Problem Set I

Solutions to many homework problems, including problems on this set, are available on the Internet or can be obtained by an LLM, either for exactly the same problem formulation or for some minor perturbation. It is *not acceptable* to copy such solutions. It is hard to make strict rules on what information from the Internet you may use and hence whenever in doubt contact Michael Kapralov. You are, however, allowed to discuss problems in groups of up to three students.

1 A new type of matroid. *(30pts)* Let $G = (V, E)$ be a directed graph, and let $S, D \subseteq V$ be sets of "starting" and "destination" vertices (not necessarily disjoint). We say that a subset I of D is *linked* to S if the vertices of I can be reached from S via vertex disjoint paths. We define $\mathcal{I}$ to be the set of all subsets of D which are linked to S. Prove that $\mathcal{I}$ is a matroid.

Hint. Let P_X be some set of paths which link X to S and let P_Y be some set of paths which link Y to S. Let $P_X(v)$ the path in P_X which ends in $v \in X$ and, correspondingly, let $P_Y(v)$ be the path in P_Y which ends in $v \in Y$. Let N be the number of *intersections* between paths in P_X and P_Y where by an "intersection" of paths p and q we mean the maximal consecutive sequence of vertices in p which is also a consecutive sequence of vertices in q. Note that this implies that two paths may have more than one intersection.

To prove the second axiom for matroids – that is, that a smaller independent set can be augmented with an element from a bigger independent set – it might be useful to show that it is possible to rebuild P_Y into a set of vertex-disjoint paths $P_{Y'}$ linking Y' to S so that $|Y'| = |Y|$, $Y' \setminus X \subseteq Y \setminus X$ and one of the following happens:

- The number of intersections of P_X and $P_{Y'}$ is bounded by $N - 1$,
- The new collection of endpoints, Y', satisfies $|X \setminus Y'| = |X \setminus Y| - 1$.

Solution: We show that $\mathcal{I}$ satisfies the two axioms for matroids.

If $X \subseteq Y$ and $Y \in \mathcal{I}$ then $X \in \mathcal{I}$.

Indeed, $Y \in \mathcal{I}$ means by definition it is possible to select a collection of vertex-disjoint paths which start in S and end in Y, and for every vertex $y \in Y$ there is a path in this collection which ends in y. By taking this set of paths and removing those of them which end in $Y \setminus X$, we show that X is linked to S as well.

If $X \in \mathcal{I}$, $Y \in \mathcal{I}$ and $|Y| > |X|$ then $\exists e \in Y \setminus X : X \cup \{e\} \in \mathcal{I}$.

Note that once we have the procedure described in the hint, we can repeat it until we get a set of paths $P_{Y'}$ for which one of the following holds:

- The number of intersections between $P_{Y'}$ and P_X is 0,
- X is a subset of the collection of endpoints of $P_{Y'}$, i.e. $X \subseteq Y'$.

We show how to augment X with an element of Y to a set linked to S separately for each of the two cases.

- In the first case, because for any $e \in Y' \setminus X$ $P_{Y'}(e)$ does not intersect any of the paths in P_X, $P_X \cup P_{Y'}(e)$ links $X \cup \{e\}$ to W and hence $X \cup \{e\} \in \mathcal{I}$. Moreover, since $Y' \setminus X \subseteq Y \setminus X$, the vertex e indeed belongs to Y.

- In the second case, because $|Y| = |Y'| > |X|$, we have that $|Y' \setminus X| \neq 0$. Additionally, since $X \subset Y'$ and $Y' \in \mathcal{I}$, we get that $X \cup \{e\} \in \mathcal{I}$ for any $e \in Y' \setminus X$ by the first axiom of matroids. And, since $Y' \setminus X \subseteq Y \setminus X$, any such vertex e belongs to Y.

