Lecture 9

Lecturers: Yair Bartal, Nova Fandina *Fall 2018/19*
For comments contact fandina@gmail.com

9.1 Prioritized and Coarse Scaling Distortions

In this section we present more measures of quality of embedding that essentially provide a stronger guarantee than the scaling distortion. We start with the prioritized distortion, that was defined and studied in [4].

Definition 9.1 (Prioritized Distortion)**.** *Let (X, d_X) be any n-point metric space, and let $\pi = (x_1, x_2, \ldots, x_n)$ be a priority ranking (an ordering) of the points of X. We say that an embedding $f : X \to Y$ has prioritized distortion α with respect to π, for a non-decreasing function $\alpha : \mathbb{N} \to \mathbb{R}^+$, if for all $1 \leq j < i \leq n$, the pair (x_j, x_i) has distortion $\alpha(j)$.*

In [2] it was shown[1] (among the other results) that the classical Bourgain result has the following prioritized counterpart.

Theorem 9.1. *For every $1 \leq p \leq \infty$, and every finite metric space (X, d_X), and any priority ranking of X, there is an embedding with prioritized distortion $O(\log j)$ into $\ell_p^{O(\log |X|)}$.*

It turns out that prioritized distortion is essentially equivalent to a strong version of scaling distortion, called *coarse scaling distortion*, in which for every point, the $(1 - \epsilon)$ fraction of the farthest points from it are preserved with the desired distortion.

Definition 9.2 (Coarse Scaling Distortion)**.** *Let $f : X \to Y$ be any embedding. For $v \in X$ and $\epsilon \in (0, 1)$, let $R(v, \epsilon) = \min\{r : |B(v, r)| \geq \epsilon n\}$. A point u is called ϵ-far from v if $d(u, v) \geq R(v, \epsilon)$. Given a function $\gamma : (0, 1) \to \mathbb{R}^+$, we say that f has a coarse-scaling distortion γ, if for every $\epsilon \in (0, 1)$, every pair $v \neq u \in X$ such that u, v are $\epsilon/2$-far from each other, has distortion at most $\gamma(\epsilon)$.*

Note that this definition indeed implies scaling distortion. Next we show the relation of prioritized and coarse-scaling distortion, which was proven in [2]:

Theorem 9.2. *Let $\mu : \mathbb{N} \to \mathbb{R}^+$ be a non-increasing function [2], s. t. $\sum_{j \geq 1} \mu(j) = 1$. Let $\mathbb{Y}$ be a subset closed family of finite metric spaces, and assume that for every finite metric space (Z, d_Z) there is a non-contractive embedding $f_Z : Z \to Y_Z$, where $(Y_Z, d_{Y_Z}) \in \mathbb{Y}$, with coarse-scaling distortion γ. Then, given a finite metric space (X, d_X) and a priority ranking $x_1, \ldots x_n$ of the points of X, there exists an embedding $f : X \to Y$, for some $(Y, d_Y) \in Y$, with prioritized distortion $\gamma(\mu(j))$.*

[1] The proof follows from the existence of $O(\log(1/\epsilon))$ - coarse-scaling Bourgain embedding, which was proven to exist in [1] (we will see this proof later in the course), and by Theorem 9.2

[2] For example, $\mu(j) \sim 1/(j \log^2 j)$, or $\mu(j) \sim 1/(j \log j \log^2 \log j)$.

Proof. Given the metric space (X, d_X) and a priority ranking $x_1, \ldots, x_n$, let $\delta = \min d_X(x_i, x_j)/2$. We define a new metric space (Z, d_Z) as follows. For every $1 \leq j \leq n$, every point x_j is replaced by a set X_j, of artificial points, of size $|X_j| = \lceil \mu(j)n \rceil$. Then, Z is the union of all X_j. The distances are defined by: for every $u \in X_j$ and $v \in X_l$, $d_Z(u, v) = d_X(x_j, x_l)$, if $j \neq l$, and $d_Z(u, v) = \delta$, otherwise. Note that $|Z| \leq \sum_{j=1}^{n} (\mu(j)n + 1) \leq 2n$.

