Due Date: December 3, 2015

Problem 1: As in the n-body problem, suppose we are given a collection of n points $P = \{p_1, \ldots, p_n\} \subset \mathbb{R}^2$, a real number $\delta > 0$ and a family $\Phi = \{(A_1, B_1), \ldots, (A_s, B_s)\}$ such that

(a) For each k, A_k and B_k are δ-separated,

(b) For any pair $p_i \neq p_j$, there is a unique k such that $p_i \in A_k$ and $p_j \in B_k$ or $p_i \in B_k$ and $p_j \in A_k$, and

(c) $S = O(n/\delta^2)$.

Build a weighted (undirected) graph $G = (V, E)$ where $V = P$ and for every pair (A_i, B_i) in the family we add one edge $e = (a_i, b_i)$, for some $a_i \in A_i$ and $b_i \in B_i$, to E with its weight being $w(e) = \|a_i - b_i\|$, the Euclidean distance between a_i and b_i.

(a) Prove that G is connected.

(b) For any $p_i, p_j \in P$, let $d_G(p_i, p_j)$ be the cost of the shortest path in G between p_i, p_j. Prove that for all $\epsilon > 0$ we can find some $\delta > 0$ such that $d_G(p_i, p_j) \leq (1 + \epsilon)\|p_i - p_j\|_2$.

(**Hint:** Given a pair of points p_i, p_j, there is a unique pair (A_k, B_k) that contains this pair. G contains the edge (a_k, b_k). Using induction show that G has a path from p_i to p_j containing the edge (a_k, b_k) of cost at most $(1 + \epsilon)\|p_i - p_j\|$.)

Problem 2: Suppose we store a set of n items in a hash-table T of size m using a random hash function, i.e., each item is stored in a random location of T.

(a) What is the probability that a cell of T stores exactly k items? You can use the inequality $\binom{a}{b} \leq (ae/b)^b$.

(b) Suppose $n = m \ln m$. Show that the probability of a cell of T storing at least $2e \ln m$ items is at most $1/m^2$.

(c) Again suppose $n = m \ln m$. Argue that with probability at least $1 - 1/m$, all cells store $O(\log m)$ items.

Problem 3: The *sum-of-squares* problem is defined as follows: Given a set $A = (a_1, \ldots, a_n)$ and integers J, k, can A be partitioned into k disjoint sets $A_1 \ldots A_k$ so that $\sum_{i=0}^{k}(\sum_{a \in A_i} a)^2 \leq J$.

Prove that the sum-of-squares problem is NP-Complete.

Problem 4: A dominating set in a graph $G = (V, E)$ is a subset $S \subseteq V$ such that each vertex of V is either in S or has a neighbor in S. The *dominating-set* problem is defined as follows: Given a graph $G = (V, E)$ and integer k, does G contain a dominating set of size at most k.

Show that the problem is NP-Complete.

Problem 5: Given a graph $G = (V, E)$ we say it can be k-*colored* if there exists a way to assign one of k colors to each vertex such there is no edge $e = (v_i, v_j)$ such that v_i and v_j have the same color.

(a) Give an $O(|V| + |E|)$-time algorithm to determine whether a graph is 2-colorable, and if so return a valid coloring.

(b) Suppose the maximum degree of a vertex in G is k. Describe an $O(|V| + |E|)$ algorithm to color G with at most $k + 1$ colors.