COMS E6232 - Problem Set #2

Alex Wong (asw2181) - asw2181@columbia.edu

March 2, 2017

Problem 1

1.Consider an instance where we have two items: i and j. Item i has $s_i = \epsilon$

It is obvious to see the OPT solution; the last 16 nodes of set L connect to a distinct node in set R, meaning that the OPT solution must contain at least 16 nodes. We also see that all 16 nodes of set R has edges to all the nodes in set L. This shows that the OPT solution would be selecting every node from set R. Thus, the approximation ratio of Greedy on this graph is $33/16 = 2.0625$.

2. Consider the graph G from part 1. We can say that set R contains k^2 nodes, all of degree $\leq$ k. We can partition set L into k sets, $L_1, ..., L_k$, where each set L_i contains $\lfloor \frac{k^2}{i} \rfloor$ nodes of degree i. From these sets, we see that $|L_1| = |R|, |L_2| = \frac{1}{2}|R|, |L_3| = \lfloor \frac{1}{3}|R| \rfloor$, and $|L_4| = \frac{1}{4}|R|$. Adding $R + \frac{1}{2}R + \lfloor \frac{1}{3}R \rfloor + \frac{1}{4}R$ forms the start of the harmonic series, which shows that $|L| \approx \ln k|R|$. Thus, this shows that the approximation ratio of Greedy is not bounded by any constant; rather, the approximation ratio is $\Omega(\log n)$ where n is the number of nodes.

Problem 2

1. **Lemma 2.1** Let T be the minimum spanning tree of the weighted complete graph G. Then cost(T) $\leq$ OPT, where OPT is the cost of the optimal s-t path.

Proof: We can prove this by contradiction. Suppose that cost(T) $>$ OPT. We know that OPT must visit every node to create a path. Thus, if OPT $<$ cost(T), it violates the very definition of a minimum spanning tree, proving that cost(T) $>$ OPT can not be true. Thus, cost(T) $\leq$ OPT. $\square$

2. **Lemma 2.2** Let U be the subset of nodes consisting of the nodes of $\{1, ..., n\}$-$\{s, t\}$ that have odd degree in T and the nodes in $\{s, t\}$ that have even degree in T. Then U has an even number of nodes.

Proof: Suppose we partition U into 2 subsets: U_{odd} and U_{even} where they contain nodes that have odd degree and nodes that have even degree, respectively. We know that in the minimum spanning tree T, there is an even number of nodes that have odd degree since $\sum\limits_{v \in T_{odd}} degree(v) = 2 \cdot \#edges$. Knowing this, if $|U_{odd}|$ is odd, that means either s or t must have an odd degree, meaning the other must have an even degree and would be contained in U_{even}, so $|U_{even}| = 1$. Thus,

$$|U| = |U_{odd}| + |U_{even}| = odd\# + 1 = even\#$$

which shows that $|U|$ is even.

If $|U_{odd}|$ is even, then s and t must both have odd degrees or both have even degrees; if both are odd, then $|U_{even}| = 0$ which trivially shows $|U|$ is even. If both are even, then $|U_{even}| = 2$ and

$$|U| = |U_{odd}| + |U_{even}| = even\# + 2 = even\#$$

which shows that $|U|$ is even. It must be noted that if $|U_{even}| = 2$ there must be other nodes in set U; U can never just contain s and t. Thus, the set U has an even number of nodes. $\square$

3. **Lemma 2.3** Let $G[U]$ be the subgraph of G induced by the subset U, M be a minimum-cost perfect matching in $G[U]$, and P^* be a minimum-cost path from s to t that visits every node exactly once. Consider the nodes of U in the order that they appear on the path P^* and let u_i be the i-th node of U in this ordering: $i = 1, ..., |U|$. Color red the edges of P^* on the subpath from u_1 to u_2,..., from u_{2k-1} to u_{2k}, where $|U| = 2k$. Color blue the other edges of P^*. Then, cost(red edges) $\geq$ cost(M).

Proof: We can prove this by contradiction. Suppose cost(red edges) $<$ cost(M). We know that red edges are essentially a perfect matching on set U. If cost(red edges)

$<$ cost(M), then M is not a minimum-cost perfect matching on set U, proving that cost(red edges) $<$ cost(M) can not be true. Thus, cost(red edges) $\geq$ cost(M). $\square$

4. **Lemma 2.4** cost(T) + cost(blue edges) $\geq 2 \cdot$ cost(M).

Proof: We first see that $T \cup$ (blue edges of P^*) creates a connected Eulerian multigraph. We know this to be true because every vertex has an even degree. A Eulerian multigraph is essentially a multigraph that contains an Eulerian tour. Within a Eulerian tour, we know that we can create two disjoint perfect matchings of U and the sum of their costs $\leq$ cost of the Eulerian tour due to triangle inequality. We also know that both of the disjoint perfect matchings can not have a cost $<$ cost(M) or else M can't be a minimum-cost perfect matching on U. Thus,

$$2 \cdot cost(M) \leq cost(2\ disjoint\ perfect\ matchings) \leq cost(T) + cost(blue\ edges)$$

$$cost(T) + cost(blue\ edges) \geq 2 \cdot cost(M)$$

$\square$

5. **Theorem 2.5** The variant of Christofides' algorithm achieves an approximation factor of 5/3 for the *s-t path metric TSP* problem.

Proof: Combining our conclusions from Lemma 2.3 and 2.4, we can say that:

$$cost(T) + cost(blue\ edges) + cost(red\ edges) \geq 3 \cdot cost(M) \tag{1}$$

We also know that $cost(blue\ edges) + cost(red\ edges) = P^* = OPT$ and from Lemma 2.1 that $cost(T) \leq OPT$. Thus, we can modify equation 1 to be:

$$cost(T) + OPT \geq 3 \cdot cost(M)$$

$$2 \cdot OPT \geq 3 \cdot cost(M)$$

$$cost(M) \leq 2/3\ OPT \tag{2}$$

Adding cost(T) to both sides of equation 2 and using Lemma 2.1 that $cost(T) \leq OPT$, we finally get:

$$cost(T) + cost(M) \leq cost(T) + 2/3\ OPT$$

$$cost(T) + cost(M) \leq OPT + 2/3\ OPT$$

$$cost(T) + cost(M) \leq 5/3\ OPT$$

Thus showing that the cost of the path computed by the variant of Christofides' algorithm is at most 5/3 times the cost of the optimal path. $\square$

6.