Metric Techniques for Approximation Algorithms **Hints and Answers for Exercises 1**
Date: August 18, 2008

2. Embedding a Metric into a (Single) Tree.

a. Minimum spanning tree.

b. View the cycle drawn as a regular polygon, and path as being wrapped along the cycle. What if for every cycle node, the edges go to different "halves" of the cycle like this:

What if there is a node where the edges go to the same "half" of the cycle" (like this)?

Can you get a path with no larger distortion where nodes don't have this property?

c. If there is a node with degree 3 or higher, show that you can do something similar to the second case above.

d. The 3-cycle. In fact, any n-cycle — you can get about $n/3$ distortion.

e. Remove the Steiner nodes. Now use (c) above.

f. The outer cycle has length $t = 4(\sqrt{n} - 1)$ and embeds with distortion 2 into C_t.

3. Tree Embeddings and Approximation Algorithms.

a. Any Euler tour of this graph.

b. Let $L^* = \sum_{i=1}^{n} d(\pi(i), \pi(i+1))$ be the length of the optimal permutation π^* on the graph G. By linearity of expectation, the expected length of this tour in the tree is at most αL^*. Hence, the expected cost L_T of the tour π_T optimal on the tree (which we can find using (a)) is at most αL^*.

By the domination property, the length of the tour π_T is only smaller on G, and hence we have a solution of expected cost α times the optimal TSP on G.

c. We used the "dominating" property when translating back the solution.

The proof fails with the weaker property. (Do you see a counterexample?)

4. A Better Padding Guarantee.

a. First, we assume that $\rho < \Delta/8$—if not, then the probability is clearly at most 1, which is at most $8\frac{\rho}{\Delta}$ and the theorem clearly holds.

Now consider the same analysis as in the lecture: the probability that the j^{th} node from x is "responsible for" cutting the ball $\mathbf{B}(x,\rho)$ is $\frac{8\rho}{\Delta} \cdot \frac{1}{j}$. However, the only nodes that can be responsible for cutting the ball $\mathbf{B}(x,\rho)$ are those which are at distance at most $\Delta/2 + \rho < \Delta$ from x. Similarly, none of the nodes at distance $< \Delta/8 \leq \Delta/4 - rho$ can be responsible for cutting the ball $\mathbf{B}(x,\rho)$. Hence, when we sum over the relevant nodes, we get the probability is at most $\frac{8\rho}{\Delta} \cdot \beta$, where

$$\beta = \beta(x,\Delta) = O\left(\log \frac{|\mathbf{B}(x,\Delta)|}{|\mathbf{B}(x,\Delta/8)|}\right).$$

b. Let us say *phase* i is when the tree-construction procedure is called with parameter i. The expected distance between x, y in the random tree is:

$$\begin{aligned}
&\sum_i \Pr[x,y \text{ separated in phase } i \mid \text{ they are not separated in phase } \geq i] \times 2^{i+1} \\
&\leq \sum_i \frac{d(x,y)}{2^{i-1}} \cdot \beta(x, 2^{i-1}) \times 2^{i+2} \\
&= 8\, d(x,y) \sum_i \beta(x, 2^{i-1}) \\
&\leq 8\, d(x,y) \left(\log|\mathbf{B}(x,2^\delta)| + \log|\mathbf{B}(x,2^{\delta-1})| + \log|\mathbf{B}(x,2^{\delta-2})|\right) \\
&\leq 24\, d(x,y)\, \log n.
\end{aligned}$$

5. A Lower Bounding Technique.

a.

b. The (expected) distortion for edge $\{u,v\}$ when trees are drawn from the probability distribution $\mathbf{q}$ is given by

$$\mathbf{E}_{T\leftarrow\mathbf{q}}[\frac{d_T(u,v)}{d(u,v)}]$$

And hence we are interested in the distribution that minimizes the (expected) distortion

$$\min_{\mathbf{q}} \max_{\{u,v\}} \mathbf{E}_T[\frac{d_T(u,v)}{d(u,v)}]$$

By the above inequality, this is lower bounded by

$$\max_{\{u,v\}\in\mathbf{p}} \min_T \mathbf{E}_{\{u,v\}}[\frac{d_T(u,v)}{d(u,v)}]$$

