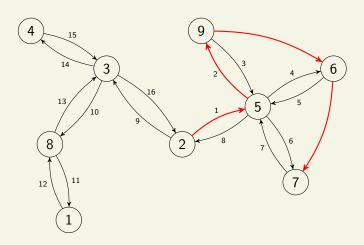


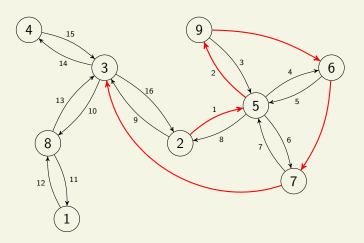


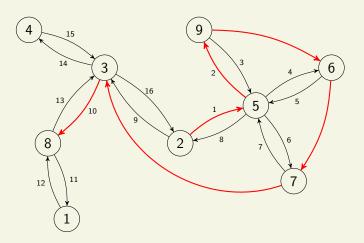


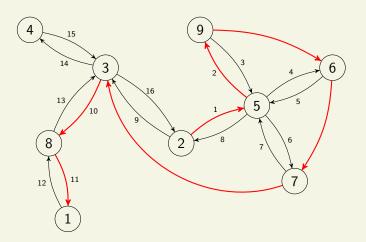


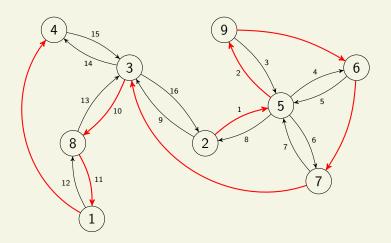


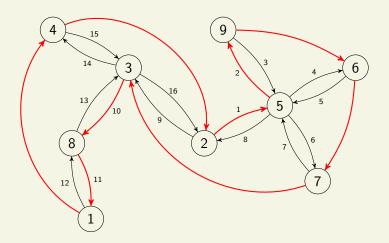












Proof. Let OPT be the cost of an optimal TSP tour. Observe that:

• $cost(T) \le OPT$. If we take any edge away from an optimal TSP tour, we have a spanning tree U. Thus $cost(T) \le cost(U) \le OPT$.

Proof. Let OPT be the cost of an optimal TSP tour. Observe that:

- $cost(T) \le OPT$. If we take any edge away from an optimal TSP tour, we have a spanning tree U. Thus $cost(T) \le cost(U) \le OPT$.
- $cost(T) = cost(E) = 2 \cdot cost(T)$.

Proof. Let OPT be the cost of an optimal TSP tour. Observe that:

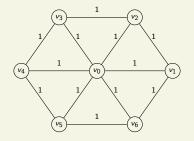
- $cost(T) \le OPT$. If we take any edge away from an optimal TSP tour, we have a spanning tree U. Thus $cost(T) \le cost(U) \le OPT$.
- $cost(T) = cost(E) = 2 \cdot cost(T)$.
- By the (extended) triangle inequality, $cost(C) \leq cost(T)$.

Proof. Let OPT be the cost of an optimal TSP tour. Observe that:

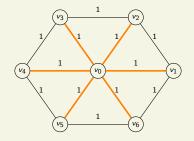
- $cost(T) \le OPT$. If we take any edge away from an optimal TSP tour, we have a spanning tree U. Thus $cost(T) \le cost(U) \le OPT$.
- $cost(T) = cost(E) = 2 \cdot cost(T)$.
- By the (extended) triangle inequality, $cost(C) \leq cost(T)$.

Putting everything together, we have

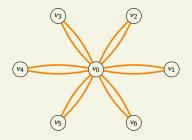
$$cost(C) \le cost(T) = 2 \cdot cost(T) \le 2 \cdot OPT.$$



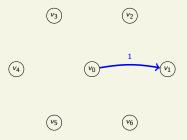
The distance for all other pairs of vertices is 2.



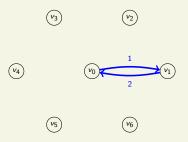
The MST consists of the orange edges and has cost 6.



