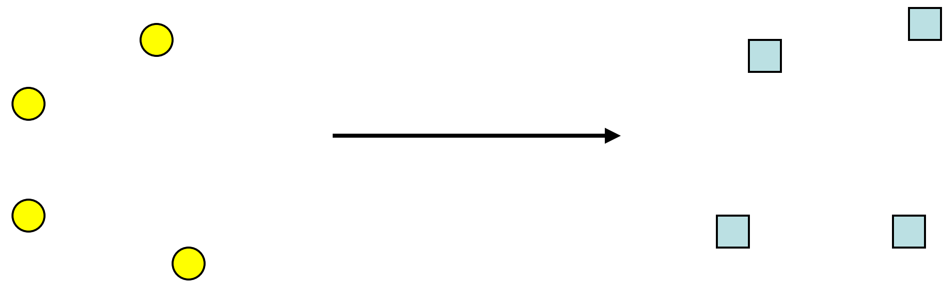
Approximation Algorithms for Embedding Problems

Piotr Indyk
MIT

Low-Distortion Embeddings

- Consider metrics (X,D_X) and (Y,D_Y)
- (X,D_X) c-embeds into (Y,D_Y) if there is a mapping $f:X \rightarrow Y$ such that, for all $p,q \in X$:

$$D_X(p,q) \leq D_Y(f(p),f(q)) \leq c\, D_X(p,q)$$



Examples of Embedding Results

- [Bourgain'85]: Any n-point metric can be embedded into d-dimensional Euclidean space with distortion $O(\log n)$
  - d can be made $O(\log^2 n)$
- [Johnson-Lindenstrauss'84]: Any n-point subset of a d-dimensional Euclidean space can be embedded into $O(\log n/\varepsilon^2)$ -dimensional Euclidean space with distortion $1+\varepsilon$

Embeddings I

- Absolute bounds: for a metric M and a class of metrics C, show that for every $M' \in C$, M' c-embeds into M
- Problem: absolute bounds very weak for embedding into, say, R^2
  - Example: uniform metric: $D(p,q)=1$ for $p \neq q$
  - Cannot be embedded into R^2 with distortion better than $\approx n^{1/2}$

    ($n^{1/2} \times n^{1/2}$ grid is near-optimal)

Embeddings II

- Relative bounds: give an algorithm that, given $M' \in C$ as an input:
  - if M' c-embeds into M,
  - then it finds an (a*c)-embedding of M' into M for some approximation factor $a>1$.
- MDS-style approach
- But, with guaranteed bounds

