8 LOW-DISTORTION EMBEDDINGS OF FINITE METRIC SPACES

Piotr Indyk and Jiří Matoušek

INTRODUCTION

An n-point metric space (X, D) can be represented by an $n \times n$ table specifying the distances. Such tables arise in many diverse areas. For example, consider the following scenario in microbiology: X is a collection of bacterial strains, and for every two strains, one is given their *dissimilarity* (computed, say, by comparing their DNA). It is difficult to see any structure in a large table of numbers, and so we would like to represent a given metric space in a more comprehensible way.

For example, it would be very nice if we could assign to each $x \in X$ a point $f(x)$ in the plane in such a way that $D(x, y)$ equals the Euclidean distance of $f(x)$ and $f(y)$. Such a representation would allow us to see the structure of the metric space: tight clusters, isolated points, and so on. Another advantage would be that the metric would now be represented by only $2n$ real numbers, the coordinates of the n points in the plane, instead of $\binom{n}{2}$ numbers as before. Moreover, many quantities concerning a point set in the plane can be computed by efficient geometric algorithms, which are not available for an arbitrary metric space.

This sounds too good to be generally true: indeed, there are even finite metric spaces that cannot be exactly represented either in the plane or in *any* Euclidean space; for instance, the four vertices of the graph $K_{1,3}$ (a star with 3 leaves) with the shortest-path metric (see Figure 8.0.1a). However, it *is* possible to embed the latter metric in a Euclidean space, if we allow the distances to be distorted somewhat. For example, if we place the center of the star at the origin in $\mathbb{R}^3$ and the leaves at $(1, 0, 0), (0, 1, 0), (0, 0, 1)$, then all distances are preserved *approximately*, up to a factor of $\sqrt{2}$ (Figure 8.0.1b).

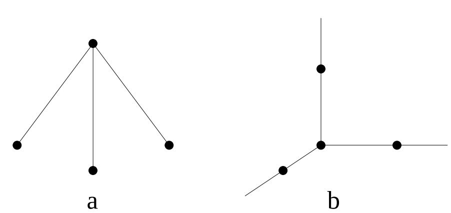

FIGURE 8.0.1
A nonembeddable metric space.

Approximate embeddings have proven extremely helpful for approximate solutions of problems dealing with distances. For many important algorithmic problems, they yield the only known good approximation algorithms.

The normed spaces usually considered for embeddings of finite metrics are the spaces ℓ_p^d, $1 \le p \le \infty$, and the cases $p = 1, 2, \infty$ play the most prominent roles.

GLOSSARY

Metric space: A pair (X, D), where X is a set of *points* and $D: X \times X \to [0, \infty)$ is a *distance function* satisfying the following conditions for all $x, y, z \in X$:

(i) $D(x, y) = 0$ if and only if $x = y$,

(ii) $D(x, y) = D(y, x)$ (symmetry), and

(iii) $D(x, y) + D(y, z) \geq D(x, z)$ (triangle inequality).

Separable metric space: A metric space (X, D) containing a countable dense set; that is, a countable set Y such that for every $x \in X$ and every $\varepsilon > 0$ there exists $y \in Y$ with $D(x, y) < \varepsilon$.

Pseudometric: Like metric except that (i) is not required.

Isometry: A mapping $f: X \to X'$, where (X, D) and (X', D') are metric spaces, with $D'(f(x), f(y)) = D(x, y)$ for all x, y.

(Real) normed space: A real vector space Z with a mapping $\|\cdot\|_Z: Z \to [0, \infty]$, the *norm*, satisfying $\|x\|_Z = 0$ iff $x = 0$, $\|\alpha x\|_Z = |\alpha| \cdot \|x\|_Z$ ($\alpha \in \mathbb{R}$), and $\|x + y\|_Z \leq \|x\|_Z + \|y\|_Z$. The metric on Z is given by $(x, y) \mapsto \|x - y\|_Z$.

ℓ_p^d: The space $\mathbb{R}^d$ with the ℓ_p-norm $\|x\|_p = \left(\sum_{i=1}^d |x_i|^p\right)^{1/p}$, $1 \leq p \leq \infty$ (where $\|x\|_\infty = \max_i |x_i|$).

Finite ℓ_p metric: A finite metric space isometric to a subspace of ℓ_p^d for some d.

ℓ_p: For a sequence $(x_1, x_2, \ldots)$ of real numbers we set $\|x\|_p = \left(\sum_{i=1}^\infty |x_i|^p\right)^{1/p}$. Then ℓ_p is the space consisting of all x with $\|x\|_p < \infty$, equipped with the norm $\|\cdot\|_p$. It contains every finite ℓ_p metric as a (metric) subspace.

Distortion: A mapping $f: X \to X'$, where (X, D) and (X', D') are metric spaces, is said to have distortion at most c, or to be a *c-embedding*, where $c \geq 1$, if there is an $r \in (0, \infty)$ such that for all $x, y \in X$,

$$r \cdot D(x, y) \leq D'(f(x), f(y)) \leq cr \cdot D(x, y).$$

If X' is a normed space, we usually require $r = \frac{1}{c}$ or $r = 1$.

Order of congruence: A metric space (X, D) has order of congruence at most m if every finite metric space that is not isometrically embeddable in (X, D) has a subspace with at most m points that is not embeddable in (X, D).

8.1 THE SPACES ℓ_p

8.1.1 THE EUCLIDEAN SPACES ℓ_2^d

Among normed spaces, the Euclidean spaces are the most familiar, the most symmetric, the simplest in many respects, and the most restricted. Every finite ℓ_2 metric embeds isometrically in ℓ_p for all p. More generally, we have the following Ramsey-type result on the "universality" of ℓ_2; see, e.g., [MS86]:

THEOREM 8.1.1 *Dvoretzky's theorem (a finite quantitative version)*

For every d and every $\varepsilon > 0$ there exists $n = n(d, \varepsilon) \le 2^{O(d/\varepsilon^2)}$ such that ℓ_2^d can be $(1+\varepsilon)$-embedded in every n-dimensional normed space.

Isometric embeddability in ℓ_2 has been well understood since the classical works of Menger, von Neumann, Schoenberg, and others (see, e.g., [Sch38]). Here is a brief summary:

THEOREM 8.1.2

(i) *(Compactness) A separable metric space (X, D) is isometrically embeddable in ℓ_2 iff each finite subspace is so embeddable.*

(ii) *(Order of congruence) A finite (or separable) metric space embeds isometrically in ℓ_2^d iff every subspace of at most $d + 3$ points so embeds.*

(iii) *For a finite $X = \{x_0, x_1, \ldots, x_n\}$, (X, D) embeds in ℓ_2 iff the $n \times n$ matrix $\big(D(x_0, x_i)^2 + D(x_0, x_j)^2 - D(x_i, x_j)^2\big)_{i,j=1}^n$ is positive semidefinite; moreover, its rank is the smallest dimension for such an embedding.*

(iv) *(Schoenberg's criterion) A separable (X, D) isometrically embeds in ℓ_2 iff the matrix $\big(e^{-\lambda D(x_i, x_j)^2}\big)_{i,j=1}^n$ is positive semidefinite for all $n \ge 1$, for any points $x_1, x_2, \ldots, x_n \in X$, and for any $\lambda > 0$. (This is expressed by saying that the functions $x \mapsto e^{-\lambda x^2}$, for all $\lambda > 0$, are* positive definite *on ℓ_2.)*

Using similar ideas, the problem of finding the smallest c such that a given finite (X, D) can be c-embedded in ℓ_2 can be formulated as a semidefinite programming problem and thus solved in polynomial time [LLR95] (but no similar result is known for embedding in ℓ_2^d with d given!).

8.1.2 THE SPACES ℓ_1^d

GLOSSARY

Cut metric: A pseudometric D on a set X such that, for some partition $X = A \dot\cup B$, we have $D(x, y) = 0$ if both $x, y \in A$ or both $x, y \in B$, and $D(x, y) = 1$ otherwise.

