**Advanced Algorithms 2024** 18.11, 2024

Sample Solutions 09

*Lecturer: Bernhard Haeupler* *Teaching Assistant: Jakob Nogler*

# 1 Tree embedding in cycle

1. Note that we consider all trees $T$ defined on the same vertex set $G$ with non-negative lengths of edges that satisfy

   i) $\forall i, j : d_G(i,j) \leq d_T(i,j)$,

   ii) $\forall i, j : d_T(i,j) \leq C \cdot d_G(i,j)$ for some value $C$.

   Consider all trees minimizing the stretch $C$ and among all of them choose $T^*$ that minimizes $\sum_{\{u,v\}\in T^*} d_{T^*}(u,v)$. First, we claim that in such a tree all lengths of edges in $T$ have the length of the corresponding shorter arc in $G$.

   **Claim 1.** *All edges $\{i,j\} \in T^*$ have length exactly $d_G(i,j)$.*

   *Proof.* Assume for contradiction that not all edges $\{i,j\} \in T^*$ have length exactly $d_G(i,j)$. Then there exists an edge $\{i,j\} \in T^*$ with $d_{T^*}(i,j) > d_G(i,j)$ and we can strictly decrease its length to $d_G(i,j)$ and get a new tree $T'$. Since the distances in $T'$ are only reduced, property ii) still holds trivially for $T'$. We now show that property i) is also satisfied. Consider any two vertices $u, v$. The length of the shortest path $P$ between $u$ and $v$ in $T'$ is the sum of all distances between intermediate vertices on unique path joining them. Since each distance between intermediate vertices in $T'$ is at least the length of the arc in $G$, we get

   $$d_{T'}(u,v) = \sum_{\{i,j\}\in P} d_{T'}(i,j) \geq \sum_{\{i,j\}\in P} d_G(i,j) \geq d_G(u,v),$$

   i.e., property i) also holds. Since properties i) and ii) are satisfied, we receive a valid tree $T'$ with a smaller sum $\sum_{\{i,j\}\in T'} d_{T'}(i,j)$, which is a contradiction to the definition of $T^*$. □

   Next, we claim that in an optimal tree $T^*$ there is no "bend".

   **Claim 2.** *In $T^*$ there are no three vertices $u, v, w \in V(T)$ such that $\{u,w\}, \{v,w\} \in E(T)$ and the shorter arcs between $u, w$ and between $v, w$ in $G$ are one subset of the other.*

   *Proof.* Assume there exists a vertex $w \in V$ and edges $\{u,w\}, \{v,w\} \in E(T)$ as in the statement of the claim. Without loss of generality we assume $vw \subset uw$, i.e., $d_G(u,v) + d_G(v,w) = d_G(u,w)$. Then we can remove the edge $\{u,w\}$ from $T$ and replace it with an edge between $u$ and $v$ of length $d_G(u,v)$. Note that after this operation the new graph $T'$ is still a tree. Consider any two vertices $s, t \in V(T^*)$. Note that $d_{T'}(s,t) \leq d_{T^*}(s,t)$, since if the path did not the edge $\{u,w\}$ it will stay the same in $T'$, while if the path used the edge $\{u,w\}$ in $T^*$, we replace $\{u,w\}$ by the pair of edges $\{u,v\}, \{v,w\}$ and get a walk of the same length in $T'$. Hence, property ii) still holds in $T'$. Moreover, the property i) holds by the same argument as in the previous claim. Since $T'$ has a smaller sum $\sum_{\{u,v\}\in T'} d_{T'}(u,v)$ compared to $T^*$, we get again a contradiction to our choice of $T^*$. □

From the second claim it follows that $T^*$ is a path, since otherwise a vertex with degree at least 3 would yield the forbidden configuration from its statement. Moreover, combining the two claims, we conclude that the path $T^*$ "wraps" clockwise or counterclockwise around the cycle, i.e., if we call its vertices $u_1, u_2, \ldots, u_n$ with an edge between $u_i$ and $u_{i+1}$ of length $d_G(u_i, u_{i+1})$, then either for all $i$ we have the length equal to the length of the corresponding clockwise arc, or for all $i$ the length is equal to the length of the corresponding counterclockwise arc. Hence, assuming the path is wrapping around the circle in clockwise direction, starting at the first vertex $u_1$ of the path, the path between vertices $u_1$ and $u_1 - 1$ in $T^*$ has length at least $n - 1$ because the path needs to wrap around the entire circle. Consequently, with the deterministic tree embedding approach we need $C \geq (n - 1)$ which is achieved by a path that we get by removing any edge from the cycle. Therefore, the deterministic embedding cannot be made to yield small stretch factor.

2. Consider a distribution over trees that we get by by removing uniformly at random one of the $n$ cycle edges and setting all lenghts of the given path to 1. We obtain a path $T$ in which all distances between vertices are clearly at least as large as in the cycle. Therefore property i) holds. For any two vertices $u, v$ in $T$, the removed edge is on the shortest path between $u$ and $v$ in the cycle with probability $\frac{d_G(u,v)}{n}$. If this happens, the shortest distance between $u$ and $v$ in $T$ equals $n - d_G(u, v)$. If the removed edge lies on the longer path in $G$, their distance in $T$ equals $d_G(u, v)$. Hence, for the expected distance between any pair of vertices $u, v$ in $T$ we have

$$\begin{aligned} \text{Exp}[d_T(u,v)] &= \underbrace{\frac{d_G(u,v)}{n}(n - d_G(u,v))}_{\text{removed edge on shortest path}} + \underbrace{\frac{n - d_G(u,v)}{n} d_G(u,v)}_{\text{removed edge on longer path}} \\ &\leq d_G(u,v) + d_G(u,v) \leq 2d_G(u,v) \end{aligned}$$

Therefore, there exists a distribution over tree embeddings that achieves a stretch 2 in a cycle of length $n$.

