CS 466/666, Spring 2015
Timothy Chan

Assignment 4 (due July 22 Wednesday 5pm)

Please read `http://www.student.cs.uwaterloo.ca/~cs466/policies.html` first for general instructions.

1. [*12 marks*] Given a set S of n unit squares in 2D, we want to assign a color to each square so that no two intersecting squares have the same color, while minimizing the number of colors used.

   Design an approximation algorithm for the problem with approximation factor at most 4, by using a grid approach. (As in class, consider the centers of the squares and assign centers to grid cells.)

2. [*12 marks*] We have stated in class that *any* on-line algorithm for bin packing can achieve asymptotic approximation factor at least 1.54. In this question, you will prove a weaker lower bound of 4/3.

   Let A be an arbitrary (deterministic) on-line algorithm. Let k be an arbitrarily large even integer.

   (a) [*4 marks*] Consider an input, consisting of k copies of 0.49 followed by k copies of 0.01. Suppose that just after the k copies of 0.49 are read, algorithm A has created p bins with two items and q bins with one item (with $2p + q = k$). Prove that on this input, algorithm A has approximation factor at least $2(1 - p/k)$.

   (b) [*4 marks*] Consider an input, consisting of k copies of 0.49 followed by k copies of 0.51. Suppose that just after the k copies of 0.49 are read, algorithm A has created p bins with two items and q bins with one item. Prove that on this input, algorithm A has approximation factor at least $1 + p/k$.

   (c) [*4 marks*] Use an adversary argument to conclude that the asymptotic approximation factor of algorithm A must be at least 4/3.

3. [*18 marks*] Consider the following job scheduling problem: we have $n/3$ processors available, each capable of handling 3 jobs; we would like to assign jobs to processors so as to minimize the overall (maximum) completion time. More precisely, given a sequence of n positive numbers $s_1, \ldots, s_n$ (where n is divisible by 3), we would like to partition the sequence into subsets $S_1, \ldots, S_{n/3}$, each containing 3 numbers, so as to minimize the quantity $\max_{j=1,\ldots,n/3} \sum_{s \in S_j} s$.

   (a) [*4 marks*] Show that there is an *online* approximation algorithm with approximation factor at most 3. [Hint: the algorithm is *really* simple. For the analysis, define $M = \max\{s_1, \ldots, s_n\}$ and compare with the optimal value...]

(b) [*4 marks*] Let k be a fixed constant and let A be a fixed set of k elements. Show that in the special case where all s_i's come from the set A (duplicates are allowed), the (offline) problem can be solved exactly in polynomial time.

(c) [*10 marks*] Now design and analyze a polynomial-time approximation scheme (PTAS) for the general (offline) problem. [Hint: use a rounding technique (but it should be simpler than the bin-packing PTAS from class).]

4. [*18 marks*] Consider the following maximization problem: We are given a collection of m subsets $A_1, \ldots, A_m$ of $\{1, \ldots, n\}$. We are also given an integer k. We would like to find a subset $S \subseteq \{1, \ldots, n\}$ of size k, maximizing the following value:

$$f(S) = \text{the number of subsets } A_j \text{ that } S \cap A_j \neq \emptyset.$$

(In other words, we want to choose k elements, maximizing the number of subsets intersected by these elements. For example: if $A_1 = \{1, 3, 4\}$, $A_2 = \{2, 3, 5\}$, $A_3 = \{3, 4, 5\}$, $A_4 = \{1, 2, 5\}$, $A_5 = \{6, 7, 8\}$ and $k = 2$, then an optimal solution is $S = \{3, 5\}$ with $f(S) = 4$.)

(a) [*5 marks*] Consider the following linear program:

$$\begin{array}{ll} \text{maximize} & z_1 + \cdots + z_m \\ \text{such that} & y_1 + \cdots + y_n \ = \ k \\ & \sum_{i \in A_j} y_i \ \geq \ z_j \qquad (j = 1, \ldots, m) \\ & 0 \leq y_1, \ldots, y_n, z_1, \ldots, z_m \leq 1. \end{array}$$

Let z_{LP} be the optimal (real) value of the linear program, and let z^* be the value of the optimal solution to the original problem. Explain the connection between this linear program and the original problem, and prove that $z_{LP} \geq z^*$.

(b) [*5 marks*] Consider the following randomized algorithm:

let $y_1, \ldots, y_n, z_1, \ldots, z_m$ be the solution to the linear program
for $\ell = 1, \ldots, k$ {
  randomly choose an element from $\{1, \ldots, n\}$ where
    i is chosen with probability $y_i/(y_1 + \cdots + y_n)$
  output this element
}

Let S be the subset of elements outputted by the algorithm. Prove that the probability that S does not intersect A_j (i.e., $S \cap A_j \neq \emptyset$) is at most $e^{-\sum_{i \in A_j} y_j}$.
[Hint: the inequality $1 - x \leq e^{-x}$ might again come in handy...]

(c) [*4 marks*] Continuing part (b), prove that the probability that S intersects A_j is at least $(1 - e^{-1})z_j$.
[Hint: you may use the fact that $(1 - e^{-x})/x$ is a decreasing function for $x \in [0, 1]$.]

(d) [*4 marks*] Prove that the above algorithm has expected approximation factor at least $1 - e^{-1} \approx 0.632$.