CS 466, Spring 2007
Timothy Chan

Assignment 5 (due July 25 Wed noon)

1. [*15 marks*] Although we have given an asymptotic PTAS for off-line bin packing, the on-line bin packing problem is of interest in various applications. In this question, you will obtain a better on-line result for a special case, with approximation factor better than the 17/10 bound for the first/best-fit algorithm mentioned in class. In our special case, we assume that *all items have values greater than* 0.25.

   (a) [*5 marks*] Let n_A be the number of items in the range $(1/2, 1]$ ("type-A items"), n_B be the number of items in the range $(1/3, 1/2]$ ("type-B items"), and n_C be the number of items in the range $(1/4, 1/3]$ ("type-C items").

   Give a simple online algorithm for bin packing that uses at most $n_A + n_B/2 + n_C/3 + 3$ bins.

   (b) [*5 marks*] For each item, define its *weight* to be 1 if it is of type A, 1/2 if it is of type B, and 1/3 if it is of type C.

   Suppose a subset of items has total value at most 1. Prove that their total weight must be at most 3/2.

   [Hint: one way is to do an exhaustive case analysis.]

   (c) [*5 marks*] Prove that the algorithm from part (a) has asymptotic approximation factor 3/2 for our special case.

   [2 bonus marks: give a rigorous proof that a generalization of the above online algorithm has asymptotic approximation factor strictly smaller than 17/10 for the case when all items have values greater than 0.1. 2 more bonus marks (difficult!): do the same for the general case without any assumptions.]

2. [*19 marks*] In this question, you will explore another "3-clustering"-like problem: Given an undirected graph $G = (V, E)$ with n vertices and three special vertices $s_1, s_2, s_3 \in V$, we want to partition V into three subsets V_1, V_2, V_3 such that $s_i \in V_i$ ($i \in \{1, 2, 3\}$), minimizing

$$c(V_1, V_2, V_3) \ := \ \text{number of edges } uv \in E \text{ such that } u \in V_i \text{ and } v \in V_j \text{ with } i \neq j.$$

   In other words, $c(V_1, V_2, V_3)$ is the number of edges *cut* by the partition. As you might guess, this problem is NP-hard.

   (a) [*3 marks*] Consider the following optimization problem, where for each vertex $v \in V$, we create 3 real variables x_v, y_v, z_v:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2} \sum_{uv \in E} (|x_u - x_v| + |y_u - y_v| + |z_u - z_v|) \\ \text{subject to} & x_v + y_v + z_v = 1, \ \ x_v, y_v, z_v \geq 0 \qquad \forall v \in V \\ & (x_{s_1}, y_{s_1}, z_{s_1}) = (1, 0, 0), \ \ (x_{s_2}, y_{s_2}, z_{s_2}) = (0, 1, 0), \ \ (x_{s_3}, y_{s_3}, z_{s_3}) = (0, 0, 1) \end{array}$$

Show that the above is a linear programming problem. [Hint: this is not completely obvious because of the absolute values, but they can be avoided by introducing extra variables. . .]

(b) [*3 marks*] Let c_{LP} denote the optimal value of the linear program. Let c^* denote the optimal value of our 3-clustering problem.

Show how a solution (V_1, V_2, V_3) to the 3-clustering problem can be converted to a solution to the linear program, and argue that $c_{LP} \leq c^*$.

(c) [*2 marks*] Consider the following randomized rounding scheme:

1. let x_v, y_v, z_v $(v \in V)$ be the solution to the LP
2. pick a random real number $\alpha \in (0, 1)$
3. let $X = \sum_{uv \in E} |x_u - x_v|$, $Y = \sum_{uv \in E} |y_u - y_v|$, and $Z = \sum_{uv \in E} |z_u - z_v|$
3. if $Z \leq X$ and $Z \leq Y$ then {
4.   for each $v \in V$ do
5.   if $x_v \geq \alpha$ then put v in V_1
6.   else if $y_v \geq \alpha$ then put v in V_2
7.   else put v in V_3
    }
8. else if $X \leq Y$ and $X \leq Z$ then
9.   similar but swap x's with z's and V_1 with V_3
10. else similar but swap y's with z's and V_2 with V_3

Show that the partition (V_1, V_2, V_3) produced by the algorithm is a feasible solution to the 3-clustering problem (i.e., check that $s_i \in V_i$).

(d) [*4 marks*] Assume line 3 holds. Show that the probability that a fixed edge uv is cut by the partition produced by part (c) is at most $|x_u - x_v| + |y_u - y_v|$.

(e) [*4 marks*] Assume line 3 holds. Show that $E[c(V_1, V_2, V_3)] \leq X + Y \leq \frac{2}{3}(X + Y + Z)$.

(f) [*3 marks*] Conclude that there is a polynomial-time algorithm for our 3-clustering problem with expected approximation factor at most 4/3.

3. [*6 marks*] Consider a variant of the lost cow problem, where there are two fences instead of one. One fence is known to be to the left and the other fence is to the right. The initial distance to either fence is unknown. The objective is to find *both* fences. Give a deterministic online algorithm with competitive ratio at most 2.

[Note: this is supposed to be easy. First determine the off-line cost. . .]