It only remains to show the construction of the procedure. Take any path p in P_X which intersects at least one of the paths in P_Y – if such a path does not exist, we are already in the case $N = 0$. Let $u \in V$ be the vertex in p which is closest to the endpoint of p out of the vertices which belong to intersections with paths in P_Y. Let q be the path in P_Y which passes through u. We define a path q' to coincide with q until q passes through u and, afterwards, to follow the path p. In particular, if u happens to be the endpoint of p, u will also be the endpoint of q'. We define $P_{Y'}$ to contain q' and all of the paths in P_Y except for q. By construction, the paths in $P_{Y'}$ are vertex-disjoint. Since $P_{Y'}$ consists of the same number of paths, $|Y| = |Y'|$. Moreover, since Y' is defined by replacing an element of Y with an element in X, we get that $Y' \setminus X \subseteq Y \setminus X$.

- If the path q, after it passes through u, intersects some paths in P_X, then the operation of replacing q with q' reduces the number of intersections by at least 1, and the first condition holds.

- If not, this means that q cannot have its endpoint in the set X, and thus replacing q with q' makes the number of points in $X \cap Y'$ greater than $X \cap Y$ by at least one, and the second condition holds.

$\square$

2 Intersection of 3 matroids. *(30pts)* We have seen in class that there exists an efficient algorithm for matroid intersection, i.e. a polynomial-time algorithm that given two matroids[1] over the same ground set, returns a common independent set of maximum size.

In this problem, you will show that finding an independent set of maximum size in the intersection of more than two matroids is unlikely to admit an efficient algorithm. More precisely, consider the following problem, called 3-Matroid-Intersection:

- **input:** three matroids $\mathcal{M}_1 = (E, \mathcal{I}_1), \mathcal{M}_2 = (E, \mathcal{I}_2), \mathcal{M}_3 = (E, \mathcal{I}_3)$ over the ground set E, and an integer k;

- **ouput**: YES if there exists a set $S \subseteq E$ of size at least k that is independent in each of $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ (i.e. $|S| \geq k$ and $S \in \mathcal{I}_1 \cap \mathcal{I}_2 \cap \mathcal{I}_3$), NO otherwise

[1] With an efficient independence oracle.

You are asked to show that 3-Matroid-Intersection is NP-hard.

Hint. In the Directed-Hamiltonian-Path problem, the input is a directed graph $G = (V, E)$, and one is asked to output YES or NO depending on whether G has a Hamiltonian path[2]. This problem is known to be NP-hard.

Solution: We reduce Directed-Hamiltonian-Path to 3-Matroid-Intersection. The ground set is the set of edges E of G. The first matroid is the graphic matroid, so $S \subseteq E$ is independent in $\mathcal{M}_1$ if it is acyclic (in the undirected sense). The second matroid is a partition matroid defined so that $S \subseteq E$ is independent in $\mathcal{M}_2$ if, for every vertex u, it contains at most one edge leaving u. Finally, the third matroid is a partition matroid defined so that $S \subseteq E$ is independent in $\mathcal{M}_2$ if, for every vertex u, it contains at most one edge going into u. In the reduction, we set $k = n - 1$.

If G has a Hamiltonian path, the set of edges used by such path is independent in all three matroids, since paths are acyclic and the in- and out-degrees are at most 1. Moreover, it must consist of $n - 1$ elements.

Conversely, we argue that if there is a set of edges S of size $n - 1$ that is independent in all three matroids, then G has a Hamiltonian path. To see this, the fact that the size of S is $n - 1$ and S is independent in $\mathcal{M}_1$ means that it is a spanning tree, so it connects all vertices. Moreover, in- and out-degrees are at most 1, thanks to independence in $\mathcal{M}_2$ and $\mathcal{M}_3$. A connected acyclic graph with in- and out-degrees bounded by 1 is a path. □

3 Transportation. *(40pts)* In the minimum transportation problem, we are given a road network (of possibly one-way streets) and a central warehouse, and the goal is to select a cheap subset of roads that enables to send goods from the central warehouse to all other vertices respecting the direction of traffic. More formally, we consider the following problem:

- **input:** a directed graph $G = (V, E)$, a vertex $c \in V$ such that there is a directed path from c to every other vertex in G, and edge weights $w_e \in \mathbb{R}$ for every $e \in E$;
- **output:** a set of edges $F^* \subseteq E$ of minimum weight such that there is a directed path from c to every other vertex in the directed graph (V, F^*).