Hypermetric inequality: A metric space (X, D) satisfies the $(2k+1)$-point hypermetric inequality (also called the $(2k+1)$-gonal inequality) if for every multiset A of k points and every multiset B of $k + 1$ points in X, $\sum_{a,a' \in A} D(a, a') + \sum_{b,b' \in B} D(b, b') \le \sum_{a \in A, b \in B} D(a, b)$. (We get the triangle inequality for $k = 1$.)

Hypermetric space: A space that satisfies the hypermetric inequality for all k.

Cocktail-party graph: The complement of a perfect matching in a complete graph K_{2m}; also called a ***hyperoctahedron graph***.

Half-cube graph: The vertex set consists of all vectors in $\{0, 1\}^n$ with an even number of 0's, and edges connect vectors with Hamming distance 2.

Cartesian product of graphs G and H: The vertex set is $V(G) \times V(H)$, and the edge set is $\{\{(u, v), (u, v')\} \mid u \in V(G), \{v, v'\} \in E(H)\} \cup \{\{(u, v), (u', v)\} \mid \{u, u'\} \in E(G), v \in V(H)\}$. The cubes are Cartesian powers of K_2.

Girth of a graph: The length of the shortest cycle.

The ℓ_1 spaces are important for many reasons, but considerably more complicated than Euclidean spaces; a general reference here is [DL97]. Many important and challenging open problems are related to embeddings in ℓ_1 or in ℓ_1^d.

Unlike the situation in ℓ_2^n, not every n-point ℓ_1-metric lives in ℓ_1^n; dimension of order n^2 is sometimes necessary and always sufficient to embed n-point ℓ_1-metrics isometrically (similarly for the other ℓ_p-metrics with $p \neq 2$).

The ℓ_1 metrics on an n-point set X are precisely the elements of the ***cut cone***; that is, linear combinations with nonnegative coefficients of cut metrics on X. Another characterization is this: A metric D on $\{1, 2, \ldots, n\}$ is an ℓ_1 metric iff there exist a measure space (Ω, Σ, μ) and sets $A_1, \ldots, A_n \in \Sigma$ such that $D(i, j) = \mu(A_i \triangle A_j)$.

Every ℓ_1 metric is a hypermetric space (since cut metrics satisfy the hypermetric inequalities), but for 7 or more points, this condition is not sufficient. Hypermetric spaces have an interesting characterization in terms of *Delaunay polytopes* of lattices; see [DL97].

ISOMETRIC EMBEDDABILITY

Deciding isometric embeddability in ℓ_1 is NP-hard. On the other hand, the embeddability of *unweighted* graphs, both in ℓ_1 and in a Hamming cube, has been characterized and can be tested in polynomial time. In particular, we have:

THEOREM 8.1.3

(i) *An unweighted graph G embeds isometrically in some cube $\{0, 1\}^m$ with the ℓ_1-metric iff it is bipartite and satisfies the pentagonal inequality.*

(ii) *An unweighted graph G embeds isometrically in ℓ_1 iff it is an isometric subgraph of a Cartesian product of half-cube graphs and cocktail-party graphs.*

A first characterization of cube-embeddable graphs was given by Djokovic [Djo73], and the form in (i) is due to Avis (see [DL97]). Part (ii) is from Shpectorov [Shp93].

ORDER OF CONGRUENCE

The isometric embeddability in ℓ_1^2 is characterized by 6-point subspaces (6 is best possible here), and can thus be tested in polynomial time (Bandelt and Chepoi [BC96]). The proof uses a result of Bandelt and Dress [BD92] of independent interest, about certain canonical decompositions of metric spaces (see also [DL97]).

On the other hand, for no $d \geq 3$ it is known whether the order of congruence of ℓ_1^d is finite; there is a lower bound of d^2 (for odd d) or $d^2 - 1$ (for d even).

8.1.3 THE OTHER p

The spaces ℓ_∞^d are the richest (and thus generally the most difficult to deal with); *every* n-point metric space (X, D) embeds isometrically in ℓ_∞^n. To see this, write $X = \{x_1, x_2, \ldots, x_n\}$ and define $f\colon X \to \ell_\infty^n$ by $f(x_i)_j = D(x_i, x_j)$.

The other $p \neq 1, 2, \infty$ are encountered less often, but it may be useful to know the cases where all ℓ_p metrics embed with bounded distortion in ℓ_q: This happens iff $p = q$, or $p = 2$, or $q = \infty$, or $1 \leq q \leq p \leq 2$. Isometric embeddings exist in all these cases. Moreover, for $1 \leq q \leq p \leq 2$, the whole of ℓ_p^d can be $(1+\varepsilon)$ embedded in ℓ_q^{Cd} with a suitable $C = C(p, q, \varepsilon)$ (so the dimension doesn't grow by much); see, e.g., [MS86]. These embeddings are probabilistic. The simplest one is $\ell_2^d \to \ell_1^{Cd}$, given by $x \mapsto Ax$ for a random ± 1 matrix A of size $Cd \times d$ (surprisingly, no good explicit embedding is known even in this case).

8.2 APPROXIMATE EMBEDDINGS OF GENERAL METRICS IN ℓ_p

8.2.1 BOURGAIN'S EMBEDDING IN ℓ_2

The mother of most embeddings mentioned in the next few sections, from both historical and "technological" points of view, is the following theorem.

THEOREM 8.2.1 *Bourgain* [Bou85]

Any n-point metric space (X, D) can be embedded in ℓ_2 (in fact, in every ℓ_p) with distortion $O(\log n)$.

We describe the embedding, which is constructed probabilistically. We set $m = \lfloor \log_2 n \rfloor$ and $q = \lfloor C \log n \rfloor$ (C a suitable constant) and construct an embedding in ℓ_2^{mq}, with the coordinates indexed by $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, q$. For each such i, j, we select a subset $A_{ij} \subseteq X$ by putting each $x \in X$ into A_{ij} with probability 2^{-j}, all the random choices being mutually independent. Then we set $f(x)_{ij} = D(x, A_{ij})$. We thus obtain an embedding in $\ell_2^{O(\log^2 n)}$ (Bourgain's original proof used exponential dimension; the possibility of reducing it was noted later), and it can be shown that the distortion is $O(\log n)$ with high probability.

This yields an $O(n^2 \log n)$ randomized algorithm for computing the desired embedding. The algorithm can be derandomized (preserving the polynomial time and the dimension bound) using the method of conditional probabilities; this result seems to be folklore. Alternatively, it can be derandomized using small sample spaces [LLR95]; this, however, uses dimension $\Theta(n^2)$. Finally, as was remarked above, an embedding of a given space in ℓ_2 with optimal distortion can be computed by semidefinite programming.

The $O(\log n)$ distortion for embedding a general metric in ℓ_2 is tight [LLR95] (and similarly for ℓ_p, $p < \infty$ fixed). Examples of metrics that cannot be embedded any better are the shortest-path metrics of constant-degree expanders. (An n-vertex graph is a ***constant-degree expander*** if all degrees are bounded by some constant r and each subset of k vertices has at least βk outgoing edges, for $1 \le k \le \frac{n}{2}$ and for some constant $\beta > 0$ independent of n.)

Another interesting lower bound is due to Linial et al. [LMN02]: The shortest-path metric of *any* k-regular graph ($k \ge 3$) of girth g requires $\Omega(\sqrt{g}\,)$ distortion for embedding in ℓ_2.

8.2.2 THE DIMENSION OF EMBEDDINGS IN ℓ_∞

If we want to embed all n-point metrics in ℓ_∞^d, there is a tradeoff between the dimension d and the worst-case distortion. The following result was proved in [Mat96] by adapting Bourgain's technique.

THEOREM 8.2.2

For an integer $b > 0$ set $c = 2b-1$. Then any n-point metric space can be embedded in ℓ_∞^d with distortion c, where $d = O(bn^{1/b}\log n)$.