I.e., for any probability distribution over the edges, if we can show that any tree $T \in \mathcal{T}$ incurs an expected distortion least β, that is a lower bound for embedding into distributions over trees.

c. Consider the uniform distribution over the edges of C_n; i.e., π picks each of the n edges with probability $1/n$. Then, for any tree T, there is at least one edge that incurs a distortion of $n/3 - 1$, and even if the others incur a distortion of 1, we get

$$\mathbf{E}_{\{u,v\}}[\frac{d_T(u,v)}{d(u,v)}] \geq \frac{1}{n} \times (\frac{n}{3} - 1) + (1 - \frac{1}{n}) \times 1 = \frac{4}{3} - o(1).$$

d. The question basically outlines the answer. The sum giving the total distortion for the edges is

$$2^{k-1} \times (2^{k+1}/3 - 1) + 2^k \times (2^k/3 - 1) + \dots$$

which is $\Omega(k2^{2k})$. Since we are taking the uniform distribution over the 4^k edges, the average value is $\approx k/3 - O(1)$.

The only worry is that we are counting the same edges repeatedly. However, for an edge with distortion 2^j, the above sum counts a distortion of $2^j + 2^{j-1} + 2^{j-2} + \ldots \leq 2 \cdot 2^j$—hence all we have to do is to divide the above value by 2, which still gives a lower bound of $\Omega(k)$.

6. Covers.

a. Draw $2\log_2 n$ random trees from the FRT distribution $\mathcal{D}$. Consider any pair $x, y \in V$; the probability that one random tree drawn from $\mathcal{D}$ incurs a distortion of more than 2α is at most $\frac{1}{2}$; hence, the probability that at least one of $2\log_2 n$ trees incurs a distortion $\leq 2\alpha$ is at least $1 - \frac{1}{n^2}$. Now a union bound over all $\binom{n}{2}$ pairs suffices.

b. Use the β-padded decomposition procedure with parameter $2\beta r$ for $t = O(\log n)$ times, and take the union of the clusters as the subsets. Each vertex lies in exactly $t = O(\log n)$ subsets, one for each partitioning (giving property (b)); each subset has diameter at most $2\beta r$ by the properties of the padded decomposition (giving property (c)); finally, since $\mathbf{B}(x, r)$ is not split by one padded decomposition with probability $\frac{r}{2\beta r}\beta = \frac{1}{2}$, a union bound ensures that all the balls $\{\mathbf{B}(x, r)\}_{x \in V}$ are not split in at least one of the t repetitions with constant probability.

7. Small Support Distributions. Consider the linear program which tries to find a distribution that minimizes the expected distortion: there's a variable $x_T \geq 0$ for each tree T in $\mathcal{T}$ (the set of trees $T = (V, E_T)$ that dominate $\mathsf{M} = (V, d)$), that captures the probability of T being output. Hence we have

$$\sum_T x_T = 1$$

and for each $\{x, y\} \in \binom{V}{2}$, we have a constraint

$$\sum_T x_T \, d_T(x, y) \leq \lambda \, d(x, y)$$

finally, we want to minimize λ. The FRT procedure implies that there is a feasible solution of value $\lambda = O(\log n)$. Since there are only $\binom{n}{2} + 1 = O(n^2)$ constraints in this LP, there must be a solution of value $O(\log n)$ that has only $O(n^2)$ non-zero variables, implying the existence of a distribution with small support.

(We claimed this LP was exponentially sized—not that it matters, but why can we consider only exponentially many trees in $\mathcal{T}$? As defined, there could be infinitely many trees.)