CSC 2420 Fall 2012, Assignment 3
Due: December 14

1. We are given a degree bound d and query access to a partial table for a function $f : Q \to Q$; namely given $\{(x_1, f(x_1), \ldots, (x_n, f(x_n)\}$ we can access any $(x_i, f(x_i))$ in one query. We consider the following question for $d << n$:

   We want to test if the partial table f is equal to a degree d polynomial p or if it is "far-away" from any degree d polynomial where by far-away we mean that $f(x_i) = p(x_i)$ for at most $(1 - 2/d)n$ of the points given in the table. Provide a randomized 1-sided error algorithm that will make $O(d)$ queries, always returning p if it exists and with probability $\geq \delta$ will determine that f is far-away from any degree d polynomial. If neither condition is true, the algorithm can give any answer.

   Analyze the probability δ that can be achieved in terms of the number of queries used.

2. Consider a set of m linear equations defined over the field GF_2 (the field of 2 elements) by $Ax = b$ where A is an m by n matrix, x is a n-vector $\{x_i\}$ and b is an m vector of constant values $\{b_i\}$. (All entries are in GF_2 and operations are in GF_2.)

   (a) Use the method of conditional probabilities to show how to derive a deterministic algorithm that will find a solution for the x_i variables so as to yield a constant approximation on the number of satisfied equations. State the approximation ratio and argue why you obtain that ratio.

   (b) State the deterministic algorithm in terms that do not mention randomization in the same way that Johnson's algorithm is expressed (and was originally derived) without any reference to randomization.

   (c) Suppose every equation had at most 2 variables. Indicate (very briefly and that is a hint) if your approximation can be significantly improved.

3. This question concerns local search for the exact Max-2-Sat problem where the input is a CNF formula with exactly 2 literals per clause. The goal is to maximize the number (or the total weight) of clauses that can be satisfied by some truth assignment. As discussed in class, Khanna et al (link to paper will be provided) consider the locality gap achieved by oblivious and non-oblivious local search. (See pages 14-17 for Max-2-Sat and in particular, the proof of the $\frac{2}{3}$ locality gap for the 1-flip neighbourhood oblivious local search starts at the bottom of page 16.) As suggested by Khanna et al, one can also obtain a locality gap of $\frac{3}{4}$ by an oblivious local search algorithm that defines the neighbourhood of a solution (i.e. truth assignment) to include flipping any single variable and also flipping all variables.

   (a) Modify the $\frac{2}{3}$ locality gap result to obtain the improved locality gap for this larger neighbourhood.

   (b) OPEN Bonus problem

   What happens when you combine the non oblivious potential function with the extended neighbourhood given above? Can you show that this does or does not improve on the $\frac{3}{4}$ approximation obtained by the non-oblivious local search algorithm (using the standard Hamming 1-flip neighbourhood)? More generally, I believe it is an open problem if there is a "combinatorial algorithm" that beats the $\frac{3}{4}$ approximation for exact Max-2-Sat. (Recall that Johnson's algorithm and the naive randomized algorithm achieve this bound.)

4. Consider the weighted max-di-cut problem. That is, we are given an edge weighted directed graph $G = (V, E)$ with weight function $w : E \to \Re^{\geq 0}$ and the goal is find a subset of vertices $S \subseteq V$ so as to maximize $f(S) = \sum_{(u,v) \in E, u \in S, v \in V \setminus S} w(u, v)$.

   (a) Show that f is a non monotone submodular set function.

   (b) What is the expected approximation ratio for the naive randomized algorithm? That is, independently place each vertex in S with probability $1/2$.

   (c) Consider the de-randomization (by the method of conditional expectations) of the naive randomized algorithm. Express the de-randomized algorithm as a deterministic greedy algorithm.

5. Complete the proof of Lemma II.2 in the Buchbinder et al paper. (I have a link to the paper on the course web page.) That is, give the missing part of the proof when $a_i < b_i$.