ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Midterm Exam, Advanced Algorithms 2018-2019

- You are only allowed to have a handwritten A4 page written on both sides.
- Communication, calculators, cell phones, computers, etc... are not allowed.
- Your explanations should be clear enough and in sufficient detail that a fellow student can understand them. In particular, do not only give pseudo-code 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 lecture notes** including theorems without reproving them. You may also refer to statements given in the exercise sheets and the homework sheet. You are however *not* allowed to refer to material from any specific solution to these exercises.
- **Do not touch until the start of the exam.**

Good luck!

Name: _____ **N° Sciper:** _____

Problem 1 / 20 points	Problem 2 / 14 points	Problem 3 / 22 points	Problem 4 / 22 points	Problem 5 / 22 points

Total / 100

1 *(20 pts)* **Duality of linear programming.** Consider the following linear program:

$$\begin{aligned} \text{Minimize} \quad & 6x_1 + 15x_2 + 8x_3 \\ \text{Subject to} \quad & 2x_1 + 6x_2 + x_3 \geq 3 \\ & x_1 + 2x_2 + 3x_3 \geq 4 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Write down its dual and the complementarity slackness conditions.

Solution: You will find the definition of the dual linear program in lecture notes 5. The dual problem of the abovementioned primal is the following:

$$\begin{aligned} \text{Maximize} \quad & 3y_1 + 4y_2 \\ \text{Subject to} \quad & 2y_1 + \ y_2 \leq 6 \\ & 6y_1 + 2y_2 \leq 15 \\ & \ y_1 + 3y_2 \leq 8 \\ & y_1, y_2 \geq 0 \end{aligned}$$

In lecture notes 5, proposition 1.5 you will find the definition of complementary slackness. In this particular case it is equivalent to saying that if (x_1, x_2, x_3) and (y_1, y_2) are feasible solutions to the primal and dual correspondingly, then :

$$\left.\begin{array}{ll} (x_1, x_2, x_3) & \text{is an optimal solution of the primal} \\ (y_1, y_2) & \text{is an optimal solution of the dual} \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{ll} x_1 > 0 & \Rightarrow 2y_1 + \ y_2 = 6 \\ x_2 > 0 & \Rightarrow 6y_1 + \ y_2 = 15 \\ x_3 > 0 & \Rightarrow \ y_1 + 3y_2 = 8 \\ y_1 > 0 & \Rightarrow 2x_1 + 6x_2 + \ x_3 = 3 \\ y_2 > 0 & \Rightarrow \ x_1 + 2x_2 + 3x_3 = 4 \end{array}\right.$$

2 *(14 pts)* **Karger's randomized algorithm for min-cut.** In class, we saw Karger's beautiful randomized algorithm for finding a minimum cut in an undirected graph $G = (V, E)$. Recall that his algorithm works by repeatedly contracting a randomly selected edge until the graph only consists of two vertices which define the returned cut. For general graphs, we showed that the returned cut is a minimum cut with probability at least $1/\binom{n}{2}$.

In this problem, we are going to analyze the algorithm in the special case when the input graph is a tree. Specifically, you should show that if the input graph $G = (V, E)$ is a spanning tree, then Karger's algorithm returns a minimum cut with probability 1.

(In this problem you are asked to show that Karger's min-cut algorithm returns a minimum cut with probability 1 *if the input graph is a spanning tree. Recall that you are allowed to refer to material covered in the lecture notes.)*

Solution: Firstly notice that for a tree the minimum cut is of size one. It is enough to consider $S = \{v\}$ where v is any leaf node. As such it has only one edge connecting it to the rest of the graph and the cut size is 1. Since a tree is connected there cannot be a 0-weight cut, so the minimum cut indeed contains only one edge.

Now let's prove that Karger's algorithm always outputs a cut with only one edge. The starting graph $G = (V, E)$ is a tree, so it satisfies $\#E = \#V - 1$. At every step Karger's algorithm decreases the number of nodes by 1, and the number of edges by ≥ 1 (in case of multiedges). At the end, since the starting graph is connected, we will end with ≥ 1 edges between the two final nodes. Since the algorithm runs for $\#V - 2$ steps, this means that it had to remove exactly 1 edge at every step, and finish with exactly 1 edge between the two final nodes in order to satisfy the above inequalities. It follows that the obtained cut will have weight 1, i.e. it will be a minimum cut. Note also that at every step of the algorithm the obtained graph is still a tree, i.e. Karger's algorithm preserves acyclicity. Since at the end we finish with two nodes we know that there can be exactly one edge between them since the graph is a tree.

