Problem Set 2

Please hand in your solutions for this problem set via email (roesner@cs.uni-bonn.de) or personally at Room 2.060 until Tuesday, 23th of October.

Problem 1
Recall that Gonzalez’ algorithm computes a sequence of centers \( c_1, c_2, \ldots \), which adds one additional center in each iteration. This way the algorithm not only computes a solution with \( k \) clusters, but also implicitly computes for each \( 1 \leq k' \leq k \) an additional clustering with \( k' \) clusters. If we set \( k = |P| \) this yields an incremental clustering.

- Show with an example that these incremental clusterings computed with Gonzalez’ algorithms are not necessarily hierarchical.

Problem 2
Incremental/hierarchical clusterings compute a \( k \)-clustering for every \( k \in [|P|] \). If we want to compare two incremental/hierarchical clusterings, one of them might have the better clustering for some \( k \in [|P|] \) while the other might have the better clustering for a different \( k' \in [|P|] \).

- Give an example of a \( k \)-center problem, where no incremental clustering has an optimal solution for all \( k' \in [|P|] \).
- Give an example of a \( k \)-center problem, where no hierarchical clustering has an optimal solution for all \( k' \in [|P|] \).
- Show that for every incremental/hierarchical clustering, in some instances of the \( k \)-center problem, there exists another incremental/hierarchical clustering that has a truly better clustering for some \( k' \in [|P|] \).

Problem 3
Given a set of elements \( U = \{1, 2, \ldots, n\} \) and a collection of \( m \) subsets \( U_i \subseteq U \ (1 \leq i \leq m) \) together with \( k \in \{1, \ldots, m\} \), the Set Cover problem asks for some the question if there exists a sub collection of at most \( k \) of these subsets, whose union contains every element of \( U \). Formally, the Set Cover problem asks to decide if there exists a set \( I \subseteq \{1, \ldots, m\} \) with \( |I| \leq k \) such that \( \bigcup_{i \in I} U_i = U \). The Set Cover problem is known to be an NP-hard problem.

- Use the Set Cover problem to show that the \( k \)-supplier problem is NP-hard.
- Furthermore show that it is NP-hard to compute an \( \alpha \)-approximation for the \( k \)-supplier problem for any \( \alpha < 3 \).