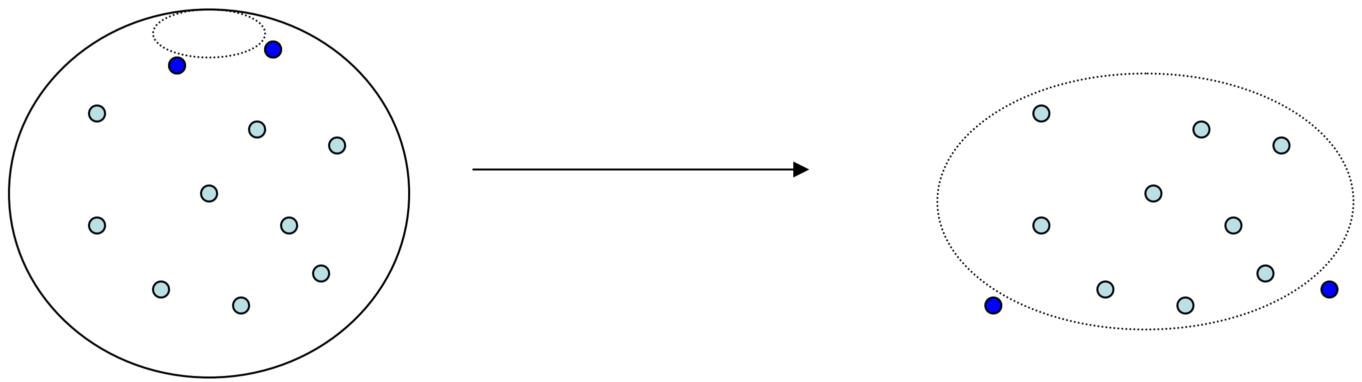
Results

Paper	From	Into	Distortion	Comments
[DGRR]+	unweighted graphs	line	$O(c^2)$	
[BIRS]=	unweighted graphs	line	>ac, a>1	Hardness
[BDGRRRS'05]	unweighted graphs	line	c	c constant
	unweighted trees	line	$O(c^{3/2} \log c)$	
	sphere	plane	3c	
[BIS'04]	unweighted graphs	trees	O(c)	
[BCIS'05]	general metrics	line	$\Delta^{3/4} c^{O(1)}$	Δ = spread
	weighted trees	line	$c^{O(1)}$	
	weighted trees	line	$\Omega(c\, n^{1/12})$	Hardness
[BCIS'06]	ultrametric	plane	$O(c^3)$	

Sphere $\to$ Plane

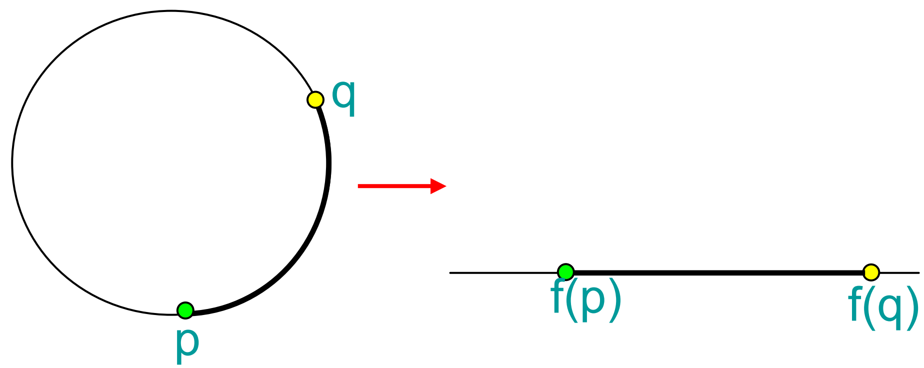
- Given $X \subseteq S^2$, $|X|=n$, approximate the min distortion of $f: X \to R^2$
- The distortion could be $\Omega(n^{1/2})$
  - Take X to be an $1/n^{1/2}$-net of S^2