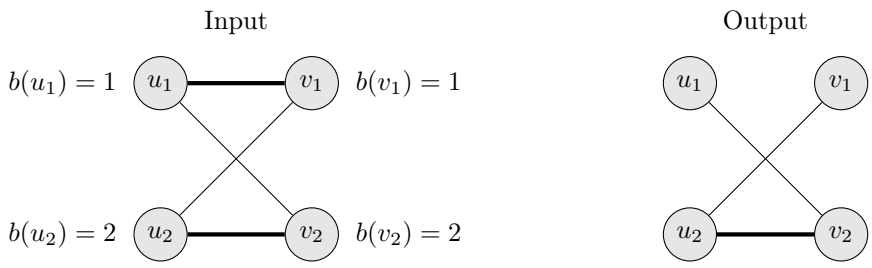
3 *(22 pts)* **Structure of extreme point solutions.** In this problem, we consider a generalization of the min-cost perfect matching problem. The generalization is called the *min-cost perfect b-matching problem* and is defined as follows:

Input: A graph $G = (V, E)$ with edge costs $c : E \to \mathbb{R}$ and degree bounds $b : V \to \{1, 2, \ldots, n\}$.

Output: A subset $F \subseteq E$ of minimum cost $\sum_{e \in F} c(e)$ such that for each vertex $v \in V$:

- The number of edges incident to v in F equals $b(v)$, i.e., $|\{e \in F : v \in e\}| = b(v)$.

Note that min-cost perfect matching problem is the special case when $b(v) = 1$ for all $v \in V$. An example with general b's is as follows:



On the left, we illustrate the input graph with the degree bounds (the b's). Thin and thick edges have cost 1 and 2, respectively. On the right, we illustrate a solution of cost $1 + 1 + 2 = 4$. It is a feasible solution since the degree of each vertex v equals $b(v)$ in the solution.

Your task is to prove the following statement: If the input graph $G = (V, E)$ is bipartite then any extreme point solution to the following linear programming relaxation (that has a variable x_e for every edge $e \in E$) is integral:

$$\begin{aligned} \textbf{Minimize} \quad & \sum_{e \in E} c(e) x_e \\ \textbf{subject to} \quad & \sum_{e \in E : v \in e} x_e = b(v) \qquad \text{for all } v \in V \\ & 0 \leq x_e \leq 1 \qquad \text{for all } e \in E. \end{aligned}$$

(In this problem you are asked to prove that every extreme point solution to the above linear program is integral assuming that the input graph G is bipartite. Recall that you are allowed to refer to material covered in the lecture notes.)

Solution:

Let x^* be an extreme point for the graph $G = (V, E)$ and let $E_f = \{e \in E : 0 < x^*_e < 1\}$. Suppose towards contradiction that $E_f \neq \emptyset$. Note that E_f must then contain a cycle as b is integral: indeed any vertex incident to an edge in E_f is incident to at least two edges in E_f. Note that since G is bipartite, hence, the cycle has even length. Let $e_1, e_2, ..., e_{2k}$ be the edges of the cycle. All these edges are fractional and we want to define y and z so that they are feasible solutions and $x^* = \frac{1}{2}(y + z)$ which will contradict the fact that x^* is an extreme point. Let y, z be

$$y_e = \begin{cases} x^*_e + \epsilon & \text{if } e \in \{e_1, e_3, e_5, ..., e_{2k-1}\} \\ x^*_e - \epsilon & \text{if } e \in \{e_2, e_4, e_6, ..., e_{2k}\} \\ x^*_e & \text{otherwise} \end{cases}$$

$$z_e = \begin{cases} x^*_e - \epsilon & \text{if } e \in \{e_1, e_3, e_5, ..., e_{2k-1}\} \\ x^*_e + \epsilon & \text{if } e \in \{e_2, e_4, e_6, ..., e_{2k}\} \\ x^*_e & \text{otherwise} \end{cases}$$