An almost matching lower bound can be proved using graphs without short cycles, an idea also going back to [Bou85]. Let $m(g, n)$ be the maximum possible number of edges of an n-vertex graph of girth $g + 1$. For every fixed $c \geq 1$ and integer $g > c$ there exists an n-point metric space such that any c-embedding in ℓ_∞^d has $d = \Omega(m(g, n)/n)$ [Mat96]. The proof goes by counting: Fix a graph G_0 witnessing $m(g, n)$, and let $\mathcal{G}$ be the set of graphs (considered with the shortest-path metric) that can be obtained from G_0 by deleting some edges. It turns out that if $G, G' \in \mathcal{G}$ are distinct, then they cannot have "essentially the same" c-embeddings in ℓ_∞^d, and there are only "few" essentially different embeddings in ℓ_∞^d if d is small.

It is easy to show that $m(g, n) = O(n^{1+1/\lfloor g/2\rfloor})$ for all g, and this is conjectured to be the right order of magnitude [Erd64]. This has been verified for $g \leq 7$ and for $g = 10, 11$, while only worse lower bounds are known for the other values of g (with exponent roughly $1 + 4/3g$ for g large). Whenever the conjecture holds for some $g = 2b - 1$, the above theorem is tight up to a logarithmic factor for the corresponding b. Unfortunately, although explicit constructions of graphs of a given girth with many edges are known, the method doesn't provide explicit examples of badly embeddable spaces.

DISTANCE ORACLES

An interesting algorithmic result, conceptually resembling the above theorem, was obtained by Thorup and Zwick [TZ01]. They showed that for an integer $b > 0$, every n-point metric space can be stored in a data structure of size $O(n^{1+1/b})$ (with preprocessing time of the same order) so that, within time $O(b)$, the distance between any two points can be approximated within a multiplicative factor of $2b-1$.

LOW DIMENSION

The other end of the tradeoff between distortion and dimension d, where d is fixed (and then all ℓ_p-norms on $\mathbb{R}^d$ are equivalent up to a constant) was investigated in [Mat90]. For all fixed $d \geq 1$, there are n-point metric spaces requiring distortion $\Omega\left(n^{1/\lfloor(d+1)/2\rfloor}\right)$ for embedding in ℓ_2^d (for $d = 2$, an example is the shortest-path metric of K_5 with every edge subdivided $n/10$ times). On the other hand, every n-point space $O(n)$-embeds in ℓ_2^1 (the real line), and $O(n^{2/d}\log^{3/2} n)$-embeds in ℓ_2^d, $d \geq 3$.

8.2.3 THE JOHNSON-LINDENSTRAUSS LEMMA: FLATTENING IN ℓ_2

The n-point ℓ_2 metric with all distances equal to 1 requires dimension $n - 1$ for isometric embedding in ℓ_2. A somewhat surprising and extremely useful result shows that, in particular, this metric can be embedded in dimension only $O(\log n)$ with distortion close to 1.

THEOREM 8.2.3 *Johnson and Lindenstrauss* [JL84]

For every $\varepsilon > 0$, any n-point ℓ_2 metric can be $(1+\varepsilon)$-embedded in $\ell_2^{O(\log n/\varepsilon^2)}$.

There is an almost matching lower bound for the necessary dimension, due to Alon (see [Mat02a]): $\Omega(\log n/(\varepsilon^2 \log(1/\varepsilon)))$.

All known proofs (see, e.g., [Ach01] for references and an insightful discussion) first place the metric under consideration in ℓ_2^n and then map it into ℓ_2^d by a random linear map $A\colon \ell_2^n \to \ell_2^d$. Here A can be a random orthogonal projection (as in [JL84]). It can also be given by a random $n \times d$ matrix with independent $N(0,1)$ entries [IM98], or even one with independent uniform random ± 1 entries. The proof in the last case, due to [Ach01], is considerably more difficult than the previous ones (which use spherically symmetric distributions), but this version has advantages in applications.

An embedding as in the theorem can be computed deterministically in time $O(n^2 d(\log n + 1/\varepsilon)^{O(1)})$ [EIO02] (also see [Siv02]).

Brinkman and Charikar [BC03] proved that no flattening lemma of comparable strength holds in ℓ_1. Namely, for every fixed $c > 1$, and every n, they exhibit an n-point ℓ_1-metric that cannot be c-embedded into ℓ_1^d unless $d = n^{\Omega(1/c^2)}$. A simpler alternative proof was found later by Lee and Naor (manuscript).

In contrast, [Ind00] showed that for every $0 < \varepsilon < 1$ and any ℓ_1-metric over $X \subset \ell_1^d$, there *is* a $k \times d$ real matrix $[a_1 \ldots a_k]^T$, $k = O(\log |X|/\varepsilon^2)$, such that for any $p, q \in X$, $\|p-q\|_1 \le \mathrm{median}(|a_1(p-q)|, \ldots, |a_k(p-q)|) \le (1+\varepsilon)\|p-q\|_1$.

8.2.4 VOLUME-RESPECTING EMBEDDINGS

Feige [Fei00] introduced the notion of ***volume-respecting*** embeddings in ℓ_2, with impressive algorithmic applications. While the distortion of a mapping depends only on pairs of points, the volume-respecting condition takes into account the behavior of k-tuples. For an arbitrary k-point metric space (S, D), we set $\mathrm{Vol}(S) = \sup_{\text{nonexpanding } f\colon S \to \ell_2} \mathrm{Evol}(f(S))$, where $\mathrm{Evol}(P)$ is the $(k-1)$-dimensional volume of the convex hull of P (in ℓ_2). Given a nonexpanding $f\colon X \to \ell_2$ for some metric space (X, D) with $|X| \ge k$, we define the k-distortion of f to be

$$\sup_{S \subseteq X, |S|=k} \left(\frac{\mathrm{Vol}(S)}{\mathrm{Evol}(f(S))} \right)^{1/(k-1)}$$

If the k-distortion of f is Δ, we call f (k, Δ)-*volume-respecting*.

If $f\colon X \to \ell_2$ is an embedding scaled so that it is nonexpanding but just so, the 2-distortion coincides with the usual distortion. But note that for $k > 2$, the isometric "straight" embedding of a path in ℓ_2 is not volume-respecting at all. In fact, it is known that for any $k > 2$, *no* $(k, o(\sqrt{\log n}))$-volume-respecting embedding of a *line* exists [DV01].

Extending Bourgain's technique, Feige proved that for every $k > 2$, every n-point metric space has a $(k, O(\log n + \sqrt{k \log n \log k}))$-volume-respecting embedding in ℓ_2.

8.3 APPROXIMATE EMBEDDING OF SPECIAL METRICS IN ℓ_p

GLOSSARY

$\mathcal{G}$-metric: Let $\mathcal{G}$ be a class of graphs and let $G \in \mathcal{G}$. Each positive weight function $w\colon E(G) \to (0, \infty)$ defines a metric D_w on $V(G)$, namely, the shortest-path metric, where the length of a path is the sum of the weights of its edges. A metric space is a $\mathcal{G}$-metric if it is isometric to a subspace of $(V(G), D_w)$ for some $G \in \mathcal{G}$ and some w.

Tree metric, planar-graph metric: A $\mathcal{G}$-metric for $\mathcal{G}$, the class of all trees or all planar graphs, respectively.

Minor: A graph G is a minor of a graph H if it can be obtained from H by repeated deletions of edges and contractions of edges.

8.3.1 TREE METRICS, PLANAR-GRAPH METRICS, AND FORBIDDEN MINORS

A major research direction has been improving Bourgain's embedding in ℓ_2 for restricted families of metric spaces.