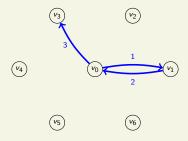
Double the edges of the MST to obtain an Eulerian graph.



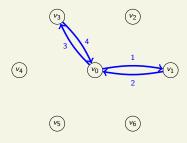
An Eulerian tour. The algorithm could find any of many Eulerian tours, but this one will give worst-case performance.

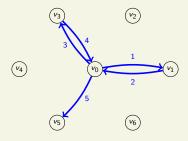


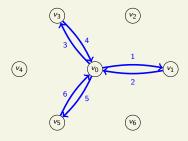
An Eulerian tour. The algorithm could find any of many Eulerian tours, but this one will give worst-case performance.

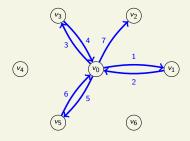


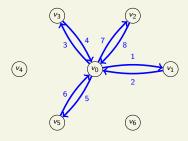
An Eulerian tour. The algorithm could find any of many Eulerian tours, but this one will give worst-case performance.

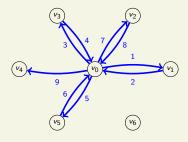


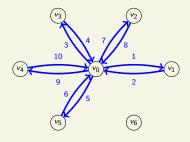


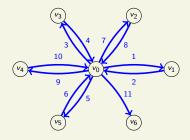


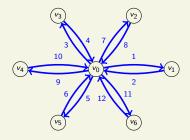


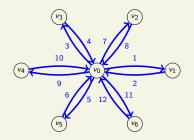


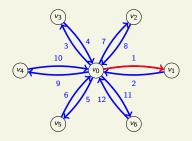


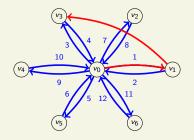


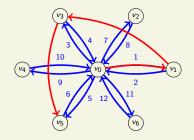


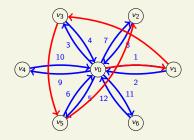


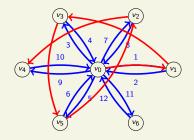


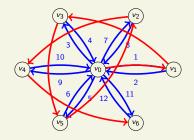


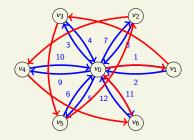


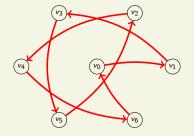


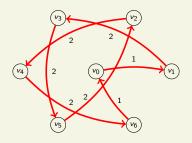




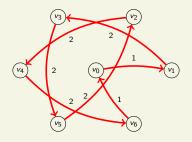






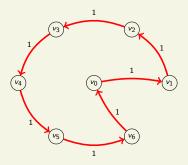


Tour cost: 2(6-1)+2=12.

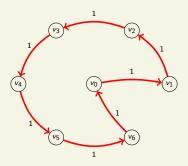


Tour cost: 2(6-1)+2=12.

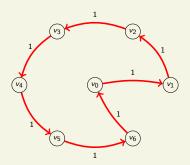
If we change 6 to n and obtain a analogous graph on n+1 vertices, the cost is 2(n-1)+2=2n.



Optimal tour cost: 7



Optimal tour cost: 7 If we change 6 to n and obtain a analogous graph on n+1 vertices, the optimal tour cost is n+1.



Optimal tour cost: 7

If we change 6 to n and obtain a analogous graph on n+1 vertices, the optimal tour cost is n+1.

The performance ratio is $\frac{12}{7} \le 2$ on the above instance. In general, for the graph on n+1 vertices, the performance ratio is:

$$\frac{2n}{n+1}$$
.

This approaches 2 as $n \to \infty$.

An improvement - Christofides' Algorithm Consider the following algorithm to compute TSP tours:

On input a TSP instance as a weighted graph G:

• Find a minimum spanning tree, T, of G.