By the assumption, there is an embedding $f_Z : Z \to Y_Z$ with coarse-scaling distortion γ. We use it to define $f : X \to Y_Z$ as follows: for every j, $f(x_j) = f_Z(u_j)$, for an arbitrary point $u_j \in X_j$. Note that $f(X) \subseteq Y_Z$, and thus $f(X) \in \mathbb{Y}$.

By the construction of Z, for every $i > j$, it holds that $X_j \subseteq B(u_j, d_Z(u_i, u_j))$, and also $X_j \subseteq B(u_i, d_Z(u_i, u_j))$, for any $u_j \in X_j$ and $u_i \in X_i$. Since $|X_j| \geq \mu(j)n \geq \frac{\mu(j)}{2}|Z|$, it holds that u_i, u_j are $\mu(j)/2$-far from each other. Therefore, $\frac{d_{Y_Z}(f(x_i), f(x_j))}{d_X(x_i, x_j)} = \frac{d_{Y_Z}(f_Z(u_i), f_Z(u_j))}{d_Z(u_i, u_j)} \leq \gamma(\mu(j))$. $\square$

For the opposite direction we have the following theorem [Exercise].

Theorem 9.3. *Let (X, d_X), (Y, d_Y) be metric spaces, then there exists a priority ranking $x_1, \ldots, x_n$ of the points of X such that the following holds: If there exists a non-contractive embedding $f : X \to Y$ with (monotone non-decreasing) prioritized distortion α, then f has coarse scaling distortion $O(\alpha(8/\epsilon))$.*

9.2 Ramsey Type Embeddings

The concept of modern Ramsey theory states that large enough systems necessarily contain highly structured sub-systems. The most famous example of this principle, perhaps is the following proposition: in every group of six people, either there are three of them that know each other, or there are three of them that do not know each other. In the metric spaces context, this phenomena can be translated to the following question: is it true that for any n-point metric space there is a large sub-space which embeds in some nice metric space (Euclidian space, trees, ultra-metrics), with small distortion?

Definition 9.3. *We say that an embedding $f : X \to Y$ is a Ramsey embedding with distortion α, if (f, G) is a partial embedding with distortion α, where $G = \binom{Z}{2}$ for some $Z \subseteq X$. We say that Z is a core of the embedding.*

Definition 9.4. *We say that Ramsey embedding is strong if $(f, Z \times X)$ is a partial embedding.*

The main result of this field is the following theorem.

Theorem 9.4 ([3, 5])**.** *For any metric on n points, for any $0 < \epsilon < 1$, there is a strong Ramsey embedding of X into an ultra-metric, with core of size at least $n^{1-\epsilon}$, and with distortion $O(\frac{1}{\epsilon})$.*

The theorem was proven in [5], where the authors presented a randomized construction of the embedding. In what follows, we present a deterministic embedding, that was constructed by Bartal.

The above theorem has an application to the Distance Oracle problem, where one is asked to build a compact data structure for a finite metric space, to answer (approximate) distance queries.

Claim 9.5. *Given Theorem 9.4, for any metric space X on n points there exists a (deterministic) data structure for the Distance Oracle problem for X with the following parameters:*

Space Complexity - $O(n^{1+\epsilon})$;

Query Time - $O(1)$ (in previous results, query time is a function of ϵ);

Approximation Guarantee - $O(\frac{1}{\epsilon})$.

We note that thisd ata structure has the best possible tradeoff (asymptotically).

