Local Global Tradeoffs in Metric Embeddings

Moses Charikar

Princeton University

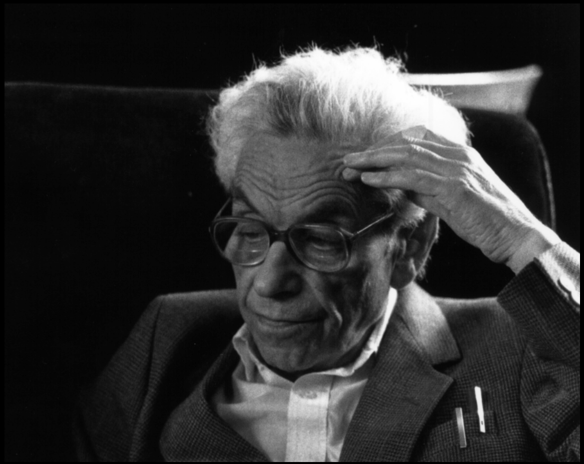
Local vs Global Question

- Local: property of subsets
- Global: property of entire set

- Does local structure imply global structure?

Erdős on Local Global

- Chromatic number not local concept
- Graphs with high chromatic number, but small subgraphs suggest otherwise



© George Paul Csicsery

1959 — **GRAPH THEORY AND PROBABILITY**
P. ERDÖS

1962 — **ON CIRCUITS AND SUBGRAPHS OF CHROMATIC GRAPHS**
P. ERDÖS

Erdős on Local Global

- Graphs with high girth and high chromatic number

A combinatorial classic – sparse graphs with high chromatic number, Jaroslav Nešetřil

- g: girth, k: chromatic number

$$\frac{\log n}{2\log k} \leq g \leq \frac{2\log n}{\log(k-2)} + 1$$

Question: Does $\lim_{n\to\infty} \frac{g_k(n)}{\log n}$ exist ?

Erdős on Local Global

- For all k, k-chromatic graph, such that $\Omega(n)$ size subsets are 3-colorable

- What if subsets are 2-colorable?

- Conjecture: exists c_k such that

  no odd cycle of length <= $c_k n^{1/k}$ ⇒ graph is k+1 colorable

- [Kierstead, Szemeredi, Trotter '84]

Local Global Questions in Other Settings

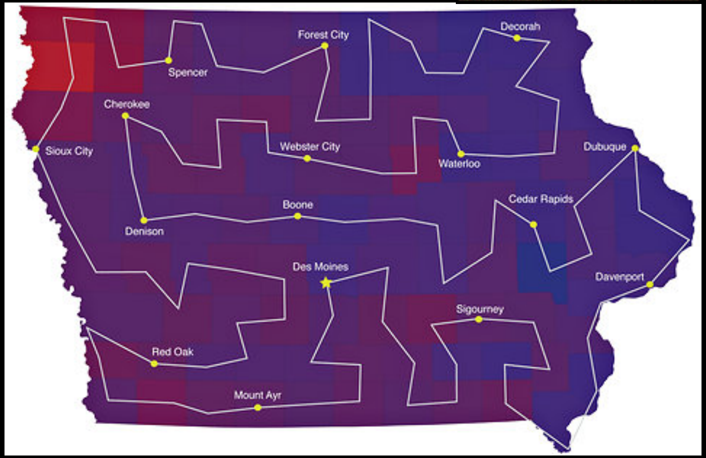
- Embedding into l_2^d
- Characterization of tree metrics
- Helly's theorem
- Ramsey theory
- Graph minors work
  – minor exclusion is local property, what does it mean for entire graph ?
- Property testing
  – infer properties of entire set from sample

Local Global Tradeoffs in Metric Embeddings

A Brief History of Optimization

- **NP-completeness** (early 70's): Many optimization problems hard to solve exactly

Traveling Salesman



- approximately optimal solution?

- **PCP theorem** (early 90's): optimization problems cannot be approximated beyond threshold

- threshold of approximability?

