Today

- $O(\log n)$ - Distortion Probabilistic Tree Embeddings
- Tree structure
- Algorithm
- Analysis

(Bartal next time?)

Recall

Sparsification Framework (to solve graph problem on G)

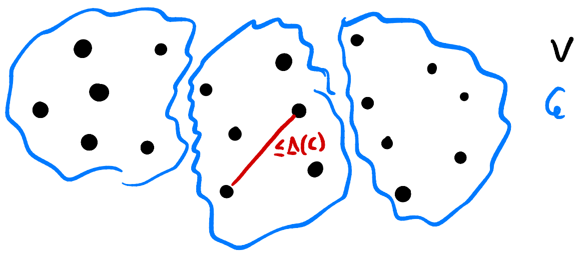
α) Find simple graph H (approximately) preserving structure of problem on G → Today: PTEs, application on hw.

β) Solve problem on H

γ) Convert solution on H to solution on G

A Clustering of V is a partition of V into $\mathcal{C} = \{C_1, C_2, C_3, \dots\}$

The diameter of C_i is $\Delta(C_i) = \max_{u,v \in V_i} d(u,v)$ and of clustering $\mathcal{C}$ is $\Delta(\mathcal{C}) = \max_i \Delta(C_i)$



An embedding of metric space (V,d) into metric space (V',d') is a function $f: V \to V'$

f has distortion α if $d(u,v) \le d'(f(u), f(v)) \le \alpha \cdot d(u,v)$ $\forall u,v \in V$

$$\sum_{i=1}^{n} \frac{1}{i} = O(\log n)$$

$$\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$$

Last time: every graph is a tree wrt min s-t cuts

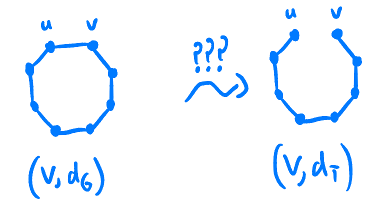
This time: every graph/metric is (approximately) a tree wrt distances

Say metric (V,d) is a <u>tree metric</u> if it embeds isometrically into (V,d_T)

where d_T is the shortest path distances in some edge-weighted tree $T=(V,E,w)$

<u>Goal</u>: embed an arbitrary metric into a tree metric w/ low distortion (b/c problems on trees are easier)

<u>Impossibility of Goal w/ $o(n)$ Distortion (Assuming $T \subseteq G$)</u>

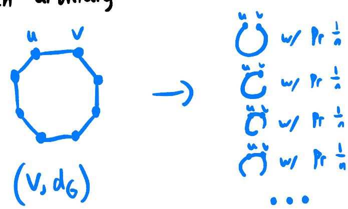


$d_G(u,v)=1$ and $d_T(u,v)\geq 1$

but $d_T(u,v) = \Omega(n)$

Also impossible even if $T \nsubseteq G$ -> See later in notes

Modified Goal: embed an arbitrary metric into a distribution over tree metrics w/ low distortion



$d_G(u,v)=1$ and $d_T(u,v)\geq 1$ $\forall T$

and $\mathbb{E}[d_T(u,v)] = \frac{n-1}{n}\cdot 1 + \frac{1}{n}\cdot\Omega(n) = O(1)$

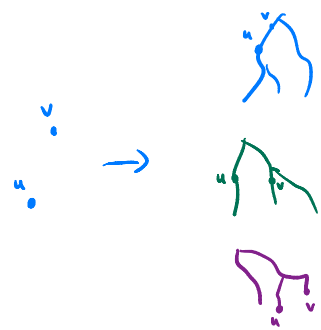
(generalizes to all pairs)

An <u>α-distortion probabilistic tree embedding</u> of (V,d) is a distribution over trees $\mathcal{T}$ containing V s.t. $\forall u,v \in V$

1) $d(u,v) \leq d_T(u,v)$ $\forall T \in \mathcal{T}$

2) $\mathbb{E}_{T\sim\mathcal{T}}[d_T(u,v)] \leq \alpha \cdot d(u,v)$

$\forall$ over trees w/ Pr>0



"FRT"

<u>Theorem</u>: every metric has a (poly-time sampleable) $O(\log n)$-distortion tree embedding

(optimal by expanders)

