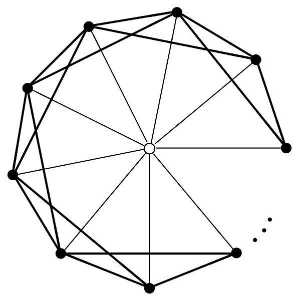


□

3.2 Metric TSP

The following is a well-studied problem in combinatorial optimization.

Problem 3.5 (Traveling salesman problem (TSP)) Given a complete graph with nonnegative edge costs, find a minimum cost cycle visiting every vertex exactly once.

In its full generality, TSP cannot be approximated, assuming $\mathbf{P} \neq \mathbf{NP}$.

Theorem 3.6 *For any polynomial time computable function $\alpha(n)$, TSP cannot be approximated within a factor of $\alpha(n)$, unless $\mathbf{P} = \mathbf{NP}$.*

Proof: Assume, for a contradiction, that there is a factor $\alpha(n)$ polynomial time approximation algorithm, $\mathcal{A}$, for the general TSP problem. We will show that $\mathcal{A}$ can be used for deciding the Hamiltonian cycle problem (which is **NP**-hard) in polynomial time, thus implying $\mathbf{P} = \mathbf{NP}$.

The central idea is a reduction from the Hamiltonian cycle problem to TSP, that transforms a graph G on n vertices to an edge-weighted complete graph G' on n vertices such that

- if G has a Hamiltonian cycle, then the cost of an optimal TSP tour in G' is n, and
- if G does not have a Hamiltonian cycle, then an optimal TSP tour in G' is of cost $> \alpha(n) \cdot n$.

Observe that when run on graph G', algorithm $\mathcal{A}$ must return a solution of cost $\leq \alpha(n) \cdot n$ in the first case, and a solution of cost $> \alpha(n) \cdot n$ in the second case. Thus, it can be used for deciding whether G contains a Hamiltonian cycle.

The reduction is simple. Assign a weight of 1 to edges of G, and a weight of $\alpha(n) \cdot n$ to nonedges, to obtain G'. Now, if G has a Hamiltonian cycle, then the corresponding tour in G' has cost n. On the other hand, if G has no Hamiltonian cycle, any tour in G' must use an edge of cost $\alpha(n) \cdot n$, and therefore has cost $> \alpha(n) \cdot n$. □

Notice that in order to obtain such a strong nonapproximability result, we had to assign edge costs that violate triangle inequality. If we restrict ourselves to graphs in which edge costs satisfy triangle inequality, i.e., consider *metric TSP*, the problem remains **NP**-complete, but it is no longer hard to approximate.

3.2.1 A simple factor 2 algorithm

We will first present a simple factor 2 algorithm. The lower bound we will use for obtaining this factor is the cost of an MST in G. This is a lower bound because deleting any edge from an optimal solution to TSP gives us a spanning tree of G.

Algorithm 3.7 (Metric TSP – factor 2)

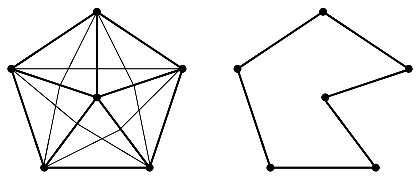
1. Find an MST, T, of G.
2. Double every edge of the MST to obtain an Eulerian graph.
3. Find an Eulerian tour, $\mathcal{T}$, on this graph.
4. Output the tour that visits vertices of G in the order of their first appearance in $\mathcal{T}$. Let $\mathcal{C}$ be this tour.

Notice that Step 4 is similar to the "short-cutting" step in Theorem 3.3.

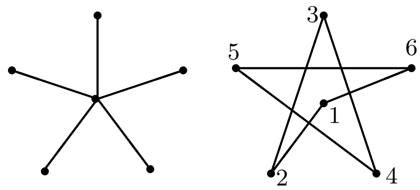
Theorem 3.8 *Algorithm 3.7 is a factor* 2 *approximation algorithm for metric TSP.*

Proof: As noted above, $\text{cost}(T) \leq \text{OPT}$. Since $\mathcal{T}$ contains each edge of T twice, $\text{cost}(\mathcal{T}) = 2 \cdot \text{cost}(T)$. Because of triangle inequality, after the "short-cutting" step, $\text{cost}(\mathcal{C}) \leq \text{cost}(\mathcal{T})$. Combining these inequalities we get that $\text{cost}(\mathcal{C}) \leq 2 \cdot \text{OPT}$. □