TREE METRICS

It is easy to show that any tree metric embeds isometrically in ℓ_1. Any n-point tree metric can also be embedded isometrically in $\ell_\infty^{O(\log n)}$ [LLR95]. For ℓ_p embeddings, the situation is rather delicate:

THEOREM 8.3.1

Distortion of order $(\log\log n)^{\min(1/2,1/p)}$ *is sufficient for embedding any n-vertex tree metric in ℓ_p ($p \in (1,\infty)$ fixed)* [Mat99]*, and it is also necessary in the worst case (for the complete binary tree;* [Bou86]*).*

Gupta [Gup00] proved that any n-point tree metric $O(n^{1/(d-1)})$-embeds in ℓ_2^d ($d \geq 1$ fixed), and for $d = 2$ and trees with unit-length edges, Babilon et al. [BMMV02] improved this to $O(\sqrt{n})$.

PLANAR-GRAPH METRICS AND OTHER CLASSES WITH A FORBIDDEN MINOR

The following result was proved by Rao, building on the work of Klein, Plotkin, and Rao.

THEOREM 8.3.2 *Rao* [Rao99]

Any n-point planar-graph metric can be embedded in ℓ_2 with distortion $O(\sqrt{\log n})$. More generally, let H be an arbitrary fixed graph and let $\mathcal{G}$ be the class of all graphs not containing H as a minor; then any n-point $\mathcal{G}$-metric can be embedded in ℓ_2 with distortion $O(\sqrt{\log n})$.

This bound is tight even for series-parallel graphs (no K_4 minor) [NR02]; the example is obtained by starting with a 4-cycle and repeatedly replacing each edge by two paths of length 2.

A challenging conjecture, one that would have significant algorithmic consequences, states that under the conditions of Rao's theorem, all $\mathcal{G}$-metrics can be c-embedded in ℓ_1 for some c depending only on $\mathcal{G}$ (but not on the number of points). Apparently, this conjecture was first published in [GNRS99], where it was verified for the forbidden minors K_4 (series-parallel graphs) and $K_{2,3}$ (outerplanar graphs).

8.3.2 METRICS DERIVED FROM OTHER METRICS

In this section we focus on metrics derived from other metrics, e.g., by defining a distance between two *sets* or *sequences* of points from the underlying metric.

GLOSSARY

Uniform metric: For any set X, the metric (X, D) is uniform if $D(p, q) = 1$ for all $p \neq q$, $p, q \in X$.

Hausdorff distance: For a metric space (X, D), the Hausdorff metric H on the set 2^X of all subsets of X is given by $H(A, B) = \min(\vec{H}(A, B), \vec{H}(B, A))$, where $\vec{H}(A, B) = \sup_{a \in A} \inf_{b \in B} D(a, b)$.

Earth-mover distance: For a metric space (X, D) and an integer $d \geq 1$, the earth-mover distance of two d-element sets $A, B \subseteq X$ is the minimum weight of a perfect matching between A and B; that is, $\min_{\text{bijective } \pi: A \to B} \sum_{a \in A} D(a, \pi(a))$.

Levenshtein distance (or ***edit distance***): For a metric space $M = (\Sigma, D)$, the distance between two strings $w, w' \in \Sigma^*$ is the minimum cost of a sequence of operations that transforms w into w'. The allowed operations are: character insertion (of cost 1), character deletion (of cost 1), or replacement of a symbol a by another symbol b (of cost $D(a, b)$), where $a, b \in \Sigma$. The total cost of the sequence of operations is the sum of all operation costs.

Fréchet distance: For a metric space $M = (X, D)$, the Fréchet distance (also called the ***dogkeeper's distance***) between two functions $f, g: [0, 1] \to X$ is defined as

$$\inf_{\pi:[0,1]\to[0,1]} \sup_{t \in [0,1]} D(f(t), g(\pi(t)))$$

where π is continous, monotone increasing, and such that $\pi(0) = 0, \pi(1) = 1$.

HAUSDORFF DISTANCE

The Hausdorff distance is often used in computer vision for comparing geometric shapes, represented as sets of points. However, even computing a single distance $H(A, B)$ is a nontrivial task. As noted in [FCI99], for any n-point metric space (X, D), the Hausdorff metric on 2^X can be isometrically embedded in ℓ_∞^n.

The dimension of the host norm can be further reduced if we focus on embedding particular Hausdorff metrics. In particular, let H_M^s be the Hausdorff metric over all s-subsets of M. Farach-Colton and Indyk [FCI99] showed that if $M = (\{1, \ldots, \Delta\}^k, \ell_p)$, then H_M^s can be embedded in $\ell_\infty^{d'}$ with distortion $1 + \varepsilon$, where $d' = O(s^2 (1/\varepsilon)^{O(k)} \log \Delta)$. For a general (finite) metric space $M = (X, D)$ they show that H_M^s can be embedded in $\ell_\infty^{s^{O(1)} |X|^\alpha \log \Delta}$ for any $\alpha > 0$ with constant distortion, where $\Delta = (\min_{p \neq q \in X} D(p, q)) / (\max_{p,q \in X} D(p, q))$.

EARTH-MOVER DISTANCE (EMD)

A very interesting relation between embedding EMD in normed spaces and embeddings in probabilistic trees (discussed below in Section 8.4.1) was discovered in [Cha02]: If a finite metric space can be embedded in a convex combination of dominating trees with distortion c (see definitions in Section 8.4), then the EMD over it can be embedded in ℓ_1 with distortion $O(c)$. Consequently, the EMD over any n-point metric can be embedded in ℓ_1 with distortion $O(\log n)$.

LEVENSHTEIN DISTANCE AND ITS VARIANTS

The Levenshtein distance is used in text processing and computational biology. The best algorithm computing the Levenshtein distance of two strings w, w', even approximately, has running time of order $|w| \cdot |w'|$ (for a constant-size Σ). Not much is known about embeddability of this metric in normed spaces, even in the simplest (but nevertheless quite common) case of the uniform metric over $\Sigma = \{0, 1\}$. It is known, however, that the Levenshtein metric, restricted to a certain set of strings, is isomorphic to the shortest path metric over $K_{2,n}$ [ADG$^+$03]; this implies that it cannot be embedded in ℓ_1 (or even the square of ℓ_2) with distortion better than $3/2 - O(1/n)$.

However, if we modify the definition of the distance by permitting the movement of an arbitrarily long contiguous block of characters as a single operation, and if the underlying metric is uniform, then the resulting ***block-edit*** metric can be embedded in ℓ_1 with distortion $O(\log l \cdot \log^* l)$, where l is the length of the embedded strings (see [MS00, CM02] and references therein). The modified metric has applications in computational biology and in string compression. The embedding of a given string can be computed in almost linear time, which yields a very fast approximation algorithm for computing the distance between two strings (the exact distance computation is NP-hard!).

FRÉCHET METRIC

The Fréchet metric is an interesting metric measuring the distances between two *curves*. From the applications perspective, it is interesting to investigate the case where $M = \ell_2^k$ and f, g are continuous, closed polygonal chains, consisting of (say) at most d segments each. Denote the set of such curves by C_d^k. It is not known whether C_d^k, under Fréchet distance, can be embedded in ℓ_∞ with finite dimension (for infinite dimension, an isometric embedding follows from the unversality of the ℓ_∞ norm). On the other hand, it is easy to check that for any bounded set $S \subset \ell_\infty^d$, there is an isometry $f \colon S \to C_{3d}^1$.

8.3.3 OTHER SPECIAL METRICS

GLOSSARY

$(1,2)$-*B metric:* A metric space (X, D) such that for any $x \in X$ the number of points y with $D(x, y) = 1$ is at most B, and all other distances are equal to 2.

Transposition distance: The (unfortunately named) metric D_T on the set of all permutations of $\{1, 2, \ldots, n\}$; $D_T(\pi_1, \pi_2)$ is the minimum number of moves of contiguous subsequences to arbitrary positions needed to transform π_1 into π_2.