Proof. Let $0 < \epsilon < 1$, we construct the data structure as follows:

Let $X_0 = X$. By Theorem 9.4 there is a strong Ramsey embedding $f_0 : X_0 \to Y_0$, with distortion $O(\frac{1}{\epsilon})$. Denote Z_0 its core, then $|Z_0| \geq n^{1-\epsilon}$. Let $X_1 = X_0 \setminus Z_0$. By the theorem there is a strong Ramsey embedding $f_1 : X_1 \to Y_1$, with distortion $O(\frac{1}{\epsilon})$. Denote Z_1 its core, then $|Z_1| \geq |X_1|^{1-\epsilon}$. Generally, $\forall i \geq 2$, let $X_i = X_{i-1} \setminus Z_{i-1}$, and $f_i : X_i \to Y_i$ is a strong Ramsey embedding with distortion $O(\frac{1}{\epsilon})$, and $|Z_i| \geq |X_i|^{1-\epsilon}$ its core. Note that $\{Z_i\}_{i=1}^{\infty}$ is a partition of X.

The size of the constructed DS. $\forall i \geq 0$, $|Y_i| = |X_i|$ (where $|Y_i|$ denotes the number of the leaves of the ultra-metric, which essentially equals to the size of the whole tree up to a constant factor). Therefore, the size of thedata structure is $\sum_{i=0}^{\infty} |Y_i| = \sum_{i=0}^{\infty} |X_i|$ (the number of steps in the construction is finite of course, but let us be inaccurate here for the sake of simplicity of presentation). Let $\beta = n^{-\epsilon}$. Thus

$|X_1| \leq |X| - |X|^{1-\epsilon} = |X| - \beta|X| = (1-\beta)|X|$,

$|X_2| \leq |X_1| - |X_1|^{1-\epsilon} = |X_1|(1 - |X_1|^{-\epsilon}) \leq |X_1|(1-\beta) \leq (1-\beta)^2|X|$.

Therefore, $|X_i| \leq (1-\beta)^i|X|$. Thus, $\sum_{i=0}^{\infty} |Y_i| = \sum_{i=0}^{\infty} |X_i| \leq |X| \sum_{i=0}^{\infty} (1-\beta)^i = nn^{\epsilon}$.

Query Algorithm. Given any $u, v \in X$, let Y_i, Y_j be the trees such that $u \in Z_i$ and $v \in Z_j$. W.l.o.g $i < j$. The algorithm finds an $lca(u,v)$ in the tree Y_i. The label of the lca is the $O(\frac{1}{\epsilon})$-approximated distance between u and v. For each $u \in X$ we store in a Hash Table the appropriate tree (during the construction of the data structure). We use an $O(1)$-time deterministic algorithm for answering lca queries ([6]) which has linear preprocessing time. □

Proof. of Theorem 9.4 We will prove a weaker version of the theorem. Namely we will prove that for any metric space X on n points, for any $0 < \epsilon < 1$ there exists a Ramsey embedding of X into an ultrametric with distortion $O(\frac{1}{\epsilon})$, and with core of size at least $n^{1-\epsilon}$.

We prove, by induction on n, that there is a Ramsey embedding with distortion $8/\epsilon$, and with core of size at least $n^{1-\epsilon}$. The idea is to partition $X = Q \dot{\bigcup} \bar{Q}$, and to define a set $P \subseteq X$ such that: $P \subseteq Q$ and $d(P, \bar{Q}) \geq \frac{8 \cdot diam(X)}{\epsilon}$.

Assume we have found such a partition of X. Then by induction's assumption, there is a Ramsey embedding $f_1 : \bar{Q} \to T_{\bar{Q}}$, with distortion $\frac{8}{\epsilon}$, and with the core $S(\bar{Q})$ of size $|S(\bar{Q})| \geq |\bar{Q}|^{1-\epsilon}$, and there is Ramsey embedding $f_2 : P \to T_P$, with distortion $\frac{8}{\epsilon}$, and with core $S(P)$ of size $|S(P)| \geq |P|^{1-\epsilon}$.

Then, we construct a Ramsey embedding $f : X \to T$ as follows. The root of T is a new vertex r, with $\Delta(r) = diam(X)$. Its children are $T_{\bar{Q}}$ and T_P. The points of $Q \setminus P$ are embedded to some arbitrary points. Note that the embedding we build is non-contractive.

