Problem Set 5

CS 331H

Due Thursday, April 27

1. Consider an interval scheduling problem where we have multiple machines, and each interval can specify which machines it can run on. That is, you have n jobs, and each job j is described by (s_j, f_j, M_j) where $M_j \subseteq [m]$ is a subset of machines and the set of intervals scheduled on each machine must be disjoint.

   (a) Show that determining whether at least k jobs can be scheduled is NP-hard via a reduction from maximum independent set. Hint: Associate each vertex with a machine, and each edge $e_i = (u_i, v_i) \in E$ with the job $(i, i+1, \{u_i, v_i\})$, and additionally for each vertex u create the job $(1, |E|+1, \{u\})$.

   (b) Show that the multiple-machine interval scheduling problem is NP-complete.

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:

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

   (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,

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

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) Separately, use a formula with two variables and four clauses to demonstrate that the integrality gap for this problem is 4/3.

(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.

3. In the BINPACKING problem, you have n items with sizes $s_1, \ldots, s_n$ in $(0, 1)$. You would like to pack them into the smallest possible number k of unit-sized bins, where the sum of the sizes of the items in each bin must be at most 1.

   (a) Show that it is NP-hard to compute an α-approximate answer to BINPACKING, for any $\alpha < 3/2$.

   Hint: show that it is NP-hard to determine if the answer is ≤ 2. Reduce from the PARTITION problem, where one is given n items $a_1, \ldots, a_n \in \mathbb{Z}$, and wants to know whether there exists a set S with

   $$\sum_{i \in S} a_i = \sum_{i \notin S} a_i.$$

   You may use that PARTITION is NP-hard.

   (b) Give a simple 2-approximation to BINPACKING.

   (c) [Optional.] Give an algorithm with a better approximation factor.