BOUNDED DISTANCE METRICS

Trevisan [Tre01] considered approximate embeddings of $(1,2)$-B metrics in ℓ_p^d (in a sense somewhat different from low-distortion embeddings). Guruswami and Indyk [GI03] proved that any $(1,2)$-B metric can be isometrically embedded in $\ell_\infty^{O(B\log n)}$.

PERMUTATION METRICS

It was shown in [CMS01] that D_T can be $O(1)$-embedded in ℓ_1; similar results were obtained for other metrics on permutations, including reversal distance and permutation edit distance.

8.4 APPROXIMATE EMBEDDINGS IN RESTRICTED METRICS

GLOSSARY

Dominating metric: Let D, D' be metrics on the same set X. Then D' dominates D if $D(x,y) \geq D'(x,y)$ for all $x, y \in X$.

Convex combination of metrics: Let X be a set, $T_1, T_2, \ldots, T_k$ metrics on it, and $\alpha_1, \ldots, \alpha_k$ nonnegative reals summing to 1. The convex combination of the T_i (with coefficients α_i) is the metric D given by $D(x,y) = \sum_{i=1}^k \alpha_i T_i(x,y)$, $x, y \in X$.

Hierarchically well-separated tree (k-HST): A 1-HST is exactly an *ultrametric*; that is, the shortest-path metric on the leaves of a rooted tree T (with weighted edges) such that all leaves have the same distance from the root. For a k-HST with $k > 1$ we require that, moreover, $\Delta(v) \leq \Delta(u)/k$ whenever v is a child of u in T, where $\Delta(v)$ denotes the diameter of the subtree rooted at v (w.l.o.g. we may assume that each non-leaf has degree at least 2, and so $\Delta(v)$ equals the distance of v to the nearest leaves). *Warning:* This is a newer definition introduced in [BBM01]. Older papers, such as [Bar96, Bar98], used another definition, but the difference is merely technical, and the notion remains essentially the same.

8.4.1 PROBABILISTIC EMBEDDINGS IN TREES

A convex combination $\overline{D} = \sum_{i=1}^r \alpha_i T_i$ of some metrics $T_1, \ldots, T_r$ on X can be thought of as a ***probabilistic metric*** (this concept was suggested by Karp). Namely, $\overline{D}(x,y)$ is the expectation of $T_i(x,y)$ for $i \in \{1, 2, \ldots, r\}$ chosen at random according to the distribution given by the α_i. Of particular interest are embeddings in convex combinations of *dominating* metrics. The domination requirement is crucial for many applications. In particular, it enables one to solve many problems over the original metric (X, D) by solving them on a (simple) metric chosen at random from $T_1, \ldots, T_r$ according to the distribution defined by the α_i.

The usefulness of probabilistic metrics comes from the fact that a sum of metrics is much more powerful than each individual metric. For example, it is not difficult to show that there are metrics (e.g., cycles [RR98, Gup01]) that cannot be embedded in tree metrics with $o(n)$ distortion. In contrast, we have the following result:

THEOREM 8.4.1 *Fakcharoenphol, Rao, and Talwar* [FRT03]

Let (X, D) be any n-point metric space. For every $k > 1$, there exist a natural number r, k-HST metrics $T_1, T_2, \ldots, T_r$ on X, and coefficients $\alpha_1, \ldots, \alpha_r > 0$ summing to 1 such that each T_i dominates D, and the (identity) embedding of (X, D) into$(X, \overline{D})$, where $\overline{D} = \sum_{i=1}^{r} \alpha_i T_i$, has distortion $O((k/\log k) \cdot \log n)$.

The first result of this type was obtained by Alon et al [AKPW95]. Their embedding has distortion $2^{O(\sqrt{\log n \log\log n})}$, and uses convex combinations of the metrics induced by spanning trees of M. A few years later Bartal [Bar96] improved the distortion bound considerably, to $O(\log^2 n)$ and later even to $O(\log n \log\log n)$ [Bar98]. The bound on the distortion in the theorem above is optimal up to a constant factor for every fixed k, since any convex combination of tree metrics embeds isometrically into ℓ_1.

The constructions in [Bar96, Bar98, FRT03] generate trees with Steiner nodes (i.e., nodes that do not belong to X). However, one can get rid of such nodes in *any* tree while increasing the distortion by at most 8 [Gup01].

An interesting extra feature of the construction of Alon et al. mentioned above is that if the metric D is given as the shortest-path metric of a (weighted) graph G on the vertex set X, then all the T_i are spanning trees of this G. None of the constructions in [Bar96, Bar98, FRT03] share this property.

The embedding algorithms in Bartal's papers [Bar96, Bar98] are randomized and run in polynomial time. A deterministic algorithm for the same problem was given in [CCG$^+$98]. The latter algorithm constructs a distribution over $O(n \log n)$ trees (the number of trees in Bartal's construction was exponential in n).

8.4.2 RAMSEY-TYPE THEOREMS

Many Ramsey-type questions can be asked in connection with low-distortion embeddings of metric spaces. For example, given classes $\mathcal{X}$ and $\mathcal{Y}$ of finite metric spaces, one can ask whether for every n-point space $Y \in \mathcal{Y}$ there is an m-point $X \in \mathcal{X}$ such that X can be α-embedded in Y, for given n, m, α.

Important results were obtained in [BBM01], and later greatly improved and extended in [BLMN03], for $\mathcal{X}$ the class of all k-HST and $\mathcal{Y}$ the class of all finite metric spaces; they were used for a lower bound in a significant algorithmic problem (metrical task systems). Let us quote some of the numerous results of Bartal et al.:

THEOREM 8.4.2 *Bartal, Linial, Mendel, and Naor* [BLMN03]

Let $R_{\mathrm{UM}}(n, \alpha)$ denote the largest m such that for every n-point metric space Y there exists an m-point 1-HST (i.e., ultrametric) that α-embeds in Y, and let $R_2(n, \alpha)$ be defined similarly with "ultrametric" replaced with "Euclidean metric."

(i) *There are positive constants C, C_1, c such that for every $\alpha > 2$ and all n,*

$$n^{1-C_1(\log\alpha)/\alpha} \leq R_{\mathrm{UM}}(n, \alpha) \leq R_2(n, \alpha) \leq C n^{1-c/\alpha}.$$

(ii) *(Sharp threshold at distortion 2) For every $\alpha > 2$, there exists $c(\alpha) > 0$ such that $R_2(n, \alpha) \geq R_{\mathrm{UM}}(n, \alpha) \geq n^{c(\alpha)}$ for all n, while for every $\alpha \in (1, 2)$, we* *have* $c'(\alpha)\log n \le R_{\mathrm{UM}}(n,\alpha) \le R_2(n,\alpha) \le 2\log n + C'(\alpha)$ *for all* n*, with suitable positive* $c'(\alpha)$ *and* $C'(\alpha)$*.*

For embedding a k-HST in a given space, one can use the fact that every ultrametric is k-equivalent to a k-HST. For an earlier result similar to the second part of (ii), showing that the largest Euclidean subspace $(1{+}\varepsilon)$-embeddable in a general n-point metric space has size $\Theta(\log n)$ for all sufficiently small fixed $\varepsilon > 0$, see [BFM86].

8.4.3 APPROXIMATION BY SPARSE GRAPHS

GLOSSARY

t-spanner: A subgraph H of a graph G (possibly with weighted edges) is a *t-spanner* of G if $D_H(u,v) \le t \cdot D_G(u,v)$ for every $u,v \in V(G)$.

Sparse spanners are useful as a more economic representation of a given graph (note that if H is a t-spanner of G, then the identity map $V(G) \to V(H)$ is a t-embedding).

THEOREM 8.4.3 *Althöfer et al.* [ADD$^+$93]

For every integer $t \ge 2$*, every* n*-vertex graph* G *has a* t*-spanner with at most* $m(t,n)$ *edges, where* $m(g,n) = O(n^{1+1/\lfloor g/2 \rfloor})$ *is the maximum possible number of edges of an* n*-vertex graph of girth* $g+1$*.*

The proof is extremely simple: Start with empty H, consider the edges of G one by one from the shortest to the longest, and insert each edge into the current H unless it creates a cycle with at most t edges. It is also immediately seen that the bound $m(t,n)$ is the best possible in the worst case.