    (each point in S^2 has a point in X within dist. $1/n^{1/2}$



Algorithm

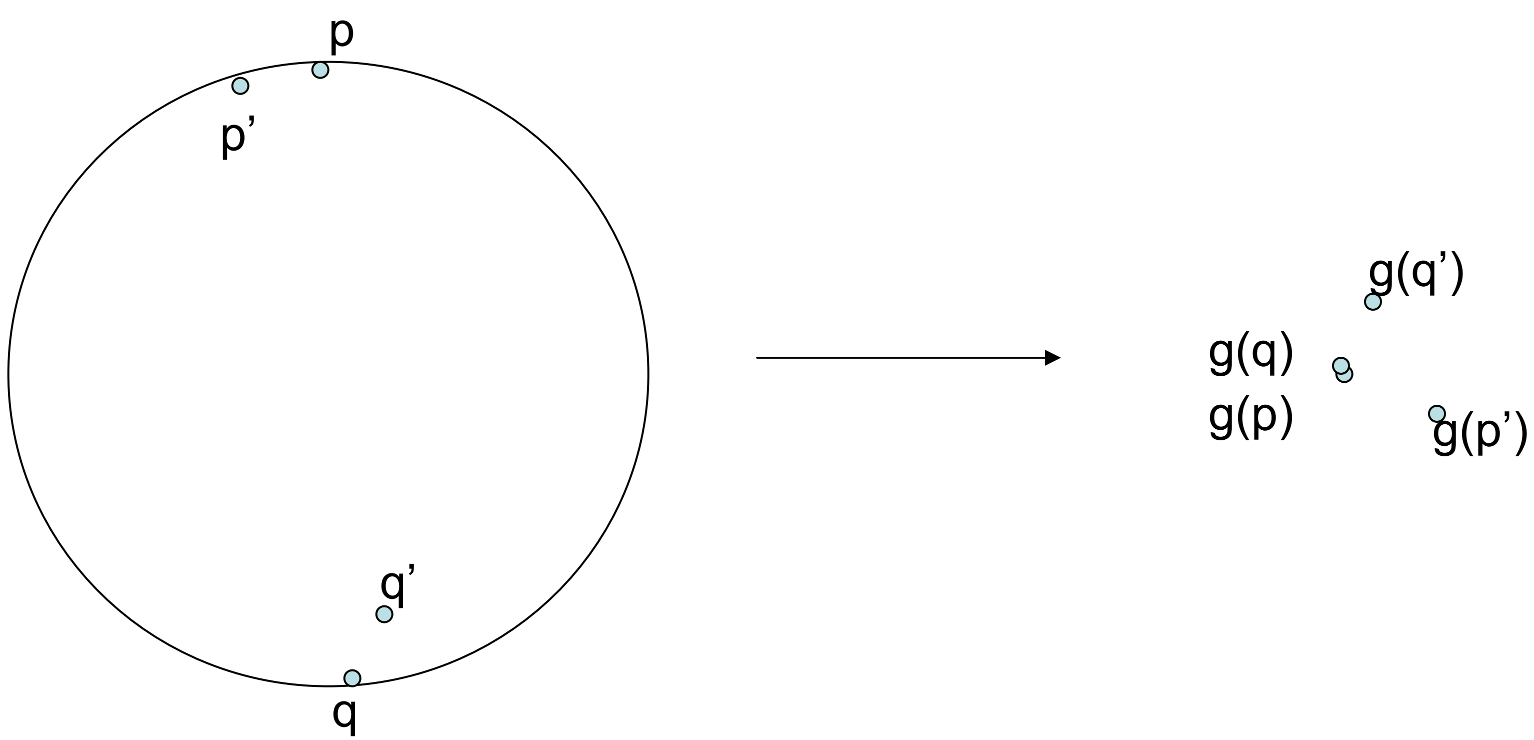
- Find largest empty cap B(p,r)
- Rotate the sphere to put p at the bottom
- Map sphere → plane:
  - "Cut the cap"
  - "Unwrap the sphere"
  - For each point q, the distance |f(p)-f(q)| equal to the geodesic distance from p to q
- Distortion: O(1/r)



Analysis – Lower bound

- The set X is an r-net of S^2
- Consider optimal $f: X \rightarrow R^2$, assume non-expansion
- Extend f to (non-expanding) $g: S^2 \rightarrow R^2$
- Borsuk-Ulam: there exist antipodal p,q for which $g(p)=g(q)$
- There exists $p',q' \in X$ with $|p'-p| \le r$,$|q-q'| \le r$

