Midterm Exam, Algorithms II 2023-2024

Do not turn the page before the start of the exam. This document is double-sided, has 6 pages, the last ones possibly blank. Do not unstaple.

- The exam consists of three parts. The first part consists of multiple-choice questions, the second part consists of a short open question, and the last part consists of three open-ended questions.
- For the open-ended questions, your explanations should be clear enough and in sufficient detail that a fellow student can understand them. In particular, do not only give pseudocode without explanations. A good guideline is that a description of an algorithm should be such that a fellow student can easily implement the algorithm following the description.
- You are allowed to refer to material covered in the lectures including algorithms and theorems (without reproving them). You are however *not* allowed to simply refer to material covered in exercises/homework.

Good luck!

Problem 1: Multiple Choice Questions (24 points)

For each question, select the correct alternative. Note that each question has **exactly one** correct answer. Wrong answers are not penalized with negative points.

1a. Matroids *(8 points).* Consider the ground set $E = \{a, b, c, d\}$. Select a collection $\mathcal{I}$ of independent sets from below such that $(E, \mathcal{I})$ is a matroid.

A. $\{\{\}, \{a\}, \{b\}, \{c\}\ \{a, b\}, \{a, c\}, \{a, d\}\}$

B. $\{\{\}, \{a\}, \{b\}, \{c\}, \{a, b\}\}$

C. $\{\{\}, \{a\}, \{b\}, \{c\}\}$

D. $\{\{\}, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, c\}, \{b, c\}, \{b, d\}, \{a, b, c\}, \{b, c, d\}\}$

E. $\{\{\}, \{a\}, \{b\}, \{c\}, \{a, b, c\}\}$

1b. Vertex-Cover relaxation *(8 points).* Consider the minimum vertex cover problem, and consider the linear programming (LP) relaxation for vertex cover you saw in class. For a graph G, denote by $\mathrm{OPT}(G)$ the cost of an optimum vertex cover for G, and by $\mathrm{OPT}_{\mathrm{LP}}(G)$ the cost of an optimal solution of the LP relaxation. Which one of the following statements is true?

A. There is a graph G so that $\mathrm{OPT}(G) \geqslant 4 \cdot \mathrm{OPT}_{\mathrm{LP}}(G)$.

B. For all n-vertex graphs G, we have $\mathrm{OPT}_{LP}(G) \geqslant n/2$.

C. If an n-vertex graph G has $\mathrm{OPT}(G) = 3n/4$, then $\mathrm{OPT}_{\mathrm{LP}}(G) \geqslant 3n/4$.

D. For all graphs G, it holds that $\mathrm{OPT}(G) \geqslant \mathrm{OPT}_{\mathrm{LP}}(G)$.

E. There exists a graph such that $\mathrm{OPT}_{\mathrm{LP}}(G) > 2 \cdot \mathrm{OPT}(G)$.

1c. Weighted Majority *(8 points).* We apply the weighted majority algorithm to aggregate the (binary) answers of 17 experts. At every step, we *divide by* 2 the weights of those experts that provided a wrong answer. Assume that c experts always provide the correct answer. What is the smallest value of c for which the total number of mistakes we make is *at most* 1, independently of the answers given by the other $17 - c$ experts?

A. 3

B. 4

C. 5

D. 6

E. 7

Problem 2: Short Open Question (10 points)

Write the dual linear program of the following linear program. No explanation is needed for your answer.

$$\begin{aligned} \min \quad & 3x_1 + 5x_2 + x_3 \\ \text{s.t.} \quad & x_1 + x_2 + x_3 \geqslant 1 \\ & x_2 - 2x_3 \leqslant 3 \\ & x_1, x_2, x_3 \geqslant 0 \end{aligned}$$

Problem 3: Extreme Point Structure (22 points)

Given a graph $G = (V, E)$ with edge-weights $w : E \to \mathbb{R}$, consider the matching problem where we wish to select a matching of maximum weight consisting of exactly k edges. We can adapt the linear program seen in class to obtain the following relaxation:

$$\begin{aligned} \text{maximize} \quad & \sum_{e \in E} x_e \cdot w(e) & \\ \text{subject to} \quad & \sum_{e \in \delta(v)} x_e \leqslant 1 & \forall v \in V \\ & \sum_{e \in E} x_e = k & \\ & x_e \geqslant 0 & \forall e \in E \end{aligned}$$

where $\delta(v)$ denotes the set of edges incident to vertex v.