An improvement - Christofides' Algorithm Consider the following algorithm to compute TSP tours:

On input a TSP instance as a weighted graph G:

- Find a minimum spanning tree, T, of G.
- Compute a minimum cost perfect matching, M, on the set of odd-degree vertices of T. A minimum cost perfect matching is a pairing (matching) of vertices at minimum total edge cost. Add M to T and obtain an Eulerian graph E.

An improvement - Christofides' Algorithm Consider the following algorithm to compute TSP tours:

On input a TSP instance as a weighted graph G:

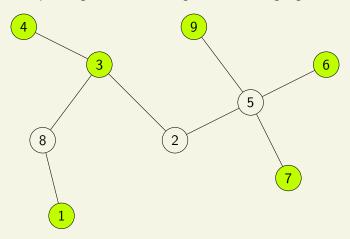
- **1** Find a minimum spanning tree, T, of G.
- Compute a minimum cost perfect matching, M, on the set of odd-degree vertices of T. A minimum cost perfect matching is a pairing (matching) of vertices at minimum total edge cost. Add M to T and obtain an Eulerian graph E.
- **3** Find an Eulerian tour, \mathcal{T} , of E.

An improvement - Christofides' Algorithm Consider the following algorithm to compute TSP tours:

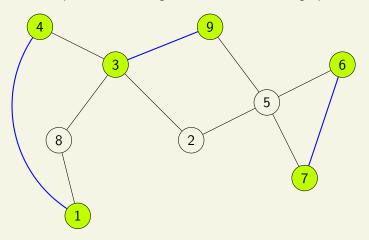
On input a TSP instance as a weighted graph G:

- Find a minimum spanning tree, T, of G.
- Compute a minimum cost perfect matching, M, on the set of odd-degree vertices of T. A minimum cost perfect matching is a pairing (matching) of vertices at minimum total edge cost. Add M to T and obtain an Eulerian graph E.
- **3** Find an Eulerian tour, \mathcal{T} , of E.
- **①** Output the tour C that visits vertices of G in the order of their first appearing in T.

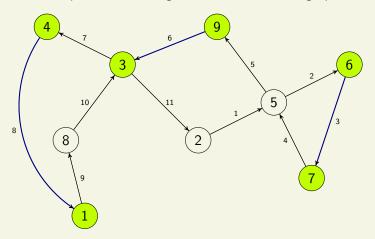
Minimum spanning tree T, odd degree vertices highlighted



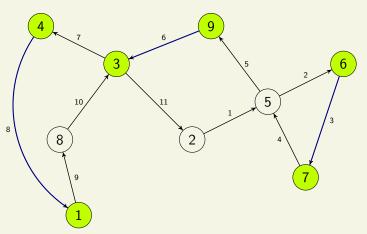
Minimum cost perfect matching added, now Eulerian graph



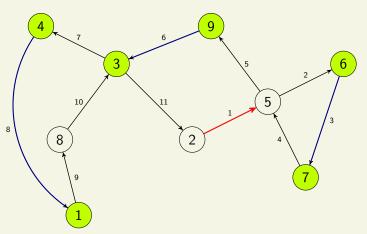
Minimum cost perfect matching added, now Eulerian graph



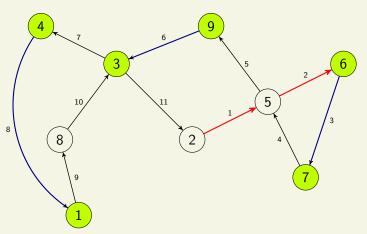
Take shortcuts to obtain final tour ${\mathcal C}$



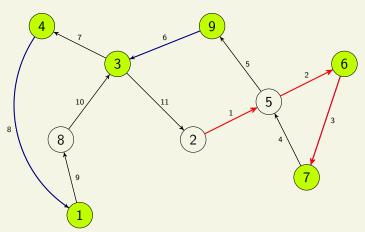
Take shortcuts to obtain final tour ${\mathcal C}$