Consider the following linear programming relaxation of this problem[3].

$$\begin{aligned} & \min \sum_{e \in E} w_e x_e \\ \text{s.t.} \quad & \sum_{e \in \delta^-(S)} x_e \geq 1, \quad \forall \emptyset \neq S \subseteq V \setminus \{c\} \qquad \text{(LP)} \\ & x_e \geq 0, \quad \forall e \in E \end{aligned}$$

[2]A Hamiltonian path is a directed simple path that visits every vertex exactly once.

[3]Throughout this problem, for a set of vertices $S \subseteq V$, we use $\delta^-(S)$ to denote the subset of edges (u, v) of G going into S, i.e. such that $u \notin S, v \in S$.

The goal of this problem is to show that the extreme points of the above linear program are integral. To do so, we will make use of the following fact.

Fact 1. *Any extreme point of a linear program can be written as the unique solution of a linear system of tight constraints*[4]*. In other words, consider a linear program over $y \in \mathbb{R}^n$ whose feasible region is given by $Ay \geq b$ where $A \in \mathbb{R}^{t \times n}, b \in \mathbb{R}^t$: then, for any extreme point y^* there is a subset $R \subseteq [t]$ of size n such that y^* is the unique solution of the linear system $A' y = b'$, where A' and b' are the restriction of A and b to the rows in R (so the rows of A' are linearly independent).*

Let $x \in \mathbb{R}^E$ be an extreme point solution to the above (LP) for the minimum transportation problem.

Assumption 2. *Without loss of generality, we assume that $x_e > 0$ for all $e \in E$. Indeed, if this was not the case, we could consider the graph G' obtained by only keeping the edges in the support of x, and the restriction of x to the edges in its support would be an extreme point of the linear program for G'.*

By virtue of this assumption, none of the constraints of the form $x_e \geq 0$ is satisfied with equality by x. Let $\mathcal{F}$ be the collection of vertex subsets such that $S \in \mathcal{F}$ if the corresponding constraint is tight for x, i.e. the constraint is satisfied with equality. In particular, the extreme point solution x is the unique solution to the linear system,

$$\sum_{e \in \delta^-(S)} x_e = 1, \qquad \forall S \in \mathcal{F}\,.$$

As a first step, you will show that the tight constraints $\mathcal{F}$ are implied by a *laminar*[5] subset of tight constraints $\mathcal{L} \subseteq \mathcal{F}$. With $\mathbb{1}_F \in \{0,1\}^E$ we denote the characteristic vector of $F \subseteq E$, defined as

$$(\mathbb{1}_F)_e \;=\; \begin{cases} 1 & \text{if } e \in F, \\ 0 & \text{if } e \notin F \end{cases}.$$

3a *(15 pts)* Show that if $S, T \in \mathcal{F}$ with $S \cap T \neq \emptyset$, then the following three conditions hold:

- $S \cup T$, $S \cap T \in \mathcal{F}$;
- $E(S \setminus T, T \setminus S) = E(T \setminus S, S \setminus T) = \emptyset$;
- $\mathbb{1}_{\delta^-(S)} + \mathbb{1}_{\delta^-(T)} = \mathbb{1}_{\delta^-(S \cup T)} + \mathbb{1}_{\delta^-(S \cap T)}$.