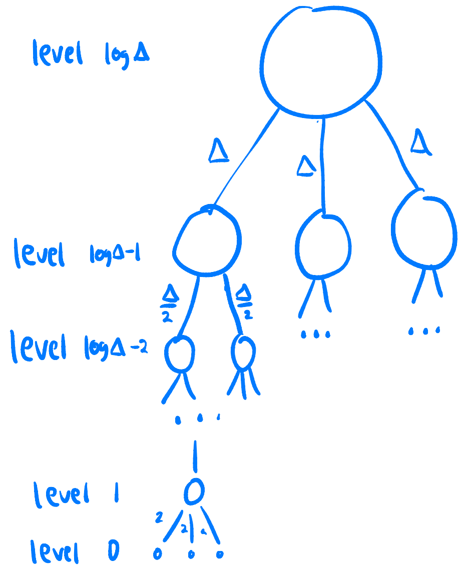
Structure of $T \in \mathcal{T}$

WLOG $d(u,v) \ge 1$ $\forall u,v \in V$ and let $\Delta := 2^x$ be min. power of 2 s.t. $d(u,v) \le \Delta$ $\forall u,v \in V$

T has

- $1 + \log\Delta$ levels
- Level i ~ a 2^{i+1}-diameter Clustering of V
  - ↳ level 0 ~ $\{\{v\} : v \in V\}$
  - ↳ level $\Delta+1$ ~ $\{V\}$
- Edge from level i to $i-1$ has length 2^i

(Each level really an LDD)



Will be clear our trees have this structure so can already show lower bound

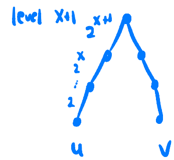
Proof of (1) for trees w/ above structure

Consider $u,v \in V$; let x be max $x \in \mathbb{Z}$ s.t. $2^{x+1} < d(u,v)$ so $2^{x+2} \ge d(u,v)$

$d_T(u,v)$ minimized if u,v in same cluster for min possible level

u,v cannot be in same cluster at level x (by diameter bound)

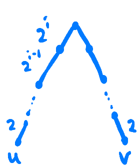
so the lowest level at which u,v in same partitions is $x+1$



so $d_T(u,v) \ge 2 \cdot \sum_{i=1}^{x+1} 2^i \ge 2 \cdot 2^{x+1} = 2^{x+2} \ge d(u,v)$

Can also make progress towards (2) just w/ above structure

Claim: if u,v have least common ancestor (lca) at level x in T then $d_T(u,v) \le 2^{x+2}$



$$d_T(u,v) = 2 \cdot \sum_{i=1}^{x} 2^i = 2 \cdot 2^x \cdot \sum_{j=1}^{x} \frac{1}{2^{j-1}} = 2 \cdot 2^x \sum_{j=0}^{x-1} \left(\tfrac{1}{2}\right)^j \underset{\sum_{i=0}^{\infty} r^i \le \frac{1}{1-r}}{\le} 2^{x+2}$$

Now need to talk about how to generate each tree

$\subseteq V$ Permutation of V $\in \mathbb{R}$

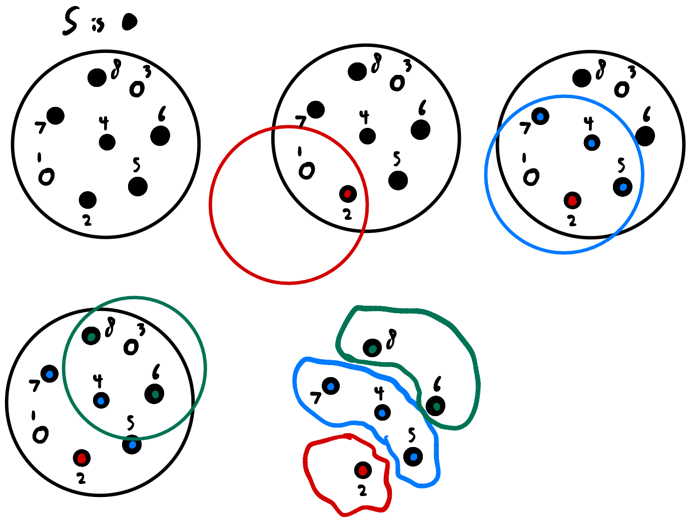
Partition(S, π, r)