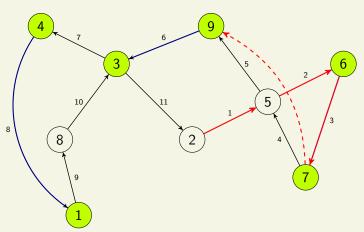
Take shortcuts to obtain final tour ${\mathcal C}$



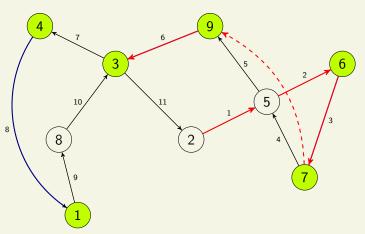
Take shortcuts to obtain final tour ${\mathcal C}$



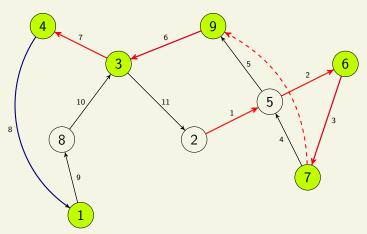
Take shortcuts to obtain final tour ${\mathcal C}$



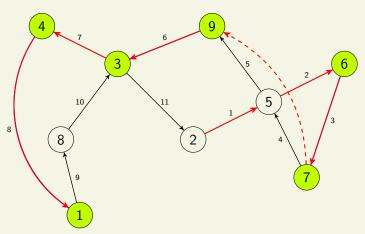
Take shortcuts to obtain final tour ${\mathcal C}$



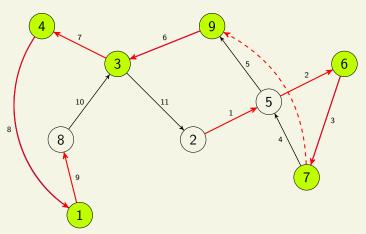
Take shortcuts to obtain final tour ${\mathcal C}$



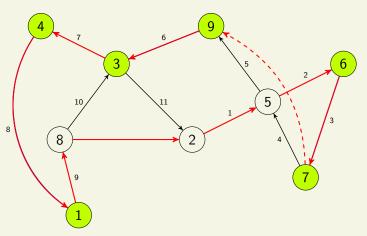
Take shortcuts to obtain final tour ${\mathcal C}$



Take shortcuts to obtain final tour ${\mathcal C}$



Take shortcuts to obtain final tour ${\mathcal C}$



Proof. Let τ be an optimal TSP tour on V.

• Let τ' be the tour on V' obtained by shortcutting τ .

Proof. Let τ be an optimal TSP tour on V.

- ullet Let au' be the tour on V' obtained by shortcutting au.
- By the triangle inequality, $cost(\tau') \leq cost(\tau)$.

Proof. Let τ be an optimal TSP tour on V.

- Let τ' be the tour on V' obtained by shortcutting τ .
- By the triangle inequality, $cost(\tau') \leq cost(\tau)$.
- Now look at τ' as the union of two perfect matchings on V', each consisting of alternate edges of τ' .

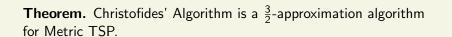
Proof. Let τ be an optimal TSP tour on V.

- Let τ' be the tour on V' obtained by shortcutting τ .
- By the triangle inequality, $cost(\tau') \leq cost(\tau)$.
- Now look at τ' as the union of two perfect matchings on V', each consisting of alternate edges of τ' .
- The cheaper of these two perfect matchings has cost $\leq \cot(\tau')/2 \leq \mathrm{OPT}/2$.

Proof. Let τ be an optimal TSP tour on V.

- Let τ' be the tour on V' obtained by shortcutting τ .
- By the triangle inequality, $cost(\tau') \leq cost(\tau)$.
- Now look at τ' as the union of two perfect matchings on V', each consisting of alternate edges of τ' .
- The cheaper of these two perfect matchings has cost $< cost(\tau')/2 < OPT/2$.