Note that $S(\bar{Q}) \cup S(P) \subseteq S(X)$ (the core of f), since by the induction's assumption, it holds that for all $x, y \in S(\bar{Q})$ or $x, y \in S(P)$, the distortion of the pair is $\frac{8}{\epsilon}$. Also, for $x \in S(\bar{Q})$ and $y \in S(P)$, it holds that $d(x, y) \geq \frac{8 \cdot diam(X)}{\epsilon}$, and $d_T(x, y) = diam(X)$, implying $dist_f(x, y) \leq \frac{8}{\epsilon}$.

It remains to show how to partition X such that $|S(X)| \geq |X|^{1-\epsilon}$. Define $w : 2^X \to \mathbb{N}$ as follows: $\forall\ T \subseteq X,\ \ w(T) = |T|$. If there exists $w^* : 2^X \to \mathbb{N}$ such that $\forall\ T \subseteq X,\ \ w^*(T) \leq w(T)$ and w^* is monotonically increasing (in a sense of inclusions of sets), then it's enough to prove that $w(S(X)) \geq w(X) \cdot (w^*(X))^{-\epsilon}$.

Therefore, we strengthen the requirement on the size of core in the induction's assumption, using w^* function. Thus, using the stronger induction's assumption, we have:

$$w(S(X)) \geq w(S(P)) + w(S(\bar{Q})) \geq^{induction's\ assumption} \geq w(P) \cdot (w^*(P))^{-\epsilon} + w(\bar{Q}) \cdot (w^*(\bar{Q}))^{-\epsilon} \geq$$

$$\geq^{(w^*(\bar{Q}))^{-\epsilon} \geq (w^*(X))^{-\epsilon}} \geq w(P) \cdot (w^*(P))^{-\epsilon} + w(\bar{Q}) \cdot (w^*(X))^{-\epsilon} \geq^{when\ is?} \geq w(X) \cdot (w^*(X))^{-\epsilon}.$$

The last inequality holds $\Leftrightarrow w(P) \cdot (w^*(P))^{-\epsilon} \geq w(X) \cdot (w^*(X))^{-\epsilon} - w(\bar{Q})(w^*(X))^{-\epsilon}$

$\Leftrightarrow w(P) \cdot (w^*(P))^{-\epsilon} \geq w(Q) \cdot (w^*(X))^{-\epsilon} \Leftrightarrow \frac{w(P)}{w(Q)} \geq \left(\frac{(w^*(X))}{(w^*(P))}\right)^{-\epsilon}$.

Thus, we have to show that there exist $Q \subseteq X$, $P \subseteq Q$, and w^* such that the above holds.

Claim 9.6. *Let X be any n point metric space. Let $0 < \Delta \leq \frac{diam(X)}{2}$, and $0 < t \in \mathbb{N}$. Then there exists a partition of $X = (Q, \bar{Q})$ such that the following holds:*

1. *$diam(Q) \leq \Delta$;*
2. *There exists $P \subseteq Q$ such that $d(P, \bar{Q}) \geq \frac{\Delta}{4t}$;*
3. *For $w^* : 2^X \to \mathbb{N}$, $w^*(T) = \max_{z \in T} w(B_T(z, \frac{diam(T)}{4}))$, it holds that $\frac{w(P)}{w(Q)} \geq \left(\frac{w^*(X)}{w^*(P)}\right)^{-\frac{1}{t}}$.*

For our theorem we need the result of the claim for $\Delta = \frac{diam(X)}{2}$, and $t = \lceil 1/\epsilon \rceil$. Also note that w^* satisfies the properties stated in the claim.