Example 3.9 A tight example for this algorithm is given by a complete graph on n vertices with edges of cost 1 and 2. We present the graph for $n = 6$ below, where thick edges have cost 1 and remaining edges have cost 2. For arbitrary n the graph has $2n-2$ edges of cost 1, with these edges forming the union of a star and an $n-1$ cycle; all remaining edges have cost 2. The optimal TSP tour has cost n, as shown below for $n = 6$:



Suppose that the MST found by the algorithm is the spanning star created by edges of cost 1. Moreover, suppose that the Euler tour constructed in Step 3 visits vertices in order shown below for $n = 6$:



Then the tour obtained after short-cutting contains $n-2$ edges of cost 2 and has a total cost of $2n-2$. Asymptotically, this is twice the cost of the optimal TSP tour. □

3.2.2 Improving the factor to 3/2

Algorithm 3.7 first finds a low cost Euler tour spanning the vertices of G, and then short-cuts this tour to find a traveling salesman tour. Is there a cheaper Euler tour than that found by doubling an MST? Recall that a graph has an Euler tour iff all its vertices have even degrees. Thus, we only need to be concerned about the vertices of odd degree in the MST. Let V' denote this set of vertices. $|V'|$ must be even since the sum of degrees of all vertices in the MST is even. Now, if we add to the MST a minimum cost perfect matching on V', every vertex will have an even degree, and we get an Eulerian graph. With this modification, the algorithm achieves an approximation guarantee of 3/2.

Algorithm 3.10 (Metric TSP – factor $3/2$)

1. Find an MST of G, say T.
2. Compute a minimum cost perfect matching, M, on the set of odd-degree vertices of T. Add M to T and obtain an Eulerian graph.
3. Find an Euler tour, $\mathcal{T}$, of this graph.
4. Output the tour that visits vertices of G in order of their first appearance in $\mathcal{T}$. Let $\mathcal{C}$ be this tour.

Interestingly, the proof of this algorithm is based on a second lower bound on OPT.

Lemma 3.11 *Let $V' \subseteq V$, such that $|V'|$ is even, and let M be a minimum cost perfect matching on V'. Then,* $\text{cost}(M) \le \text{OPT}/2$.

Proof: Consider an optimal TSP tour of G, say τ. Let τ' be the tour on V' obtained by short-cutting τ. By the triangle inequality, $\text{cost}(\tau') \le$

cost(τ). Now, τ' is the union of two perfect matchings on V', each consisting of alternate edges of τ. Thus, the cheaper of these matchings has cost $\leq$ cost$(\tau')/2 \leq$ OPT/2. Hence the optimal matching also has cost at most OPT/2. □

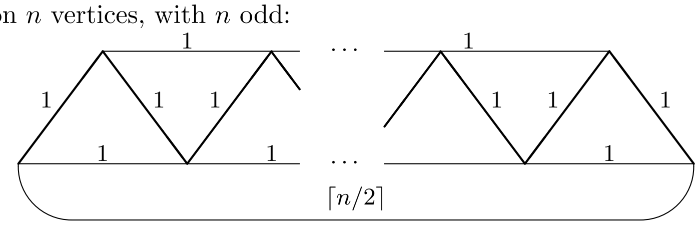
Theorem 3.12 *Algorithm 3.10 achieves an approximation guarantee of* 3/2 *for metric TSP.*

Proof: The cost of the Euler tour,

$$\mathrm{cost}(\mathcal{T}) \leq \mathrm{cost}(T) + \mathrm{cost}(M) \leq \mathrm{OPT} + \frac{1}{2}\mathrm{OPT} = \frac{3}{2}\mathrm{OPT},$$

where the first inequality follows by using the two lower bounds on OPT. Using the triangle inequality, $\mathrm{cost}(\mathcal{C}) \leq \mathrm{cost}(\mathcal{T})$, and the theorem follows. □

Example 3.13 A tight example for this algorithm is given by the following graph on n vertices, with n odd:



Thick edges represent the MST found in step 1. This MST has only two odd vertices, and by adding the edge joining them we obtain a traveling salesman tour of cost $(n-1) + \lceil n/2 \rceil$. In contrast, the optimal tour has cost n. □

Finding a better approximation algorithm for metric TSP is currently one of the outstanding open problems in this area. Many researchers have conjectured that an approximation factor of 4/3 may be achievable.