For each $u \in V$ in order of π

  Let $B_u := S \cap B(u, r)$

  $S \leftarrow S \setminus B_u$

Return $\{B_u : u \in V\}$



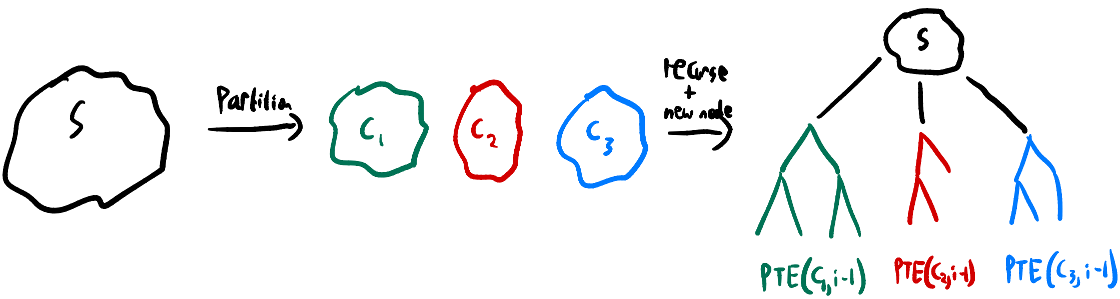
PTE(S, π, r_0, i)

Let $r_i := 2^i \cdot r_0$

$\mathcal{C} \leftarrow$ Partition(S, π, r_i)

Create node S as root

Add length 2^i edge from S to root of PTE$(C, \pi, r_0, i-1)$ $\forall C \in \mathcal{C}$



PTE(V)

Let $r_0 \sim U[\frac{1}{2}, 1)$

Let π be a uniformly random permutation of V

Return PTE$(V, \pi, r_0, \log \Delta)$

Note: satisfies structure by construction

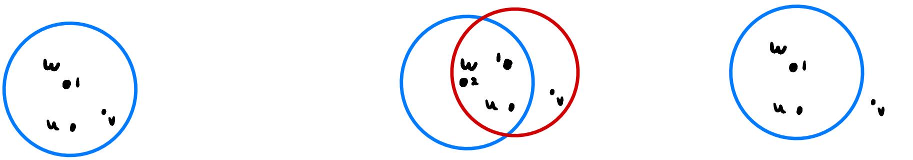
Fix $u, v \in V$

Say w <u>settles $\{u,v\}$ on level i</u> if it is the first vertex of π

w/ $|B(w,r_i) \cap \{u,v\}| \neq \emptyset$ → let $S_{i,w}$ be event

Say w <u>cuts $\{u,v\}$ on level i</u> if $|B(w,r_i) \cap \{u,v\}| = 1$ → let $C_{i,w}$ be event

E.g.



Settles, doesn't cut   Cuts, doesn't settle   Settles and cuts

Observe: lca at level i iff i is max i s.t. $\exists w$ s.t. $C_{i,w} \wedge S_{i,w}$

Order vertices by $d(w,\{u,v\}) := \min(d(w,u), d(w,v))$ and let X_w be position of w in order

<u>Claim 1:</u> $\Pr(S_{i,w} \mid C_{i,w}) \leq \frac{1}{X_w}$ $\forall i, w$

<u>Claim 2:</u> $\sum_i 2^{i+2} \cdot \Pr(C_{i,w}) \leq O(d(u,v))$ $\forall w \in V$

<u>Proof of (2) using Claims</u>

The lca of u,v is at level i iff i is the max i s.t.

$\exists w$ s.t. $C_{i,w} \wedge S_{i,w}$ (by construction)

lca at level i → $d_T(u,v) \leq 2^{i+2}$ (by claim) $\max \leq \sum_i$ and $\mathbb{1}(\exists w \ldots) \leq \sum_w \mathbb{1}(\ldots)$

Thus, $d_T(u,v) = \max_i 2^{i+2} \cdot \mathbb{1}(\exists w : C_{i,w} \wedge S_{i,w}) \leq \sum_i 2^{i+2} \cdot \sum_w \mathbb{1}(C_{i,w} \wedge S_{i,w})$

Taking $\mathbb{E}$, $\mathbb{E}[d_T(u,v)] \leq \sum_i 2^{i+2} \sum_w \Pr(C_{i,w} \wedge S_{i,w}) = \sum_w \sum_i 2^{i+2} \cdot \Pr(C_{i,w}) \cdot \Pr(S_{i,w} \mid C_{i,w})$

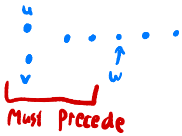
$$\leq \sum_w \frac{1}{X_w} \sum_i 2^{i+2} \cdot \Pr(C_{i,w}) \quad \text{(Claim 1)}$$

$$\leq O(d(u,v)) \sum_w \frac{1}{X_w} \quad \text{(Claim 2)}$$

$$= O(\log n) \cdot d(u,v) \quad \left(\sum_{i=1}^{n} \frac{1}{i} = O(\log n)\right)$$

