Weekplan: Approximate Distance Oracles

Philip Bille   Inge Li Gørtz   Christian Wulff-Nilsen*

References and Reading

[1] Approximate distance oracles, M. Thorup, U. Zwick, Journal of the ACM, 2005.

[2] Undirected single-source shortest paths with positive integer weights in linear time, M. Thorup, Journal of the ACM, 1999.

[3] Faster Algorithms for All-Pairs Approximate Shortest Paths in Undirected Graphs, S. Baswana and T. Kavitha, SIAM J. Computing, 2010

This week's curriculum is the following part of [1]:

- Section 1 until and including the first three paragraphs on Page 2 (you can ignore the last sentence of Theorem 1.1)
- Section 3 until and including Section 3.4. The $O(n^2)$ preprocessing time bound for metric spaces is only cursory

You should know that the results extend from metric spaces to graphs (see Theorem 1.1) and that the analysis for bounding space, stretch, and query time is the same for graphs as for metric spaces; only the analysis of preprocessing time differs. The lecture will focus on graphs only.

1 Exact Distance Oracle

Given an undirected graph G with non-negative edge weights, we want a data structure that answers queries for $\delta(u, v)$ for vertices u, v, where $\delta(u, v)$ is the shortest path distance in G between u and v.

1 Exercise. Consider trivial solutions:

- With no preprocessing of the graph, what query-time can you obtain? How much space is needed?
- Can this be done in $O(n^2)$ space and $O(1)$ query time? What is your preprocessing time?

*Based in part on the Weekplan from 2024 by Eva Rotenberg

2 Shortest Path Metric

In the lecture, we relied on the triangle inequality for shortest path distances.

2 Exercise. Prove that this inequality indeed holds for shortest path distances, i.e., show that for all $u, v, w \in V$,

$$\delta(u, w) \leq \delta(u, v) + \delta(v, w)$$

3 Specialized Oracle of Thorup and Zwick

3 Exercise We showed that the specialized approximate distance oracle presented in the lecture requires expected space $O(n^{3/2})$.

- What is its query-time?
- Show that its stretch is 3.

4 Exercise One version of Markov's inequality states that for any non-negative random variable X and for any $t > 0$,

$$\mathrm{P}[X \geq t] \leq \frac{E[X]}{t}$$

- With this inequality, show how to modify the construction of the specialized approximate distance oracle to get $O(n^{3/2})$ *worst-case* space.

4 General Oracle of Thorup and Zwick

5 Exercise The approximate distance oracle of [1] presented in this lecture does not work if G is a weighted *directed* graph. Point to where the argument breaks down.

6 Exercise Show that the query algorithm terminates no later than at index $k - 1$.

7 Exercise Show that the two distances, whose sum is returned in the last line of the query algorithm, are stored by the algorithm.

8 Exercise Show how to achieve $O(k)$ time for the query algorithm of Thorup and Zwick. You may rely on the existence of a hash table with $O(1)$ access time and $O(s)$ space where s is the number of elements stored.

As mentioned in the lecture, the Thorup-Zwick oracle can be constructed in $O(kmn^{1/k})$ time. In the next exercise, we will not prove this but only cover a small part of the construction needed to achieve this time bound.

We assume non-negative integer edge weights in the following exercise since this is needed to achieve the bound. This can be extended to arbitrary non-negative edge weights at the cost of an extra $\log n$-factor in the construction time.

9 Exercise Assume there is an $O(m)$ time algorithm for calculating the single-source-shortest path tree of a graph (Thorup [2]).

- For any $i < k$, show how to compute $p_i(v)$ for all $v \in V$ in $O(m)$ total time (not $O(m)$ time per v)

5 Modelling

10 Exercise In [1], the authors state that "the US road network is a planar graph". Indeed, a planar undirected weighted graph models a large road network. Can you think of different ways of modelling a large road network as a graph? What are their advantages and disadvantages?

6 Restricted Oracle

11 Exercise Let $S \subset V$ be a subset of vertices. Assume we only want to answer approximate distance-queries in G between vertices in S.

- For which vertices $v \in V$ is the bunch $B(v)$ needed in order to answer such queries?
- What is the space for this "restricted" oracle?
- Using a different data structure, can you remove the dependency on $n = |V|$ from the space bound?

7 Spanners

In the following, δ_H denotes the shortest path distance function for a graph H.

Let $G = (V, E)$ be an undirected graph with non-negative edge weights and let $t \geq 1$. A t-spanner of G is a subgraph $S = (V, E')$ such that for all $u, v \in V$,

$$\delta_S(u, v) \leq t\delta_G(u, v).$$

Fact I: for any integer $k \geq 1$ and any graph G with m edges and n vertices, G contains a $(2k-1)$-spanner S with $O(kn^{1+1/k})$ edges, and S can be found in $O(km + n)$ time.

12 Exercise Let $\varepsilon > 0$ be a given constant. Combine the above fact with the approximate distance oracle of Thorup and Zwick to obtain an approximate distance oracle with $O(1)$ stretch, $O(1)$ query time, $O(n^{1+\varepsilon})$ space, and $O(m + n^{1+\varepsilon})$ construction time.