Rabinovich and Raz [RR98] proved that there are (unweighted) n-vertex graphs G that cannot be t-embedded in graphs (possibly weighted) with fewer than $m(\Omega(t),n)$ edges (for t sufficiently large and n sufficiently large in terms of t). Their main tool is the following lemma, proved by elementary topological considerations: If H is a simple unweighted connected n-vertex graph of girth g and G is a (possibly weighted) graph on at least n vertices with $\chi(G) < \chi(H)$, then H cannot be c-embedded in G for $c < g/4 - 3/2$; here $\chi(G)$ denotes the ***Euler characteristic*** of a graph G, which, for G connected, equals $|E(G)| - |V(G)| + 1$.

8.5 ALGORITHMIC APPLICATIONS OF EMBEDDINGS

In this section we give a brief overview of the scenarios in which embeddings have been used in the design of algorithms and for determining computational complexity. For a more detailed survey, see [Ind01].

The most typical scenario is as follows. Suppose we have a problem defined over a set of points in a metric space M. If the metric space is "complex" enough, the problem is likely to be NP-hard. To solve the problem, we embed the metric in a "simple" metric M', and solve the problem there. This gives an approximation algorithm for the original problem, whose approximation factor depends on the distortion of the embedding.

The implementation of this general paradigm depends on "complex" and "simple" metrics M and M'. The most frequent scenarios are as follows:

1. *General metrics $\rightarrow$ tree metrics.* This approach uses the theorems of [Bar98, FRT03], which enable the embedding of an arbitrary finite metric space, in a "probabilistic" way, in tree metrics, with low distortion. It is not difficult to see that if the goal of the original problem is to minimize a linear function of the interpoint distances, then the properties guaranteed by the above embedding are sufficient to show that given a c-approximation algorithm for HST's (or trees, resp.), one can construct an $O(c \log n \log\log n)$-approximation (or $O(c \log n)$-approximation, resp.) algorithm for the original metric. Since the random choice of a tree does not depend on the function to be optimized, this approach works even if the optimization function is not known in advance. Thus, this approach has been very successful for both *offline* and *online* problems. In particular, it led to a polylog(n)-competitive algorithm [BBBT97] for metrical task systems, resolving a long-standing conjecture. In the latter paper, the embedding in HST's (as opposed to general trees) is crucial to obtain the result.

2. *General metric $\rightarrow$ low-dimensional normed spaces.* In this case we use Bourgain's or Matoušek's theorem to obtain a low-dimensional approximate representation of a metric. Since the host metric is low-dimensional, each point can be represented using a small number of bits. This has interesting consequences for approximate proximity-preserving labeling [Pel99, GPPR01].

3. *Specific metrics $\rightarrow$ normed spaces.* This approach uses the results of Section 8.3.2, which provide embeddings of certain metrics (e.g., Hausdorff or Levenshtein metrics) in normed spaces. This enables the use of algorithmic tools designed for normed spaces (see, e.g., Chapter 39 of this Handbook) for problems defined over more complex metrics.

4. *High-dimensional spaces $\rightarrow$ low dimensional spaces.* Here, we use dimensionality reduction techniques, notably the Johnson-Lindenstrauss theorem. In this way, we reduce the dimension of the original space to $O(\log n)$, which yields significant savings in the running time and/or space. The improvement is particularly impressive if an algorithm for the original problem uses space/time *exponential* in the dimension (see, e.g., Chapter 39).

   We note, however, that for most applications, the embedding properties listed in the statement of Theorem 8.2.3 are not sufficient. Instead, one must often use additional properties of the embedding, such as:

   - The embedding is chosen at random, independently of the input point set. This property is crucial in situations where not all points are known in advance (e.g., for the nearest neighbor problem).
   - The mapping is linear. This property is used, e.g., for dimensionality reduction theorems for hyperplanes (i.e., when the input set can consist of points, lines, planes etc.) [Mag02], and for low-space computation as described below.

- The coefficients of the mapping matrix are chosen independently of each other (this property holds for *some* but not *all* proofs of dimensionality reduction theorems). This property is useful, e.g., if we want to obtain deterministic versions of dimensionality reduction theorems [Ind00, Siv02, EIO02], which have applications to the derandomization of approximation algorithms based on semidefinite programming.

5. *"Complex" normed spaces → "simple" normed spaces.* The "complexity" of a normed space clearly depends on the problem we want to solve. For example, if we want to find the *diameter* of a set of points, it is very helpful if the interpoint distances are induced by the l_∞^d norm. In this case, the diameter of the point set is equal to the maximum diameter of all one-dimensional point sets, obtained by projecting the (d-dimensional) points onto one of the coordinates. Thus approach gives an $O(nd)$ time for computing the diameter in l_∞^d. However, from Section 8.1 we know that the space l_1^d can be isometrically embedded in $l_\infty^{2^{d-1}}$. Thus, we obtain a linear-time (assuming constant dimension) algorithm for computing the diameter in the l_1 norm. Other embedding results described in Section 8.2 have similar algorithmic applications as well.

A second type of result involves using the embeddings in the "reverse" directions, in order to derive *lower* bounds. Specifically, in order to show a hardness result for a metric M', it suffices to show that a given problem is hard (to approximate) in a metric M that can be embedded in M'. This approach has been used to prove the following results:

- In [Tre01, GI03], it was shown that certain geometric problems (e.g., TSP) are hard to approximate even in $\Theta(\log n)$ dimensions. This was achieved by embedding $(1,2)$-B metrics (known to be the "hard" cases) in $l_p^{O(\log n)}$.
- In [BBM01], it was shown that certain online problems (metrical task systems) do not have $\Omega(\log n / \log^{O(1)} \log n)$-competitive algorithms. This was achieved by showing that "large" HST metrics can be embedded in arbitrary finite metrics, and proving a lower bound for HST metrics.

Finally, embeddings can be used for problems that, at first sight, do not seem to be "metric" in nature. Notable examples of such an application are approximation algorithms for graph problems, such as the algorithm of [LLR95] for the sparsest cut problem and for graph bandwidth [Fei00]. In particular, the former problem can be phrased as finding a cut metric minimizing a certain objective function. Although the problem is NP-hard, its relaxation that requires finding just a metric (minimizing the same objective function) can be solved in polynomial time via linear programming. The algorithm proceeds by embedding the solution metric in l_1 (with low distortion) and decomposing it into a convex combination of cut metrics. It can be shown that that one of those cut metrics provides an approximate solution to the sparsest cut problem.

Another area whose relation to embeddings is not *a priori* apparent is low-space computing. A prototypical example of such a problem is a data structure that maintains a d-dimensional vector x (under increments/decrements of x's coordinates). When queried, the data structure reports an approximate value of $\|x\|_p$. In particular, the case $p = 0$ corresponds to maintaining an approximate number of nonzero coordinates. Alternatively, one could request a succinct (e.g., piecewise constant with few pieces) approximation of x, viewed as a function from $\{1,\ldots,d\}$ into the reals. Such problems are motivated by database applications.

In order to obtain low-storage algorithms solving such problems, we can apply dimensionality reduction techniques to reduce the dimension, while approximately preserving important properties of x (e.g., its norm, or its best succinct approximation). In this way, we only need to store the image Ax of x. Since the update operations on x are linear, they can be easily transformed into operations on Ax. One also has to ensure that there is no need to store the description of A explicitly; this is done by showing that a "pseudorandom" matrix A is good enough [AMS99, Ind00].

