Problem Set 10

CS 331H

Due Monday, April 29

1. Set L_n be the set of n-state Turing machines over a binary alphabet. For each machine $x \in L_n$, let $T_x \in \mathbb{Z}^+ \cup \{\infty\}$ denote the number of steps x takes, when run on an initially blank tape, before halting. (If x never halts, set T_x to ∞.) As defined in class, $S(n) = \max_{T_i < \infty} T_i$ is the nth busy beaver number.

   We say that an n-state Turing machine x is a "busiest beaver" if $T_x = S(n)$. Consider the problem of `BusiestBeaver`(x) which takes a Turing machine x and returns `True` if x is a busiest beaver, and `False` otherwise.

   Prove that `BusiestBeaver` is uncomputable.

2. In the BinPacking problem, you have n items with positive integer sizes $s_1, \ldots, s_n$ and a bin size B. You would like to pack them into the smallest possible number k of bins, where the sum of the sizes of the items in each bin must be at most B; the output is this number k.

   (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 or ≥ 3. Reduce from the Partition problem considered in the last problem set.

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

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

3. In the MAX-SAT problem, you have a collection of m clauses $C_1, \dots, C_m$ on n variables $x_1, \dots, x_n$. Each clause C_i is the OR of some positive literals $P_i \subset [n]$ and a disjoint set of 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 $j \in N_i$.

   We also 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.

   In this problem we consider the special case of MAX-$\leq$ 2-SAT, where you are additionally guaranteed that every clause will involve only one or two literals.

   (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 of the linear program. 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. (Recall that we assume C_i involves at most two literals.)

   Hint: You may use the fact that, for any two real numbers $a, b \in [0, 1]$,

$$a + b - ab \geq \frac{3}{4} \min(a + b, 1).$$

   (d) Conclude with an expected 3/4 approximation algorithm to MAX-$\leq$ 2-SAT.

   (e) Use a formula with two variables and four clauses to demonstrate that the integrality gap for MAX-$\leq$ 2-SAT is 4/3.

   (Recall that the "integrality gap" is the ratio between the LP optimum and the integer LP optimum.)

[Everything below here is optional.]

(f) The MAX-$\geq$ 3-SAT problem is the same as MAX-SAT, except that every clause C_i involves ≥ 3 literals.

Give a simple randomized algorithm that produces an expected 7/8-approximation.

(g) Now give an expected 3/4 approximation to MAX-SAT for general formulae, by showing that the *average* performance of the part (d) and part (f) methods is good enough for every instance.