Let x^* be an extreme point of the above linear program. Consider the graph G' which is the subgraph of G that contains only the edges with $x^*_e > 0$. **Prove that G' contains no cycles of even length.** A cycle has an even length if it has an even number of edges.

Problem 4: Prize-Collecting Vertex Cover and Duality (22 points)

The prize-collecting vertex cover problem is a generalization of vertex cover in which we are not obligated to cover all edges, but must pay a penalty for those left uncovered. A formal definition is as follows:

Input: An undirected graph $G = (V, E)$ with a penalty $p_e \geqslant 0$ for every edge $e \in E$.

Output: A subset $C \subseteq V$ of vertices so as to minimize $|C| + \sum_{e \in E : e \cap C = \emptyset} p_e$.

To formulate a linear programming relaxation, we associate a variable x_v for every vertex $v \in V$, and a variable z_e for every edge $e \in E$. The intended meaning of these variables is that x_v indicates whether $v \in C$ and z_e indicates whether e pays a penalty, i.e., is not covered by C. We then arrive at the following linear programming (LP) relaxation and its dual:

(Primal) LP Relaxation

$$\begin{array}{lll} \textbf{minimize} & \sum_{v \in V} x_v + \sum_{e \in E} p_e \cdot z_e & \\ \textbf{subject to} & x_u + x_v + z_e \geqslant 1 & \text{for } e = \{u, v\} \in E \\ & x_v \geqslant 0 & \text{for } v \in V \\ & z_e \geqslant 0 & \text{for } e \in E \end{array}$$

(Dual)

$$\begin{array}{lll} \textbf{maximize} & \sum_{e \in E} y_e & \\ \textbf{subject to} & \sum_{e \in \delta(v)} y_e \leqslant 1 & \text{for } v \in V \\ & y_e \leqslant p_e & \text{for } e \in E \\ & y_e \geqslant 0 & \text{for } e \in E \end{array}$$

Recall that $\delta(v)$ denotes the set of edges incident to vertex $v \in V$.

We will analyze a simple and very fast primal-dual algorithm for the prize-collecting vertex cover problem. The algorithm maintains a feasible dual solution y initially set to $y_e = 0$ for every $e \in E$. It then iteratively improves the dual solution until every edge $e \in E$ not covered by the set $C = \{v \in V : \sum_{e \in \delta(v)} y_e = 1\}$ corresponds to a tight constraint $y_e = p_e$. Note that C consists of those vertices whose constraints in the dual are tight and the algorithm only stops when the edges not covered by C correspond to tight dual constraints. The formal description of the algorithm is as follows:

1) Initialize the dual solution y to be $y_e = 0$ for every $e \in E$.
2) While there is an edge e with $y_e < p_e$ and that is *not* covered by $C = \{v \in V : \sum_{e \in \delta(v)} y_e = 1\}$, i.e., $e \cap C = \emptyset$:
   - Increase y_e until one of the dual constraints (corresponding to u, v or e) becomes tight.
3) Return $C = \{v \in V : \sum_{e \in \delta(v)} y_e = 1\}$.

Prove that the primal-dual algorithm has an approximation guarantee of 2. That is, show that the returned set C has value

$$|C| + \sum_{e \in E | e \cap C = \emptyset} p_e$$

at most twice the value of an optimal solution. **Partial credits will be given to solutions that bound the approximation guarantee by 3.**

Problem 5: Edge-Disjoint Spanning Trees (22 points)

Given a graph $G = (V, E)$, **design and analyze a polynomial-time algorithm** that does the following: Construct three spanning trees of G that share no edges, or report that this task is impossible (i.e., that G does not have three edge-disjoint spanning trees).