TABLE 8.5.1 A summary of approximate embeddings

FROM	TO	DISTORTION	REFERENCE
any	ℓ_p, $1 \le p < \infty$	$O(\log n)$	[Bou85]
constant-degree expander	ℓ_p, $p < \infty$ fixed	$\Omega(\log n)$	[LLR95]
k-reg. graph, $k \ge 3$, girth g	ℓ_2	$\Omega(\sqrt{g}\,)$	[LMN02]
any	$\ell_\infty^{O(bn^{1/b}\log n)}$	$2b-1$, b=1, 2, ...	[Mat96]
some	$\Omega(n^{1/b})$-dim'l. normed space	$2b-1$, b=1, 2, ... (Erdős's conj.!)	[Mat96]
any	ℓ_1^1	$\Theta(n)$	[Mat90]
any	ℓ_p^d, d fixed	$O(n^{2/d}\log^{3/2} n)$, $\Omega\left(n^{1/\lfloor(d+1)/2\rfloor}\right)$	[Mat90]
ℓ_2 metric	$\ell_2^{O(\log n/\varepsilon^2)}$	$1+\varepsilon$	[JL84]
ℓ_1 metric	$\ell_1^{n^\alpha}$, $0<\alpha<1$	$\Omega(\alpha^{-1/2})$	[BC03]
planar or forbidden minor	ℓ_2	$O(\sqrt{\log n}\,)$	[Rao99]
series-parallel	ℓ_2	$\Omega(\sqrt{\log n}\,)$	[NR02]
planar	$\ell_\infty^{O(\log^2 n)}$	$O(1)$	implicit in [Rao99]
outerplanar or series-parallel	ℓ_1	$O(1)$	[GNRS99]
tree	ℓ_1	1	(folklore)
tree	$\ell_\infty^{O(\log n)}$	1	[LLR95]
tree	ℓ_2	$\Theta((\log\log n)^{1/2})$	[Bou86, Mat99]
tree	ℓ_2^d	$O(n^{1/(d-1)})$	[Gup00]
tree, unit edges	ℓ_2^2	$\Theta(\sqrt{n}\,)$	[BMMV02]
Hausdorff metric over (X,D)	$\ell_\infty^{\lvert X\rvert}$	1	[FCI99]
Hausd. over s-subsets of (X,D)	$\ell_\infty^{s^{O(1)}\lvert X\rvert^\alpha \log\Delta}$	$c(\alpha)$	[FCI99]
Hausd. over s-subsets of ℓ_p^k	$\ell_\infty^{s^2(1/\varepsilon)^{O(k)}\log\Delta}$	$1+\varepsilon$	[FCI99]
EMD over (X,D)	ℓ_1	$O(\log\lvert X\rvert)$	[Cha02, FRT03]
Levenshtein metric	ℓ_1	$\ge 3/2$	[ADG+03]
block-edit metric over Σ^d	ℓ_1	$O(\log d\cdot\log^* d)$	[MS00, CM02]
(1,2)-B metric	$\ell_\infty^{O(B\log n)}$	1	[GI03]; for ℓ_p cf. [Tre01]
any	convex comb. of dom. trees (HSTs)	$O(\log n)$	[FRT03]
any	convex comb. of spanning trees	$2^{O(\sqrt{\log n\log\log n})}$	[AKPW95]

8.6 OPEN PROBLEMS AND WORK IN PROGRESS

The time of writing of this chapter (2002) seems to be a period of particularly rapid development in the area of low-distortion embeddings of metric spaces. Many significant results have recently been achieved, and some of them are still unpublished (or not yet even written). We have tried to mention at least some of them, but it is clear that some parts of the chapter will become obsolete very soon.

Instead of stating open problems here, we refer to a list recently compiled by the second author [Mat02b]. It is available on the Web, and it might occassionally be updated to reflect new developments.

8.7 SOURCES AND RELATED MATERIAL

Discrete metric spaces have been studied from many different points of view, and the area is quite wide and diverse. The low-distortion embeddings treated in this chapter constitute only one particular (although very significant) direction. For recent results in some other directions the reader may consult [Cam00, DDL98, DD96], for instance. For more detailed overviews of the topics surveyed here, with many more references, the reader is referred to Chapter 15 in [Mat02a] (including proofs of basic results) and [Ind01] (with emphasis on algorithmic applications), as well as to [Lin02]. Approximate embeddings of normed spaces are treated, e.g., in [MS86]. A recent general reference for isometric embeddings, especially embeddings in ℓ_1, is [DL97].

RELATED CHAPTERS

Chapter 39: Nearest neighbors in high-dimensional spaces

REFERENCES

[Ach01] D. Achlioptas. Database-friendly random projections. In *Proc. 20th Annu. ACM SIGACT-SIGMOD-SIGART Sympos. Princip. Database Syst.*, pages 274–281, 2001.

[ADD+93] I. Althöfer, G. Das, D.P. Dobkin, D. Joseph, and J. Soares. On sparse spanners of weighted graphs. *Discrete Comput. Geom.*, 9:81–100, 1993.

[ADG+03] A. Andoni, M. Deza, A. Gupta, P. Indyk, and S. Raskhodnikova. Lower bounds for embedding of edit distance into normed spaces. In *Proc. 14th Annu. ACM-SIAM Sympos. Discrete Algor.*, 2003.

[AKPW95] N. Alon, R.M. Karp, D. Peleg, and D. West. A graph-theoretic game and its application to the k-server problem. *SIAM J. Comput.*, 24:78–100, 1995.

[AMS99] N. Alon, Y. Matias, and M. Szegedy. The space complexity of approximating the frequency moments. *J. Comput. Syst. Sci.*, 58:137–147, 1999.

[Bar96] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *Proc. 37th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 184–193, 1996.

[Bar98] Y. Bartal. On approximating arbitrary metrics by tree metrics. In *Proc. 30th Annu. ACM Sympos. Theory Comput.*, pages 161–168, 1998.

[BBBT97] Y. Bartal, A. Blum, C. Burch, and A. Tomkins. A polylog(n)-competitive algorithm for metrical task systems. In *Proc. 29th Annu. ACM Sympos. Theory Comput.*, pages 711–719, 1997.

[BBM01] Y. Bartal, B. Bollobás, and M. Mendel. Ramsey-type theorems for metric spaces with applications to online problems. In *Proc. 42nd Annu. IEEE Sympos. Found. Comput. Sci.*, pages 396–405, 2001.

[BC96] H.-J. Bandelt and V. Chepoi. Embedding metric spaces in the rectilinear plane: a six-point criterion. *Discrete Comput. Geom.*, 15:107–117, 1996.

[BC03] B. Brinkman and M. Charikar. On the impossibility of dimension reduction in ℓ_1. In *Proc. 35th Annu. ACM Sympos. Theory Comput.*, 2003.

[BD92] H.-J. Bandelt and A. Dress. A canonical decomposition theory for metrics on a finite set. *Adv. Math.*, 92:47–105, 1992.

[BFM86] J. Bourgain, T. Figiel, and V. Milman. On Hilbertian subsets of finite metric spaces. *Israel J. Math.*, 55:147–152, 1986.

[BLMN03] Y. Bartal, N. Linial, M. Mendel, and A. Naor. On metric Ramsey-type phenomena. In *Proc. 35th Annu. ACM Sympos. Theory Comput.*, 2003.

[BMMV02] R. Babilon, J. Matoušek, J. Maxová, and P. Valtr. Low-distortion embeddings of trees. In *Proc. Graph Drawing 2001.* Springer-Verlag, Berlin, 2002.

[Bou85] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52:46–52, 1985.

[Bou86] J. Bourgain. The metrical interpretation of superreflexivity in Banach spaces. *Israel J. Math.*, 56:222–230, 1986.

[Cam00] P. Cameron, editor. *Discrete Metric Spaces.* Selected papers from the 3rd International Conference held in Marseille, September 15–18, 1998. European J. Combin., 21 (6), 2000.

[CCG+98] M. Charikar, C. Chekuri, A. Goel, S. Guha, and S.A. Plotkin. Approximating a finite metric by a small number of tree metrics. In *Proc. 39th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 379–388, 1998.