Therefore the optimal perfect matching has cost \leq OPT/2.



$$cost(C) \leq cost(T)$$

$$cost(C) \leq cost(T)
= cost(T) + cost(M)$$

$$cost(C) \leq cost(T)$$

$$= cost(T) + cost(M)$$

$$\leq OPT + \frac{1}{2}OPT$$

$$cost(C) \leq cost(T)
= cost(T) + cost(M)
\leq OPT + \frac{1}{2}OPT
= \frac{3}{2}OPT. \square$$

Proof. We have

$$cost(C) \leq cost(T)
= cost(T) + cost(M)
\leq OPT + \frac{1}{2}OPT
= \frac{3}{2}OPT. \square$$

Christofides' algorithm (1976) was the best approximation algorithm for Metric TSP until 2020!

In an 86 page paper, Karlin, Klein, and Gharan (2020) gave an algorithm with approximation ratio

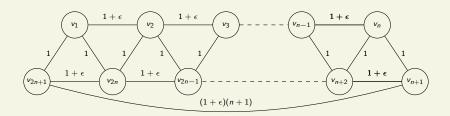
$$\frac{3}{2} - 10^{-36}$$

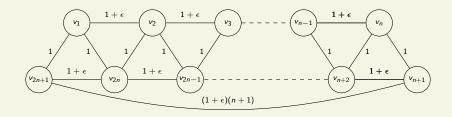
Proof. We have

$$cost(C) \leq cost(T)
= cost(T) + cost(M)
\leq OPT + \frac{1}{2}OPT
= \frac{3}{2}OPT. \square$$

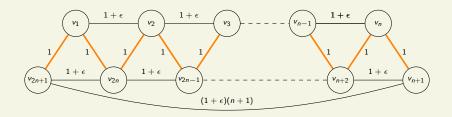
Christofides' algorithm (1976) was the best approximation algorithm for Metric TSP until 2020!

In an 86 page paper, Karlin, Klein, and Gharan (2020) gave an algorithm with approximation ratio

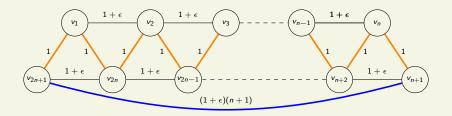




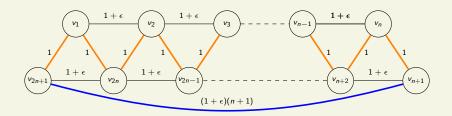
All other edge costs are the cost of the shortest path using the above edges.



The MST consists of the orange edges and has cost 2n.

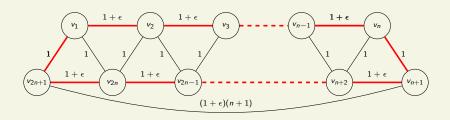


There are two odd degree vertices, v_{n+1} and v_{2n+1} . They are matched by the minimum cost perfect matching algorithm.

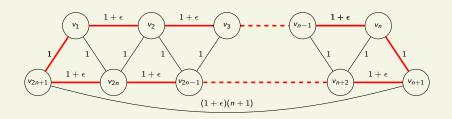


There are two odd degree vertices, v_{n+1} and v_{2n+1} . They are matched by the minimum cost perfect matching algorithm.

Tour cost: $2n + (1 + \epsilon)(n + 1) = (3 + \epsilon)n + (1 + \epsilon)$



The optimal tour has cost $(2n-1)(1+\epsilon)+2=2(1+\epsilon)n+(1-\epsilon)$.



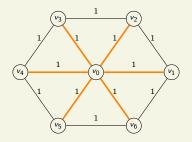
The optimal tour has cost $(2n-1)(1+\epsilon)+2=2(1+\epsilon)n+(1-\epsilon)$.

Performance ratio:

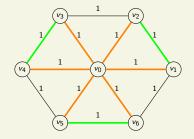
$$\frac{(3+\epsilon)n+(1+\epsilon)}{2(1+\epsilon)n+(1-\epsilon)}.$$

This approaches $\frac{3}{2}$ as $n \to \infty$ and $\epsilon \to 0$.