Approximation Algorithms

- Hard optimization problems (min/max)

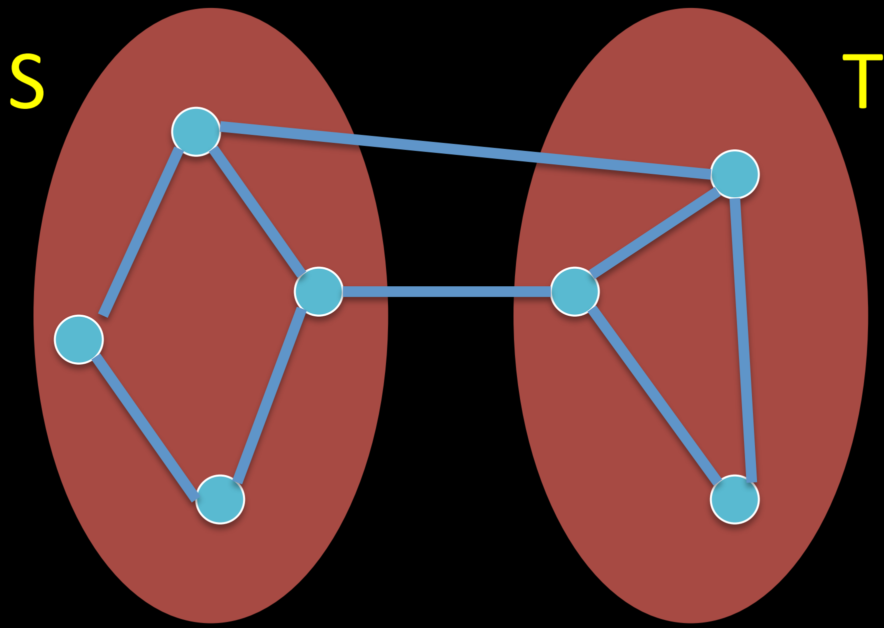
Relax and round paradigm

- "Relax" to obtain tractable problem
- "Round" solution to relaxation to solve original

- How good is the fractional solution ?

$$\text{integrality gap} = \max \frac{\text{value(integer soln)}}{\text{value(fractional soln)}}$$

