Dr. Alexander Munteanu
Dennis Köhn, Dr. Vibha Sahlot

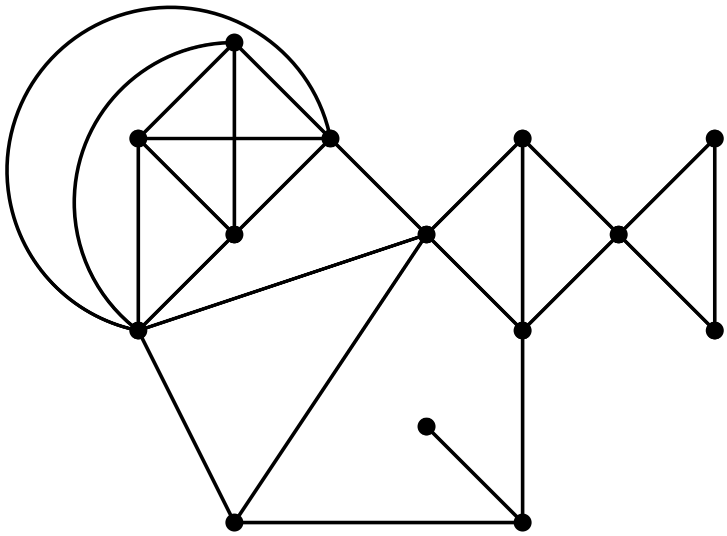
Winter term 2024/2025

5. Home Exercises

Efficient Algorithms

Ex. 1 *(4 points)*

Consider the graph below. Decompose the graph into its strongly connected components, i.e., find all its 1-strongly connected components, its 2-strongly connected components etc. What is the induced subgraph with the highest strong connectivity? Explain your answers.



Ex. 2 *(4+4 points)*

Recall the definition of a (k, ε)-coreset for the k-median problem. That is, given a weighted set $P \subseteq \mathbb{R}^d$, a weighted set $S \subset \mathbb{R}^d$ is a (k, ε)-coreset of P for k-median if for every set C of k centers the following holds:

$$(1 - \varepsilon) \cdot \mathrm{cost}(P, C) \leq \mathrm{cost}(S, C) \leq (1 + \varepsilon) \cdot \mathrm{cost}(P, C).$$

In the lecture, we constructed coresets in the streaming setting using the merge and reduce technique. Here, we required the coresets to be mergeable, i.e., that given two (k, ε)-coresets S_1, S_2 for two sets P_1, P_2 we have that $S_1 \cup S_2$ is a (k, ε)-coreset for $P_1 \cup P_2$.

In the following we want to prove that the coreset definition satisfies mergeability and we further look at a variant of the k-median problem, where this is not the case.

a) Given two disjoint sets $P_1, P_2 \subseteq \mathbb{R}^d$. Assume that S_1 with $w_1 : S_1 \to \mathbb{R}$ and S_2 with $w_2 : S_2 \to \mathbb{R}$ are (k, ε)-coresets for P_1 and P_2, respectively. Show that $S_1 \cup S_2$ with $(w_1 + w_2) : S_1 \cup S_2 \to \mathbb{R}$, $(w_1 + w_2)(x) := w_1(x) + w_2(x)$ is a (k, ε)-coreset for $P_1 \cup P_2$.

b) Consider the *capacitated* k-median problem. Here an additional parameter b is given that bounds the size of the clusters. That is, in the capacitated k-median problem we are given a set $P \subseteq \mathbb{R}^d$, $k \in \mathbb{N}$ and $b \in \mathbb{N}$ and we want to find k centers $c_1, \ldots, c_k$ and clusters $C_1, \ldots C_k$ with $\dot{\bigcup}_{i=1}^k C_i = P$ s.t. $\sum_{i=1}^k \operatorname{cost}(c_i, C_i)$ is minimized and additionally such that we have for each cluster that $|C_i| \leq b$.

   We again say that a set S with $w : S \to \mathbb{R}$ is a (k, ε)-coresets for a set $P \subseteq \mathbb{R}^d$ if for every feasible solution the costs on S and on P differ by at most a factor of $(1+\varepsilon)$.

   Show that for the capacitated k-median problem such coresets are not mergeable.

Ex. 3 *(4 + 2 + 2 points)*

Recall the two-dimensional 1-Center Problem, where we are given a set $P = \{p_1, \ldots, p_n\}$ of points in $\mathbb{R}^2$ and we want to compute the disk of smallest radius containing P. The center of this disk is called the 1-Center of P. We say that a set S is an ϵ-coreset for the 1-Center Problem if for every choice of a center c we have

$$(1-\epsilon)\operatorname{cost}(P, c) \leq \operatorname{cost}(S, c) \leq (1+\epsilon)\operatorname{cost}(P, c),$$

where $\operatorname{cost}(P, c) = \max_{p \in P} ||p - c||$.

a) Design an algorithm that constructs in $O(n)$ time an ϵ-coreset of size $O(1/\epsilon^2)$ for the 1-Center Problem.

b) Analyse the running time of your algorithm.

c) Prove the correctness of your algorithm.