Notice that the degree constraints are still satisfied by y and z as we are alternating between increasing and decreasing the edge values in a cycle of even length. More formally suppose that e_i and e_{i+1} are the edges incident to v. Depending on parity of i we have $y_{e_i} = x^*_{e_i} + \epsilon$ and $y_{e_{i+1}} = x^*_{e_{i+1}} - \epsilon$. Or we have $y_{e_i} = x^*_{e_i} - \epsilon$ and $y_{e_{i+1}} = x^*_{e_{i+1}} + \epsilon$. Therefore degree constraints still satisfied:

$$\sum_{e \in E: v \in e} y_e = \sum_{e \in E: v \in e} x^*_e = b(v), \text{ and}$$
$$\sum_{e \in E: v \in e} z_e = \sum_{e \in E: v \in e} x^*_e = b(v)$$

Hence, to ensure feasibility, we need to choose such a small ϵ so as to guarantee that all y_e and z_e are in $[0, 1]$. For example $\epsilon = \min\{x^*_e, (1 - x^*_e) : e \in E_f\}$ gives that both y and z are feasible. Now one can easily see that $x^* = \frac{1}{2}(y + z)$ which contradicts the assumption that x^* is an extreme point.

4 *(22 pts)* **Probabilistic analysis.** Consider the following algorithm RANDOM-CHECK that takes as input two subsets $S \subseteq E$ and $T \subseteq E$ of the same ground set E.

RANDOM-CHECK(S, T)

1. For each element $e \in E$, independently of other elements randomly set

$$x_e = \begin{cases} 1 & \text{with probability } 1/3 \\ 0 & \text{with probability } 2/3 \end{cases}$$

2. **if** $\sum_{e \in S} x_e = \sum_{e \in T} x_e$ **then**
3. **return** true
4. **else**
5. **return** false

Note that RANDOM-CHECK(S, T) returns true with probability 1 if $S = T$. Your task is to analyze the probability that the algorithm returns true if $S \neq T$. Specifically prove that RANDOM-CHECK(S, T) returns true with probability at most $2/3$ if $S \neq T$.

(In this problem you are asked to prove that RANDOM-CHECK*(S, T) returns true with probability at most $2/3$ if $S \neq T$. Recall that you are allowed to refer to material covered in the lecture notes.)*

Solution: We solve the problem using "deferred decision" technique. First, note that simply one has

$$\sum_{e \in S} x_e = \sum_{e \in S \setminus T} x_e + \sum_{e \in S \cap T} x_e$$

and

$$\sum_{e \in T} x_e = \sum_{e \in T \setminus S} x_e + \sum_{e \in T \cap S} x_e.$$

So, we have

$$\Pr\left[\sum_{e \in S} x_e = \sum_{e \in T} x_e \mid S \neq T\right] = \Pr\left[\sum_{e \in S \setminus T} x_e = \sum_{e \in T \setminus S} x_e \mid S \neq T\right].$$

Note that since $S \neq T$, then $(S \setminus T) \cup (T \setminus S) \neq \emptyset$, and therefore $\exists f \in (S \setminus T) \cup (T \setminus S)$. Without lose of generality suppose that $f \in (S \setminus T)$. Then,

$$Pr\left[\sum_{e \in S \setminus T} x_e = \sum_{e \in T \setminus S} x_e \mid S \neq T\right] = Pr\left[x_f = \sum_{e \in T \setminus S} x_e - \sum_{e \in (S \setminus T) \setminus \{f\}} x_e \mid S \neq T\right].$$

At this point, assume that we know that values of x_e's for all $e \in E \setminus \{f\}$, so $c := \sum_{e \in T \setminus S} x_e - \sum_{e \in (S \setminus T) \setminus \{f\}} x_e$ is fixed. We just need to note that $\Pr[x_f = c | S \neq T] = \Pr[x_f = c] \leq \frac{2}{3}$ for any $c \in \mathbb{R}$ by assumption of the question (line 1 of RANDOM-CHECK(S, T)). So, the claim holds.

(This page is intentionally left blank.)

5 *(22 pts)* **Matroids.** Design a polynomial-time algorithm for the matroid matching problem:

Input: A bipartite graph $G = (A \cup B, E)$ and two matroids $\mathcal{M}_A = (A, \mathcal{I}_A)$, $\mathcal{M}_B = (B, \mathcal{I}_B)$.