Sparsest Cut



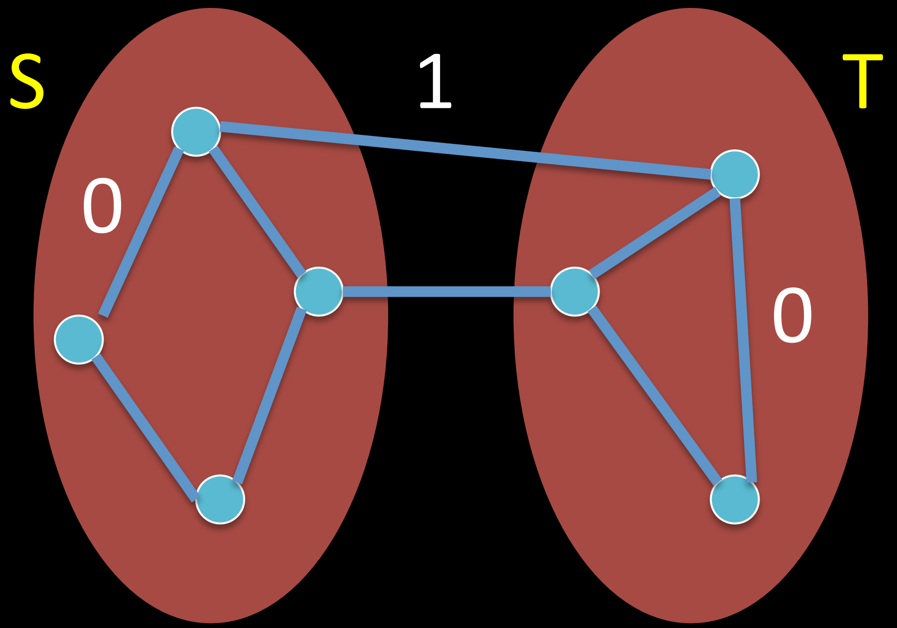
$$\min \frac{|E(S,T)|}{|S|\,|T|}$$

Generalized Sparsest Cut

Weight functions
α, β on edges

$$\min \frac{\alpha(S,T)}{\beta(S,T)}$$

Cut (semi)Metrics

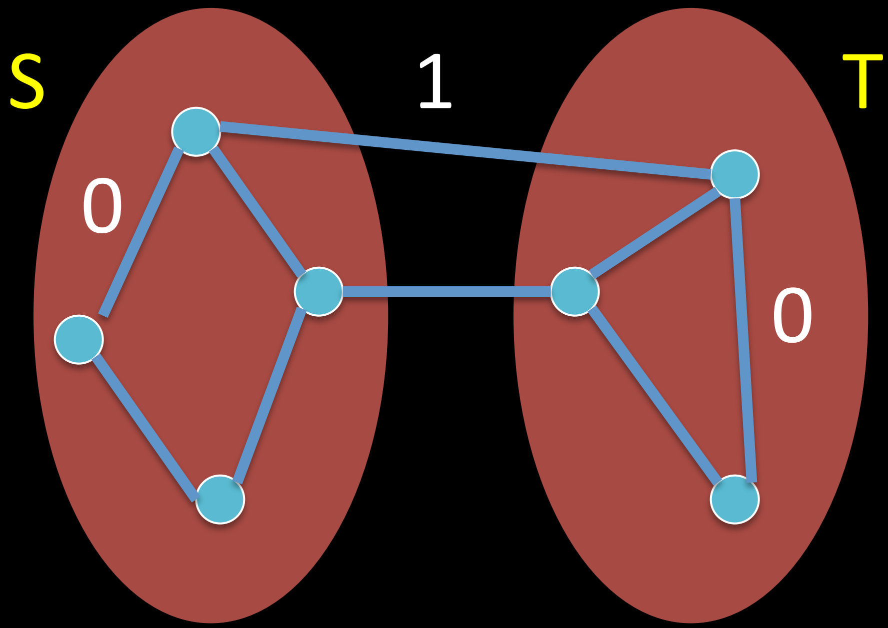


$$\min \frac{\Sigma\alpha(u,v)d(u,v)}{\Sigma\beta(u,v)d(u,v)}$$

d(.,.) is {0,1} metric

- Exact reformulation of Sparsest Cut

Relax to general metric



Linear Programming

$$\min \frac{\Sigma \alpha(u,v)d(u,v)}{\Sigma \beta(u,v)d(u,v)}$$

d(.,.) is ~~{0,1}~~ metric

- Metric can be mapped to distribution over cut metrics with log(n) distortion [Bourgain '85]
- log(n) approximation for Sparsest Cut [Linial, London, Rabinovich '94] [Aumann, Rabani '94]

Metric Embeddings

[Bourgain '85]

Metric *d* on n points can be mapped to ℓ_1

$$d(x,y) \leq ||f(x) - f(y)||_1 \leq O(\log n) d(x,y)$$

- Integrality gap of Sparsest Cut relaxation is exactly the distortion of embedding into ℓ_1

- Better approximation by constraining general metrics further?
- YES: $\sqrt{\log n}$ approximation via semidefinite programming

Lift and Project Hierarchies

- Systematic, iterative procedures to strengthen mathematical programming relaxations
- [Lovász, Schrijver ‘91]
  [Sherali Adams, ‘90]
  [Lasserre, 01]

- In k rounds, enforce constraints on all subsets of size k
  - solution in time $n^{O(k)}$
- Does local structure imply global structure?

Local Global Tradeoffs for Metrics

- [Arora, Bollobás, Lovász, Newman, Rabani, Rabinovich, Vempala, '06]
- Suppose every subset of k points in metric space embeddable into ℓ_1 with distortion D.
- Min distortion for embedding entire space into ℓ_1?

- Lower bound: $(\log n)^{\Omega(1/k)}$ for D=1
- Upper bound: $O(D\,(n/k)^2)$

Local Global Tradeoffs for Metrics

- [C, Makarychev, Makarychev '07]
- Lower Bounds

$D = 1$ $\Omega\left(\frac{\log n}{\log k + \log \log n}\right)$

$D = 1 + \delta$ $\Omega\left(\frac{\log(n/k)}{\log(1/\delta)}\right)$

$D \geq 3/2$ $\Omega(D \log(n/k))$

- Upper Bound $O(D \log(n/k))$

Large distortion even if subsets of size $n^{o(1)}$ embed isometrically

Lower bound: Roadmap

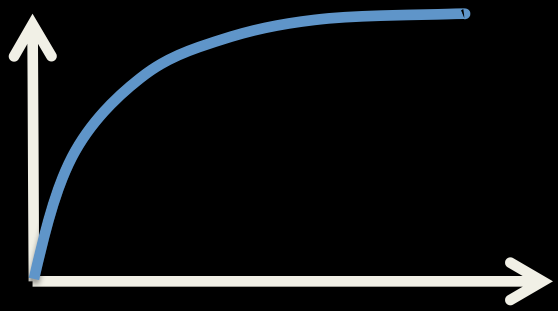
- Constant degree expander
- High global distortion

- Subgraphs of expander are sparse
- Sparse graphs embed well

New metric

- Expander with new metric

$$\rho(u,v) = 1 - (1 - \mu)^{d(u,v)}$$



- Every embedding of (G, ρ) into ℓ_1 requires distortion

$$\Omega\left(\frac{\log(n/k)}{\log(1/\delta)}\right)$$

- Every subset of X of size k embeds into ℓ_1 with distortion $1+\delta$

Global distortion

- 3-regular expander G, girth $\Omega(\log n)$
- [LLR] Min distortion for embedding it into ℓ_1 is

