Metric TSP

- Recall that a function $d : [n]^2 \mapsto \mathbb{R}^+$ is called a metric if

$$\begin{aligned} d(i,j) &= d(j,i) && \forall i \neq j \\ d(i,j) &\leq d(i,k) + d(k,j) && \forall i,j,k \text{ distinct} \end{aligned}$$

- The problem METRIC-TSP-OPT is:

  Instance: A metric $d : [n]^2 \mapsto \mathbb{R}^+$
  Solution: A permutation $\sigma : [n] \mapsto [n]$
  Objective: Minimize

$$\sum_{i=1}^{n-1} d(\sigma(i), \sigma(i+1)) + d(\sigma(n), \sigma(1))$$

- An alternate formulation is: given a complete graph $G = (V, E)$ and a (metric) weight function $d : E \mapsto \mathbb{R}^+$, find a minimum-weight Hamiltonian cycle in G. Since $d(i,j) = d(j,i)$ we may as well assume that G is an undirected graph.

A 2-approximation algorithm

- The algorithm is as follows: given a complete, undirected graph $G = (V, E)$ with metric edge weights $d : E \mapsto \mathbb{R}^+$,
  1. Find a minimum spanning tree T of G (we'll consider T to be a set of edges, i.e. $T \subseteq E$)
  2. Let $T' = \{(i,j), (j,i) \mid \{i,j\} \in T\}$ be the graph obtained by replacing each edge of T with two directed edges.
  3. Note that $G' = (V, T')$ is Eulerian. Find an Euler tour C of G'.
  4. Now C visits every vertex of G, but it isn't a valid TSP tour because it visits vertices more than once. "Shortcut" C to get a valid tour C' as follows: suppose $C = v_1, \ldots, v_{2n-2}, v_1$. If there exist indices $1 \leq i < j \leq 2n-2$ such that $v_i = v_j$, then remove v_j from the list. The cost of the new cycle has increased by $d(v_{j-1}, v_{j+1})$, and has decreased by $d(v_{j-1}, v_j) + d(v_j, v_{j+1})$. Since d is a metric we have $d(v_{j-1}, v_{j+1}) \leq d(v_{j-1}, v_j) + d(v_j, v_{j+1})$, so the total cost has not increased by shortcutting. We repeat this process until we obtain a tour $v_1, \ldots, v_n, v_1$ with no repeated vertices. Call this tour C'
  5. C' is a valid TSP tour; the algorithm returns C'.

- Let OPT be an optimal TSP tour, i.e. a Hamiltonian cycle in G with minimum total cost. Let OPT$'$ be obtained from OPT by deleting any edge. Since the edge weights are positive we have $d(\text{OPT}') \leq d(\text{OPT})$.

Now OPT' is a Hamiltonian path, and is therefore a spanning tree of G. Since T is a minimum-cost spanning tree we have $d(T) \leq d(\mathrm{OPT}') \leq d(\mathrm{OPT})$. Also, since C traverses each edge of T' exactly once, if we view C as a walk in T it traverses each edge of T exactly twice, and thus $d(C) = 2 \cdot d(T) \leq 2 \cdot d(\mathrm{OPT})$. Finally, we have observed that $d(C') \leq d(C)$, so we have

$$d(C') \leq d(C) = 2 \cdot d(T) \leq 2 \cdot d(\mathrm{OPT})$$

which proves the approximation ratio of 2.

A 3/2-approximation algorithm (Christofides' algorithm)

- The algorithm is a modification of the previous one: given an undirected, complete graph $G = (V, E)$ with metric edge weights $d : E \mapsto \mathbb{R}^+$,
    1. Find a minimum spanning tree T of G
    2. Let $S \subseteq V$ be the set of vertices having odd degree in T. Note that $|S|$ is even.
    3. Find a minimum-weight perfect matching M on the vertices of S. Recall that a matching is a set of edges that do not share any endpoints, and a matching is called perfect if every vertex is an endpoint of one edge in the matching. Since S contains an even number of vertices, and G contains all edges between vertices in S, S has a perfect matching. Finding a minimum-weight perfect matching can be done in polynomial time, but the algorithm is beyond the scope of this course.
    4. Let E' be the union, with repetitions, of T and M (i.e. if an edge is in both T and M we'll include it twice in E'). Let $G' = (V, E')$ be the corresponding multigraph (a multigraph is like a graph, except that multiple edges between pairs of vertices are allowed).
    5. Now G' is an undirected Eulerian multigraph, since we added one extra edge to every odd-degree vertex in T and the degrees of the other vertices are unchanged. Find an Euler tour C of G'.
    6. Shortcut C to get a valid TSP tour C', and return C'.
- From the analysis of the 2-approximation algorithm, we have that $d(T) \leq d(\mathrm{OPT})$. Thus $d(C') \leq d(C) = d(T) + d(M) \leq d(\mathrm{OPT}) + d(M)$. To prove an approximation ratio of 3/2 we need to show that $d(M) \leq d(\mathrm{OPT})/2$.
- Let OPT' be a tour of S obtained by starting with OPT, and shortcutting past the vertices that are not in S. Then $d(\mathrm{OPT}') \leq d(\mathrm{OPT})$, since d is a metric.
- OPT' can be decomposed into two matchings M_1 and M_2: suppose that OPT' traverses the edges $e_1, e_2, \ldots, e_{|S|}$, and let $M_1 = \{e_1, e_3, \ldots, e_{|S|-1}\}$ and $M_2 = \{e_2, e_4, \ldots, e_{|S|}\}$. Note that both M_1 and M_2 are perfect matchings on S.

- Since $d(\mathrm{OPT}') = d(M_1) + d(M_2)$ we have that either $d(M_1) \leq d(\mathrm{OPT}')/2$ or $d(M_2) \leq d(\mathrm{OPT}')/2$. As M is a minimum-weight perfect matching on S,

$$d(M) \leq \min\{d(M_1), d(M_2)\} \leq d(\mathrm{OPT}')/2 \leq d(\mathrm{OPT})/2$$

- Thus $d(C') \leq d(M) + d(T) \leq (3/2) \cdot d(\mathrm{OPT})$, which proves the approximation ratio.