Lower bound ctd

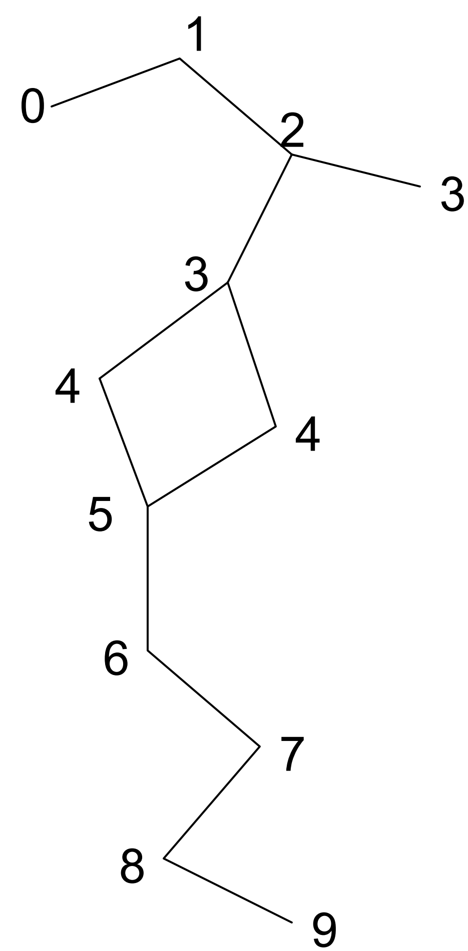


- Distortion is at least

  $||p'-q'|| \,/\, ||g(p')-g(q')|| \geq (2-2r)/2r = \Omega(1/r)$

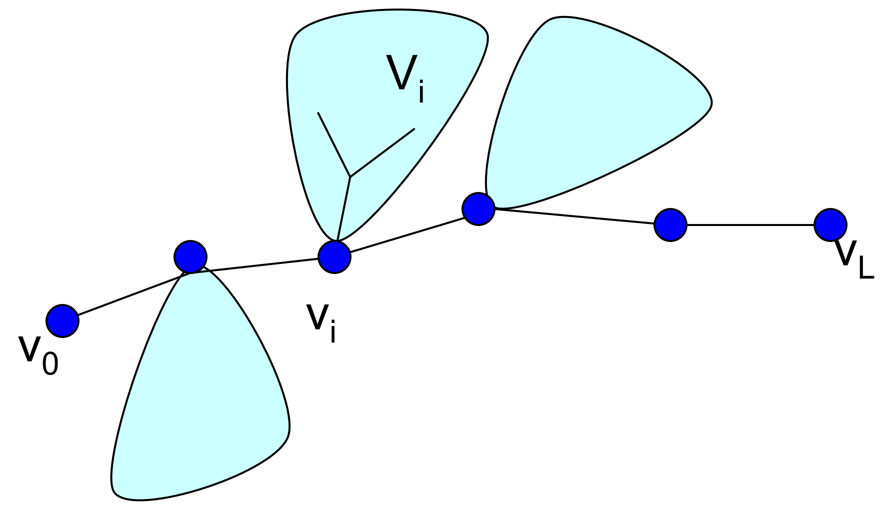
Unweighted graphs into a line

- Intuition:
  - Assume we want to embed an "almost line metric" induced by (V,E)
  - Metric should be "long and thin"
  - Distances from one endpoint should be a good approximation of the embedding



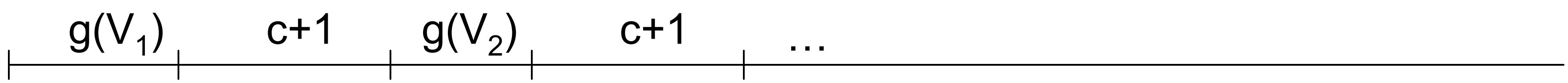
Algorithm

- Assume optimal embedding $f:V \rightarrow R$
- Guess:
  - v_0= leftmost node in $f(V)$
  - v_L = rightmost node in $f(V)$
- Compute the shortest path $P=v_0, v_1, \ldots v_L$ from v_0 to v_L
- $V_i = \{v \in V: D(v,v_i)=D(v,P)\}$



Algorithm ctd.

- Compute g:

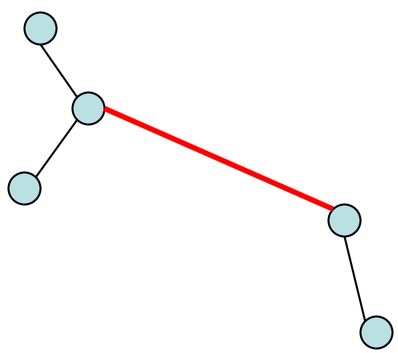


- Can prove each $|V_i| = O(c^2)$
- Each $g(V_i)$ has diameter and distortion $O(|V_i|)$ …

MST Embedding

- … because one always get distortion of $O(n)$ [Mat'90]:
  - Compute an MST T of the metric $M=(X,D)$
  - Split T into T_1, T_2 by removing longest edge e
  - Construct g:

    $g(T_1)$, $length(e)$, $g(T_2)$
  - Distortion:
    - $cost(T_1), cost(T_2) \leq n\ length(e)$
    - $length(g(T)) = O(cost(T))$
    - For $p \in T_1, q \in T_2$, distortion of $D(p,q)$ is $\leq length(g(T))/length(e) = O(n)$



Conclusions

- Approximation algorithms for min distortion embedding
- Guarantees somewhat limited, but provable
- For more info, see

http://publications.csail.mit.edu/abstracts/abstracts05/low/low.html