  Ω(avg distance / length of edge)

- New metric $\rho(u,v) = 1 - (1 - \mu)^{d(u,v)}$

- Distortion $$\frac{\Omega(1)}{1-(1-\mu)} = \Omega\left(\frac{\log(n/k)}{\log(1/\delta)}\right)$$

Embedding Subgraphs

- Trees embed isometrically into ℓ_1
- Embedding easy if subgraphs are acyclic
  - Too strong: subgraph size bounded by girth
- Exploit sparsity of subgraphs

Multicuts

- Construct a distribution on partitions
  Goal: $\Pr(u,v \text{ separated}) \approx \rho(u,v) = 1 - (1 - \mu)^{d(u,v)}$
- High level idea: remove every edge with probability μ
- The shortest path between u and v survives with probability $(1 - \mu)^{d(u,v)}$
- If the shortest path was the only path between u and v, then u,v separated with prob.

$$\rho(u,v) = 1 - (1 - \mu)^{d(u,v)}$$

l - path decomposable expanders

- **H is l - path decomposable** if every 2-connected subgraph contains a path (each vertex has degree 2) of length l

- [Arora, Bollobás, Lovász, Tourlakis ,'06] 3-regular expander G, girth $\Omega(\log n)$, every subgraph H of size at most k is $\Omega(\log(n/k))$ - path decomposable

- Probabilistic method:
  Expanders with sparse subgraphs
  Sparsity + girth => path decomposable

Multicuts

- H is l - path decomposable, L = l/9, $\mu \leq 1/L$

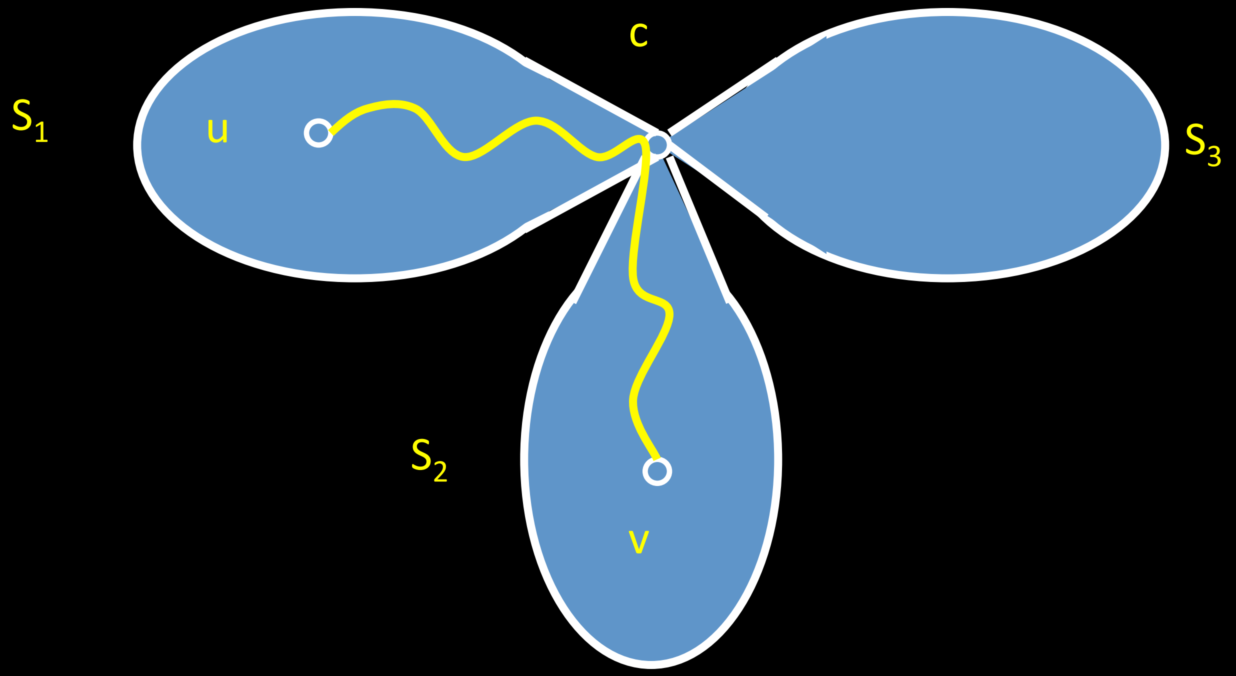
- Distribution on multicuts:
  - d(u,v) ≤ L, Pr(u,v separated) = $1 - (1-\mu)^{d(u,v)}$
  - d(u,v) > L, Pr(u,v separated) ≥ $1 - (1-\mu)^{L}$