Let's see how Christofides' Algorithm performs on the tight instances for the 2-approximation algorithm.

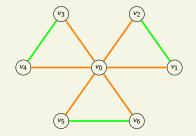


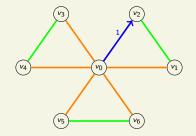
The MST consists of the orange edges.

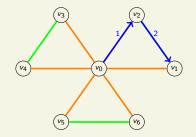


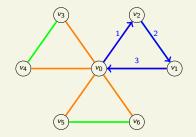
The MST consists of the orange edges.

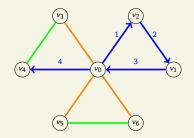
The green edges are a minimum cost perfect matching of the odd degree vertices in the MST.

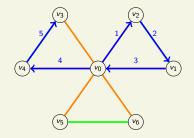


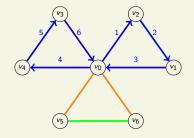


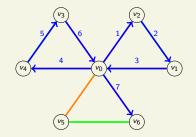


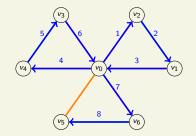


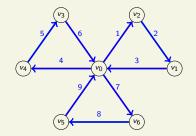


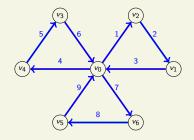


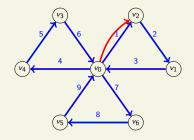


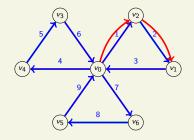


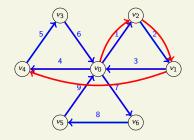


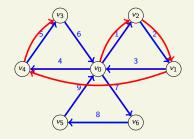


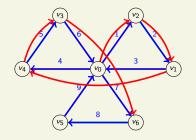


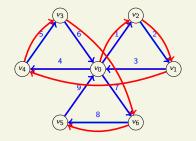


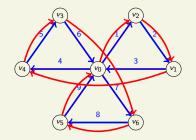


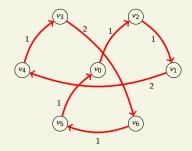




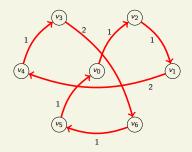






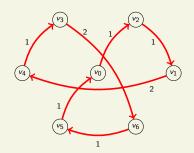


Christofides' algorithm tour cost: 9



Christofides' algorithm tour cost: 9

Optimal tour cost: 7

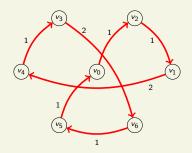


Christofides' algorithm tour cost: 9

Optimal tour cost: 7

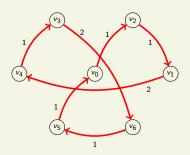
Performance ratio: $\frac{9}{7} \le \frac{3}{2}$

The performance depends on the Eulerian tour. With some Eulerian tours, the algorithm will obtain an optimal TSP tour.



For the analogous graph on n+1 vertices with n even, Christofides' algorithm will obtain a tour with cost at most

$$\underbrace{2}_{v_0 \text{ edges}} \cdot 1 + \underbrace{\frac{n}{2}}_{\text{edges for pairs edges between pairs}} \cdot 2 = \frac{3n}{2}.$$



For the analogous graph on n+1 vertices with n even, Christofides' algorithm will obtain a tour with cost at most

$$\underbrace{2}_{v_0 \text{ edges}} \cdot 1 + \underbrace{\frac{n}{2}}_{\text{edges for pairs edges between pairs}} \cdot 2 = \frac{3n}{2}.$$

The optimal tour has cost n+1, so the performance ratio is

$$\frac{\frac{3n}{2}}{n+1} = \frac{3}{2} \cdot \frac{n}{n+1}.$$