## 2 Steiner Forest

We use the tree embedding algorithm from the lecture (chapter 5.1) in order to create a tree $T$, on the same set of vertices than $G$, that has two properties:

i) for all vertices $u, v : d_G(u, v) < d_T(u, v)$

ii) for all vertices $u, v : \mathbb{E}(d_T(u, v)) \leq O(log(n)) \cdot d_G(u, v)$

We now can find the shortest paths between all terminals $(s_i, t_i)$ in $T$; their union gives the optimal algorithm for the tree $T$, i.e., $OPT(T)$. Now for every edge $e = \{u, v\}$ in $OPT(T)$ we consider the shortest path from $u$ to $v$ in $G$ and add it to $ALG(G)$. We claim that $ALG(G)$ is at most $O(log(n)) \cdot OPT(G)$ in expectation. First, we observe that

$$ALG(G) \leq OPT(T)$$

because of property i) of the tree. In other words, each edge between two vertices $u$ and $v$ in the optimal solution in $T$ is replaced by a path in $G$ that has length at most $d_T(u, v)$. So, their union has total length at most $OPT(T)$. Next, using property $ii)$ of the tree we have that

$$\mathbb{E}[OPT(T)] \leq O(\log(n)) \cdot OPT(G).$$

To see this, we define a solution $S$ in $T$ as follows: For each edge $e = \{u, v\}$ in $OPT(G)$, the path from $u$ to $v$ in $T$ is added to $S$. Note that every edge in $OPT(G)$ is replaced by a path of distance at most $O(log(n)) \cdot d_G(u,v)$ in expectation. Hence, the total distance in $S$ is at most $O(log(n)) \cdot OPT(G)$ in expectation. Finally, $\mathbb{E}[OPT(T)] \leq \mathbb{E}[S]$ as the optimal solution is always at most a specific solution. Putting the above inequalities together yields

$$\mathbb{E}[ALG(G)] \leq \mathbb{E}[OPT(T)] \leq O(\log(n)) \cdot OPT(G),$$

as needed.

# 3 Analyze the Ball-Carving with Exponential Clocks

1. In this exercise we are asked to give an upper bound $UB$ on the diameter($= 2 \cdot r_v$) of every ball $B_v$ with high probability. In other words we search $UB$ such that

$$\Pr(\forall v \in G : 2 \cdot r_v \leq UB) \geq 1 - \frac{1}{\text{poly}(n)}.$$

We will show the following statement which is equivalent to the previous one by negating.

$$\Pr(\exists v \in G : 2 \cdot r_v > UB) \leq \frac{1}{\text{poly}(n)}. \tag{1}$$

First of all, we know that the radius of a vertex follows the exponential distribution with density function $f(x) = \beta \cdot e^{-\beta x}$. Hence, $\Pr(r_v > x) = e^{-\beta x}$. Therefore, we can write

$$\Pr(\exists v : 2 \cdot r_v > UB) \leq n \cdot \Pr(\text{fixed } \ v : 2 \cdot r_v > UB) = n \cdot \Pr\left(r_v > \frac{UB}{2}\right) = n \cdot e^{-\frac{\beta \cdot UB}{2}}$$

Now, we would like to have that

$$n \cdot e^{-\frac{\beta \cdot UB}{2}} \leq \frac{1}{\text{poly}(n)} \tag{2}$$

By taking the logarithm of both sides and isolating $UB$ on one side, (2) can be written as

$$-UB \leq \frac{2 \cdot \log\left(\frac{1}{\text{poly}(n) \cdot n}\right)}{\beta}$$

which can be further transformed into

$$UB \geq \frac{2 \cdot \log(\text{poly}(n) \cdot n)}{\beta} = \Theta\left(\frac{\log(n)}{\beta}\right)$$

Finally, $UB = \Theta\left(\frac{\log(n)}{\beta}\right)$ satisfies (1). Therefore, (1) is also satisfied and we found tight upper bound on the diameter of each $B_v$.

2. Let us assume that a vertex $w$ lies on the shortest path from $u$ to $v$ but $w \in B_z$ for some $z \neq v$. By the choice of $B_z$ we know that

$$(r_z - d(z,w)) > (r_v - d(v,w)).$$

Furthermore,

$$d(u,v) = d(u,w) + d(w,v)$$

because $w$ lies on the shortest path from $u$ to $v$. Also,

$$d(z,u) \leq d(z,w) + d(w,u)$$

by the triangle inequality. Combining these three properties we get

$$(r_z - d(z,u)) \geq (r_z - d(z,w) - d(w,u)) > (r_v - d(v,w) - d(w,u)) = (r_v - d(v,u)).$$

Therefore, $u$ should also be in $B_z$ which is a contradiction.