- Distortion

$$\frac{1}{1-(1-\mu)^{L}} = 1 + O(e^{-\mu L})$$

Distribution on multicuts

- H has cut vertex c



- Sample multicuts independently in S_i

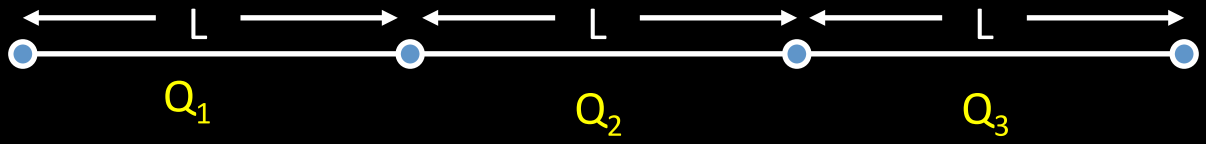
Pr[u,v not separated] = Pr[u,c not separ] * Pr[v,c not separ]

$$= (1-\mu)^{d(u,c)} (1-\mu)^{d(v,c)} = (1-\mu)^{d(u,v)}$$

Long paths

- $d(u,v) \leq L$, $\Pr(u,v \text{ separated}) = 1-(1-\mu)^{d(u,v)}$
- $d(u,v) > L$, $\Pr(u,v \text{ separated}) \geq 1-(1-\mu)^{L}$
- The end points are always separated!

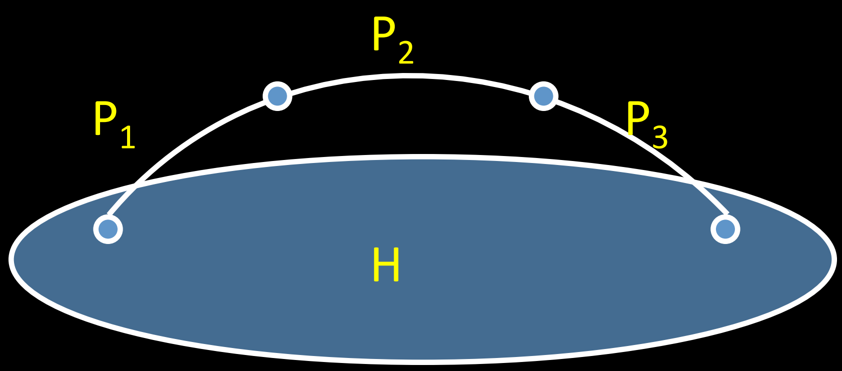
- Can be done for path of length 3L



- Cut edges "independently" with probability μ
- Decisions for Q_1 and Q_3 not independent

Distribution on multicuts

- H has a path of length $l = 9L$



- Divide path into 3 parts P_1, P_2, P_3
- Sample multicuts independently in H, P_1, P_2, P_3
- Computation same as before

Isometric local embeddings

- Every subset of size k embeds isometrically into ℓ_1
- Entire metric requires distortion

$$\Omega\left(\frac{\log n}{\log k + \log\log n}\right)$$

- Main idea:
  - make distortion very close to 1; δ = c/k.log(n)
  - add a *uniform* metric: ρ'(u,v) = ρ(u,v) + μ
  - Near isometric embedding can be corrected to embed new distance exactly

Applications

- **Sherali-Adams Hierarchy:** Integrality gap for many problems after n^{δ} rounds.

- Construct local distributions on solutions
- **Key challenge**: distributions for subsets must be consistent on intersection

- (Modular) Solution: Use lower bound construction with sqrt(distance): embeddable into ℓ_2
  - ℓ_2 embedding uniquely defined by pairwise distances
  - distribution on solutions from ℓ_2 embedding

Conclusion and Questions

- Every subset of size k isometrically embeddable into l_1 : global distortion ?

$$O(\log(n/k)) \text{ versus } \Omega\left(\frac{\log n}{\log k + \log\log n}\right)$$

- **Negative type metrics: d(u,v) = $||x_u - x_v||_2^2$**
  [Khot, Saket '09] [Raghavendra, Steurer '09]

1. isometric embedding of $2^{O(\log\log n)^{\delta}}$ size sets
2. Global distortion $\Omega(\log\log n)^{\gamma}$