Proof of Claim 1

Consider w ordered by $d(w, \{u,v\})$



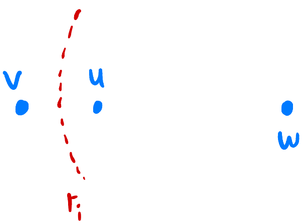
w only settles $\{u,v\}$ if it precedes all w' s.t. $X_{w'} < X_w$

Since π is a uniformly random permutation, this happens w/ $Pr \le \frac{1}{X_w}$

Proof of Claim 2

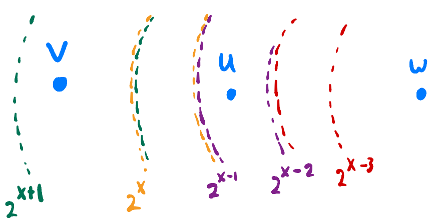
WLOG suppose $d(u,w) \le d(v,w)$

w cuts $\{u,v\}$ at level i only if $u \in B(w, r_i)$, $v \notin B(w, r_i)$



This happens only if $d(u,w) \le r_i \le d(v,w)$

But $r_i \in [2^{i-1}, 2^i)$ so $Pr(C_{iw}) \ne 0$ only if $[2^{i-1}, 2^i) \cap [d(u,w), d(v,w)] \ne \emptyset$



This is only true if 2^{i-1} or $2^i \in [d(w,u), d(w,v)]$

$\Updownarrow$

$i-1$ or $i \in [\log d(w,u), \log d(w,v)]$

Thus $\sum_i 2^{i+2} Pr(C_{iw}) = O(1) \cdot \sum_{i=\log d(w,u)}^{\log d(w,v)} 2^i = O\left(\sum_{i=0}^{\log d(w,v)} 2^i - \sum_{i=0}^{\log d(w,u)} 2^i\right) = O(d(w,v) - d(w,u))$

triangle inequality

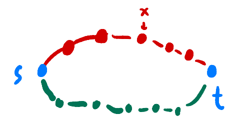
$\le O(d(u,v) + d(u,\cancel{w}) - d(\cancel{w},u))$

Embedding a Cycle into a Tree Requires $\Omega(n)$ Distortion

Claim: If $G=(V,E)$ is an edge-weighted cycle then embedding (V,d_G) into (V,d_P) for $P=(V,E,w)$ a path

Let $s,t \in V$ be an arbitrary pair s.t. $d_G(s,t) = \frac{n-1}{2} \approx \frac{n}{2}$

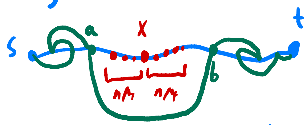
Color the vertices + edges in G between s,t red + green by path



Let x be the red vertex s.t. $d_G(x,s)=d_G(x,t) \approx \frac{n}{4}$

Notice $d_G(x,y) \geq \frac{n}{4}$ $\forall$ green y so by non-contractingness of embedding

every vertex within $\frac{n}{4}$ of x in P is red



But then $\exists$ a green edge $\{a,b\}$ as above $\rightarrow$ distortion $\Omega(n)$

Fact [Gupta]: if (V,d) is a tree metric, then $\forall U \subseteq V$ (U,d) embeds into a tree metric w/ distortion ≤ 8

Claim: If $G=(V,E)$ is an edge-weighted cycle then embedding (V,d_G) into a tree metric (V',d_T) requires distortion $\Omega(n)$

AFSOC $\exists$ embedding of (V,d_G) into (V',d_T) for T a tree w/ distortion $o(n)$

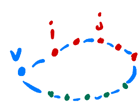
By fact, $\exists$ embedding of (V,d_G) into (V,d_T) w/ distortion $8 \cdot o(n) = o(n)$

Let T be tree w/ min edge weight Σ achieving this distortion

T must be a path

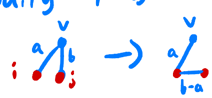
AFSOC $\exists$ v s.t. $\deg_T(v) \geq 3$

Color vertices of G red + green as



WLOG v has ≥ 2 red neighbors i,j w/ i closer to v in G

Modify T as



— This leaves shortest paths intact

— But reduces Σ edge weights so T a path

But then (V,d_G) was embedded w/ distortion $o(n)$ into (V,d_P) for $P=T$ a path, >< Claim