[Cha02] M. Charikar. Similarity estimation techniques from rounding. In *Proc. 34th Annu. ACM Sympos. Theory Comput.*, pages 380–388, 2002.

[CM02] G. Cormode and S. Muthukrishnan. The string edit distance matching problem with moves. In *Proc. 13th Annu. ACM-SIAM Sympos. Discrete Algor.*, pages 667–676, 2002.

[CMS01] G. Cormode, M. Muthukrishnan, and C. Sahinalp. Permutation editing and matching via embeddings. In *Proc. 28th Internat. Colloq. Automata Lang. Program. (ICALP)*, pages 481–492, 2001.

[DD96] W. Deuber and M. Deza, editors. *Discrete Metric Spaces.* Papers from the conference held in Bielefeld, November 18–22, 1994. European J. Combin., 17 (2–3), 1996.

[DDL98] W. Deuber, M. Deza, and B. Leclerc, editors. *Discrete Metric Spaces.* Papers from the International Conference held at the Université Claude Bernard, Villeurbanne, September 17–20, 1996. Discrete Math., 192 (1–3), 1998.

[Djo73] D.Z. Djokovic. Distance preserving subgraphs of hypercubes. *J. Combin. Theory Ser. B*, 14:263–267, 1973.

[DL97] M.M. Deza and M. Laurent. *Geometry of Cuts and Metrics.* Volume 15 of *Algor. Combin.* Springer-Verlag, Berlin, 1997.

[DV01] J. Dunagan and S. Vempala. On Euclidean embeddings and bandwidth minimization. *Proc. 5th Workshop on Randomization and Approximation*, pages 229–240, 2001.

[EIO02] L. Engebretsen, P. Indyk, and R. O'Donnell. Derandomized dimensionality reduction with applications. In *Proc. 13th Annu. ACM-SIAM Sympos. Discrete Algor.*, pages 705–712, 2002.

[Erd64] P. Erdős. Extremal problems in graph theory. *Theory of Graphs and Its Applications (Proc. Sympos. Smolenice, 1963)*, pages 29–36, 1964.

[FCI99] M. Farach-Colton and P. Indyk. Approximate nearest neighbor algorithms for Hausdorff metrics via embeddings. In *Proc. 40th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 171–179, 1999.

[Fei00] U. Feige. Approximating the bandwidth via volume respecting embeddings. *J. Comput. System Sci.*, 60:510–539, 2000.

[FRT03] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *Proc. 35th Annu. ACM Sympos. Theory Comput.*, 2003.

[GI03] V. Guruswami and P. Indyk. Embeddings and non-approximability of geometric problems. In *Proc. 14th Annu. ACM-SIAM Sympos. Discrete Algor.*, 2003.

[GNRS99] A. Gupta, I. Newman, Yu. Rabinovich, and A. Sinclair. Cuts, trees and ℓ_1-embeddings of graphs. In *Proc. 40th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 399–409, 1999.

[GPPR01] C. Gavoille, D. Peleg, S. Perennes, and R. Raz. Distance labeling in graphs. *Proc. 12th Annu. ACM-SIAM Sympos. Discrete Algor.*, pages 210–219, 2001.

[Gup00] A. Gupta. Embedding tree metrics into low dimensional Euclidean spaces. *Discrete Comput. Geom.*, 24:105–116, 2000.

[Gup01] A. Gupta. Steiner nodes in trees don't (really) help. In *Proc. 12th Annu. ACM-SIAM Sympos. Discrete Algor.*, pages 220–227, 2001.

[IM98] P. Indyk and R. Motwani. Approximate nearest neighbors: Towards removing the curse of dimensionality. In *Proc. 30th Annu. ACM Sympos. Theory Comput.*, pages 604–613, 1998.

[Ind00] P. Indyk. Stable distributions, pseudorandom generators, embeddings and data stream computation. In *Proc. 41st Annu. IEEE Sympos. Found. Comput. Sci.*, pages 189–197, 2000.

[Ind01] P. Indyk. Algorithmic applications of low-distortion embeddings. In *Proc. 42nd Annu. IEEE Sympos. Found. Comput. Sci.*, pages 10–33, 2001.

[JL84] W.B. Johnson and J. Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. *Contemp. Math.*, 26:189–206, 1984.

[Lin02] N. Linial. Finite metric spaces—combinatorics, geometry and algorithms. In volume III of *Proc. Internat. Congress Math.*, Beijing, 2002, pages 573–586.

[LLR95] N. Linial, E. London, and Yu. Rabinovich. The geometry of graphs and some its algorithmic applications. *Combinatorica*, 15:215–245, 1995.

[LMN02] N. Linial, A. Magen, and A. Naor. Euclidean embeddings of regular graphs—the girth lower bound. *Geom. Funct. Anal.*, 12:380–394, 2002.

[Mag02] A. Magen. Dimensionality reductions that preserve volumes and distance to affine spaces, and their algorithmic applications. In *Proc. 6th RANDOM*, pages 239–253, 2002.

[Mat90] J. Matoušek. Bi-Lipschitz embeddings into low-dimensional Euclidean spaces. *Comment. Math. Univ. Carolin.*, 31:589–600, 1990.

[Mat96] J. Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel J. Math.*, 93:333–344, 1996.

[Mat99] J. Matoušek. On embedding trees into uniformly convex Banach spaces. *Israel J. Math*, 114:221–237, 1999.

[Mat02a] J. Matoušek. *Lectures on Discrete Geometry*. Springer-Verlag, New York, 2002.

[Mat02b] J. Matoušek, editor. Open problems, Workshop on Discrete Metric Spaces and their Algorithmic Applications, Haifa, March 3–7, 2002. KAM Series (Tech. Report), Department of Applied Mathematics, Charles University, Prague, 2002. Available at `http://kam.mff.cuni.cz/~matousek/haifaop.ps`.

[MS86] V.D. Milman and G. Schechtman. *Asymptotic Theory of Finite Dimensional Normed Spaces*. Volume 1200 of *Lecture Notes in Math.* Springer-Verlag, Berlin, 1986.

[MS00] S. Muthukrishnan and C. Sahinalp. Approximate nearest neighbors and sequence comparison with block operations. In *Proc. 32nd Annu. ACM Sympos. Theory Comput.*, pages 416–424, 2000.

[NR02] I. Newman and Yu. Rabinovich. A lower bound on the distortion of embedding planar metrics into Euclidean space. *Discrete Comput. Geom.*, 29:77–81, 2003.

[Pel99] D. Peleg. Proximity-preserving labeling schemes and their applications. *Proc. 25th Workshop on Graph-Theoretic Aspects of Comput. Sci.*, volume 1665 of *Lecture Notes in Comput. Sci.*, Springer-Verlag, New York, pages 30–41, 1999.

[Rao99] S. Rao. Small distortion and volume respecting embeddings for planar and Euclidean metrics. In *Proc. 15th Annu. ACM Sympos. Comput. Geom.*, pages 300–306, 1999.

[RR98] Yu. Rabinovich and R. Raz. Lower bounds on the distortion of embedding finite metric spaces in graphs. *Discrete Comput. Geom.*, 19:79–94, 1998.

[Sch38] I.J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44:522–53, 1938.

[Shp93] S.V. Shpectorov. On scale embeddings of graphs into hypercubes. *European J. Combin.*, 14:117–130, 1993.

[Siv02] D. Sivakumar. Algorithmic derandomization from complexity theory. In *Proc. 34th Annu. ACM Sympos. Theory Comput.*, pages 619–626, 2002.

[Tre01] L. Trevisan. When Hamming meets Euclid: The approximability of geometric TSP and MST. *SIAM J. Comput.*, 30:475–485, 2001.

[TZ01] M. Thorup and U. Zwick. Approximate distance oracles. In *Proc. 33rd Annu. ACM Sympos. Theory Comput.*, pages 183–192, 2001.