Problem Set 5

CS 331H

Due **Thursday**, April 21

1. A *matching* on a graph is a set of edges such that no vertex is the endpoint of two different edges in the set. In the *maximum matching* problem, you would like to find a matching S with the largest possible size.

   A *maximum* matching is a matching of maximum size. A *maximal* matching is a matching S that cannot be extended by adding another edge: for every edge in the original graph, at least one endpoint is an endpoint of some edge in S.

   (a) If $S^* \subseteq E$ is a maximum matching and $S \subseteq E$ is a maximal matching, show that $|S| \geq |S^*|/2$.

   (b) Give a simple $O(m)$ greedy algorithm that gives a (1/2)-approximation for the maximum matching problem.

   (c) Given a maximal matching S, show how to construct a vertex cover of size $2|S|$.

   (d) Use the above to get an $O(m)$ time algorithm for a 2-approximation for *vertex cover*. How does this compare to the LP rounding we considered in class?

2. In the MAX-SAT problem, you have a collection of m clauses $C_1, \ldots, C_m$ on n variables $x_1, \ldots, x_n$. Each clause C_i is the OR of some positive literals $P_i \subset [n]$ and negative literals $N_i \subset [n]$; the value of the clause for a given assignment of variables is 1 if any $x_j = 1$ for $j \in P_i$ or if $x_j = 0$ for any $j \in N_i$.

   Additionally, suppose that we have a weight $w_i \geq 0$ associated with each clause. The goal of the MAX-SAT problem is to find an assignment that maximizes the sum of the weights of satisfied clauses.

   (a) Show that the following integer linear program gives the correct answer:

$$\begin{aligned}
&\max \sum w_i z_i \\
&\text{s.t. } z_i \le \sum_{j \in P_i} x_j + \sum_{j \in N_i} (1 - x_j) \quad \forall i \in [m] \\
&\quad z_i, x_j \in \{0, 1\}
\end{aligned}$$

(b) State the LP relaxation of the above integer linear program.

(c) Suppose you have (x, z) satisfying the constraints to the linear program. Suppose that each clause C_i involves at most 2 variables. Give a randomized rounding scheme such that, for each clause C_i,

$$\Pr[C_i \text{ satisfied}] \ge 3z_i/4$$

over the randomness in the rounding.

(d) Conclude with an expected 3/4 approximation to MAX-SAT on formulae with clauses of at most 2 variables.

(e) Demonstrate that the integrality gap for this problem is 4/3, using a formula with two variables.

(f) Now suppose every clause had at least 2 variables, rather than at most two variables. Show that randomly assigning values to variables gives an expected 3/4 approximation to MAX-SAT in this setting.

(g) Now give an expected 3/4 approximation to MAX-SAT for arbitrary formulae. **Hint:** show that the average of the above two methods will work for any formula; alternatively, show that one of the two will.

(h) (Optional) Try to give a polynomial time algorithm that will give a 3/4 approximation with high probability.