Proof of Claim 9.6. We construct $Q = B_X(v, r)$ as follows. Find $v \in X$ such that the quantity $\frac{w(B_X(v, \frac{\Delta}{2}))}{w(B_X(v, \frac{\Delta}{4}))}$ is minimized. We can find such v in $O(n^2)$ time. To choose r we consider a sequence of the balls in X: $Q_i = B_X(v, \frac{\Delta}{4}((1 + \frac{i}{t}))$, $0 \leq i \leq t$. Choose i such that

$$\frac{w(Q_i)}{w(Q_{i-1})} \leq \left(\prod_{j=1}^{t} \frac{w(Q_j)}{w(Q_{j-1})}\right)^{\frac{1}{t}} = \left(\frac{w(Q_t)}{w(Q_0)}\right)^{\frac{1}{t}} = \left(\frac{w(B((v, \frac{\Delta}{2})))}{w(B(v, \frac{\Delta}{4}))}\right)^{\frac{1}{t}}.$$

Note that such i must exist, as otherwise we will obtain that $\prod_{j=1}^{t} \left(\frac{w(Q_j)}{w(Q_{j-1})}\right) > \prod_{j=1}^{t} \left(\frac{w(Q_j)}{w(Q_{j-1})}\right)$. If there are several appropriate i's, choose the minimal one.

We define $Q = Q_i$, $P = Q_{i-1}$. By the definition of Q_i's it holds that $d(P, \bar{Q}) \geq \frac{\Delta}{4t}$.

In addition, $\frac{w(P)}{w(Q)} \geq \left(\frac{w(B(v,\frac{\Delta}{2}))}{w(B(v,\frac{\Delta}{4}))}\right)^{-\frac{1}{t}} \geq^{\textit{have to show}} \geq \left(\frac{w^*(X)}{w^*(P)}\right)^{-\frac{1}{t}}$, which is true iff $\frac{w^*(X)}{w^*(P)} \geq \frac{w(B(v,\frac{\Delta}{2}))}{w(B(v,\frac{\Delta}{4}))}$.

By the definition of w^*, it holds that $w(B(v, \frac{\Delta}{2})) \leq w^*(X)$, since $\frac{\Delta}{2} \leq diam(X)/4$. Note that this holds for any $v \in X$. Let $u \in Q$ such that $w(B_Q(u, \frac{diam(Q)}{4}))$ is maximal. Therefore

$$w^*(P) \leq^{w^* \textit{ is monotonically increasing}} \leq w^*(Q) =^{\textit{by the choice of } u} = w(B_Q(u, \frac{diam(Q)}{4})) \leq$$
$$\leq^{diam(Q) \leq \Delta} \leq w(B_Q(u, \frac{\Delta}{4})) \leq w(B_X(u, \frac{\Delta}{4})).$$

Therefore
$$\left(\frac{w^*(X)}{w^*(P)}\right) \geq \frac{w(B(u,\frac{\Delta}{2}))}{w(B(u,\frac{\Delta}{4}))} \geq^{\textit{by the choice of } v} \geq \frac{w(B(v,\frac{\Delta}{2}))}{w(B(v,\frac{\Delta}{4}))}.$$ □

This completes the proof of the theorem. □

References

[1] Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. STOC '06, pages 271–286, 2006.

[2] Yair Bartal, Arnold Filtser, and Ofer Neiman. On notions of distortion and an almost minimum spanning tree with constant average distortion. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 873–882, 2016.

[3] Yair Bartal, Nathan Linial, Manor Mendel, and Assaf Naor. On metric ramsey-type phenomena. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, pages 463–472, 2003.

[4] Michael Elkin, Arnold Filtser, and Ofer Neiman. Prioritized metric structures and embedding. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing*, STOC '15, pages 489–498, New York, NY, USA, 2015. ACM.

[5] Manor Mendel and Assaf Naor. Ramsey partitions and proximity data structures. *CoRR*, abs/cs/0511084, 2005.

[6] Baruch Schieber and Uzi Vishkin. On finding lowest common ancestors: Simplification and parallelization. *SIAM Journal on Computing*, 17(6):1253–1262, 1988.