Output: A matching $M \subseteq E$ of maximum cardinality satisfying:

(i) the vertices $A' = \{a \in A : \text{there is a } b \in B \text{ such that } \{a, b\} \in M\}$ of A that are matched by M form an independent set in $\mathcal{M}_A$, i.e., $A' \in \mathcal{I}_A$; and

(ii) the vertices $B' = \{b \in B : \text{there is an } a \in A \text{ such that } \{a, b\} \in M\}$ of B that are matched by M form an independent set in $\mathcal{M}_B$, i.e., $B' \in \mathcal{I}_B$.

We assume that the independence oracles for both matroids $\mathcal{M}_A$ and $\mathcal{M}_B$ can be implemented in polynomial-time. Also to your help you may use the following fact without proving it.

Fact (obtaining a new matroid by copying elements). Let $\mathcal{M} = (N, \mathcal{I})$ be a matroid where $N = \{e_1, \ldots, e_n\}$ consists of n elements. Now, for each $i = 1, \ldots, n$, make k_i copies of e_i to obtain the new ground set

$$N' = \{e_1^{(1)}, e_1^{(2)}, \ldots, e_1^{(k_1)}, e_2^{(1)}, e_2^{(2)}, \ldots, e_2^{(k_2)}, \ldots, e_n^{(1)}, e_n^{(2)}, \ldots, e_n^{(k_n)}\},$$

where we denote the k_i copies of e_i by $e_i^{(1)}, e_i^{(2)}, \ldots, e_i^{(k_i)}$. Then $(N', \mathcal{I}')$ is a matroid where a subset $I' \subseteq N'$ is independent, i.e., $I' \in \mathcal{I}'$, if and only if the following conditions hold:

(i) I' contains at most one copy of each element, i.e., we have $|I' \cap \{e_i^{(1)}, \ldots, e_i^{(k_i)}\}| \leq 1$ for each $i = 1, \ldots, n$;

(ii) the original elements corresponding to the copies in I' form an independent set in $\mathcal{I}$, i.e., if $I' = \{e_{i_1}^{(j_1)}, e_{i_2}^{(j_2)}, \ldots, e_{i_\ell}^{(j_\ell)}\}$ then $\{e_{i_1}, e_{i_2}, \ldots, e_{i_\ell}\} \in \mathcal{I}$.

Moreover, if the independence oracle of $(N, \mathcal{I})$ can be implemented in polynomial time, then the independence oracle of $(N', \mathcal{I}')$ can be implemented in polynomial time.

(In this problem you are asked to design and analyze a polynomial-time algorithm for the matroid matching problem. You are allowed to use the above fact without any proof and to assume that all independence oracles can be implemented in polynomial time. Recall that you are allowed to refer to material covered in the lecture notes.)

Solution: We use the given fact to create two matroids. First, using $\mathcal{M}_A$ we create $\mathcal{M}'_A$ as follows: For $a \in A$, create copies $a^{(b)}$ for each $b \in$ such that $(a, b) \in E$, and let $\mathcal{M}'_A$ be the matroid obtained from $\mathcal{M}_A$ using these new copies of vertices. Similarly obtain $\mathcal{M}'_B$ from $\mathcal{M}_B$ by copying each $b \in B$ to get $b^{(a)}$'s for each a such that $(a, b) \in E$. Notice that the ground set of $\mathcal{M}'_A$ has a one-to-one correspondence with E, and so do the ground set of $\mathcal{M}'_B$. Thus, w.l.o.g., we can assume that both the matroids $\mathcal{M}'_A$ and $\mathcal{M}'_B$ are defined on the common ground set E.

Now we can use matroid intersection, which runs in polynomial time, to find a maximum independent set in the intersection of the two matroids (since we have polynomial-time independence oracles for all the matroids we consider).

To see that this provides the required answer, note that any valid matching in two matroids $\mathcal{M}_A$ and $\mathcal{M}_B$ correspond to an independent set in the intersection of $\mathcal{M}'_A$ and $\mathcal{M}'_B$, and vice-versa (this follows from the given fact).

(This page is intentionally left blank.)