Midterm Exam

Approximation Algorithms: CS583, Spring 2026
Tuesday, March 10, 2026

Name:	
NetID:	

Instructions and Policy:

- This is a closed-book exam but you are allowed a 1 page cheat sheet.
- **Please read the entire exam before writing anything.** There are five numbered problems. All of them are of equal value.
- **You have 120 minutes (2 hours) for the exam. I expect you to do four of the given five problems but do as many as you can. Read all problems once before starting.**
- Write your netid clearly on each page of your exam. Start each problem on a fresh page.
- You are allowed to use existing algorithms from the lectures/notes as a black box but it is important to explain your reduction and how to use the known algorithm and its properties to derive properties for your algorithm.

Probabilistic Inequalities

- Markov's inequality: For a *non-negative* random variable X and $t > 0$, $\mathrm{P}[X \geq t] \leq \mathrm{E}[X]/t$.
- Chebyshev's inequality. For a random variable X with finite variance, $\mathrm{P}[|X - \mathrm{E}[X]| \geq a] \leq \mathrm{Var}[X]/a^2$.
- Chernoff bound for sum of *non-negative* bounded random variables. Let $X_1, \ldots, X_k$ be k independent binary random variables such that, for each $i \in [1, k]$, $\mathrm{E}[X_i] = \Pr[X_i = 1] = p_i$. Let $X = \sum_{i=1}^{k} X_i$. Then $\mathrm{E}[X] = \sum_i p_i$.
  - Upper tail bound: For any $\mu \geq \mathrm{E}[X]$ and any $\delta > 0$,
  $$\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \right)^{\mu}$$
  - Lower tail bound: For any $0 < \mu \leq \mathrm{E}[X]$ and any $0 < \delta < 1$,
  $$\Pr[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}} \right)^{\mu}$$

  The above bounds can be simplified when $0 \leq \delta < 1$, as follows:
  $$\Pr[X \geq (1+\delta)\mu] \leq e^{\frac{-\delta^2 \mu}{3}} \text{ and } \Pr[X \leq (1-\delta)\mu] \leq e^{\frac{-\delta^2 \mu}{2}}$$
- Chernoff bound for sum of bounded random variables. Let $X_1, \ldots, X_k$ be k independent random variables such that, for each $i \in [1, k]$, $X_i \in [-1, 1]$. Let $X = \sum_{i=1}^{k} X_i$. For any $a > 0$,
$$\Pr[|X - \mathrm{E}[X]| \geq a] \leq 2 \exp\left(\frac{-a^2}{2k} \right).$$

Problem 1. Consider the following variant of Dominating Set. We are given a graph $G = (V, E)$ and for each vertex v a non-negative integer requirement $r_v \geq 1$. Each vertex v has a non-negative weight w_v. The goal is to choose an integer x_u for each $u \in V$ to satisfy the covering requirement for every $v \in V$: $\sum_{u \in N^+(v)} x_u \geq r_v$. Here $N^+(v) = \{u \mid (u, v) \in E\} \cup \{v\}$ is the extended neighborhood of v. The objective is to minimize the cost $\sum_{u \in V} w_u x_u$. Describe an $O(\log n)$ approximation for this problem where $n = |V|$.

Problem 2. A graph $G = (V, E)$ is k-degenerate if every subgraph of G has a vertex of degree k or less.

- Describe a simple greedy algorithm and show that it obtains a $1/k$ approximation for the maximum independent set problem in G.
- Obtain an $\Omega(1/k)$-approximation for maximum weight independent set in G.

Problem 3. In the 2-dimensional Knapsack Cover problem we are given n vectors $v_1, v_2, \ldots, v_n$ where each $v_i = (x_i, y_i)$ is a 2-dimensional non-negative vector, and also a 2-dimensional knapsack capacity vector (B_1, B_2). Each v_i has a cost c_i. Assume all input numbers are non-negative integers. The goal is to find a minimum cost subset of the given vectors that cover the given 2-d knapsack. A set of vectors S cover the knapsack if $\sum_{i \in S} x_i \geq B_1$ and $\sum_{i \in S} y_i \geq B_2$. Describe a PTAS for this problem. A constant factor approximation will get you partial credit. *Hint:* You saw a related problem in the homework. Use similar ideas but for cost minimization.

Problem 4. Consider the unrelated machine scheduling problem. We are given n jobs and m machines. Each job i has a processing time of $p_{i,j} \in [0, \infty)$ on machine j. We want to schedule the jobs to minimize the makespan (maximum load over machines).

- Write an LP relaxation for this problem using allocation variables $x_{i,j}$ which indicate whether job i is assigned to machine j. Explain why the straightforward relaxation suffers from a large integrality gap.
- Suppose we guessed the optimum value to be L. How would you modify the LP relaxation to avoid the bad integrality gap?
- What approximation guarantee can you show if you use a simple randomized rounding strategy? Be careful to explain whether the ratio depends on n, m, or both.

Problem 5. Consider the American bin packing problem. We are given n items each with a size $s_i \in (0, 1]$. Each item has a color: red, blue, or white. We want to pack the items into as few bins of size 1 as possible but we cannot pack a red item and a blue item into the same bin. Describe and analyze a simple constant factor approximation for this problem. The smaller the ratio the better. Note that it is easy to obtain a 6-approximation so you need to do better than that.