Solution: Note that

$$\mathbb{1}_{\delta^-(S)} + \mathbb{1}_{\delta^-(T)} = \mathbb{1}_{\delta^-(S \cup T)} + \mathbb{1}_{\delta^-(S \cap T)} + \mathbb{1}_{E(S \setminus T, T \setminus S)} + \mathbb{1}_{E(T \setminus S, S \setminus T)}\,. \tag{1}$$

[4] A constraint is said to be tight for a feasible solution if it is satisfied with equality.

[5] Remember that a collection of sets $\mathcal{L}$ is called laminar if for all $S, T \in \mathcal{L}$ either $S \subseteq T$, or $T \subseteq S$, or $S \cap T = \emptyset$.

Thus,

$$x(\delta^-(S))+x(\delta^-(T)) = x(\delta^-(S\cup T))+x(\delta^-(S\cap T))+x(E(S\backslash T, T\backslash S))+x(E(T\backslash S, S\backslash T))\,. \tag{2}$$

Since S and T are in $\mathcal{F}$, they are tight. Combining then (2) with the non-negativity of x, we get

$$2 = x(\delta^-(S)) + x(\delta^-(T)) \geq x(\delta^-(S\cup T)) + x(\delta^-(S\cap T)) \geq 2\,, \tag{3}$$

where the last equality follows by feasibility of x. Hence

$$x(\delta^-(S\cup T)) + x(\delta^-(S\cap T)) = 2$$

and therefore $S\cup T$ and $S\cap T$ are tight (i.e. they are in $\mathcal{F}$). From (3) and (2), we also get

$$x(\delta^-(S))+x(\delta^-(T)) = x(\delta^-(S\cup T))+x(\delta^-(S\cap T)) \implies x(E(S\backslash T, T\backslash S))+x(E(T\backslash S, S\backslash T)) = 0\,.$$

As we assumed that no edge has $x_e = 0$, we conclude that $E(S\backslash T, T\backslash S) = E(T\backslash S, S\backslash T) = \emptyset$. This fact, together with (1), gives the third item. $\square$

3b *(15 pts)* Let $\mathcal{L} \subseteq \mathcal{F}$ be a maximal laminar subfamily[6] of $\mathcal{F}$. Show that

$$\operatorname{span}\{\mathbb{1}_{\delta^-(S)} \mid S \in \mathcal{L}\} = \operatorname{span}\{\mathbb{1}_{\delta^-(S)} \mid S \in \mathcal{F}\}\,.$$

Hint: Suppose that this is not true. Among all sets $T \in \mathcal{F}$ with $\mathbb{1}_{\delta^-(T)} \notin \operatorname{span}\{\mathbb{1}_{\delta^-(S)} \mid S \in \mathcal{L}\}$ choose one that minimizes the number of sets in $\mathcal{L}$ for which T violates laminarity, i.e. pick $T \in \mathcal{F}$ as to minimize $|\{L \in \mathcal{L} \mid L\cap T \neq \emptyset, L\setminus T \neq \emptyset, T\setminus L \neq \emptyset\}|$. Use the result of the previous question to arrive at a contradiction.

Solution: Let T as in the hint. Note that $|\{L \in \mathcal{L} \mid L\cap T \neq \emptyset, L\setminus T \neq \emptyset, T\setminus L \neq \emptyset\}| \geq 1$, otherwise the maximality of $\mathcal{L}$ is violated. Then, let $S \in \mathcal{L}$ such that $S\cap T \neq \emptyset, S\setminus T \neq \emptyset, T\setminus S \neq \emptyset$. Using the result of **3a**, we have

$$\mathbb{1}_{\delta^-(S)} + \mathbb{1}_{\delta^-(T)} = \mathbb{1}_{\delta^-(S\cup T)} + \mathbb{1}_{\delta^-(S\cap T)} \iff \mathbb{1}_{\delta^-(T)} = -\mathbb{1}_{\delta^-(S)} + \mathbb{1}_{\delta^-(S\cup T)} + \mathbb{1}_{\delta^-(S\cap T)}\,.$$

Hence, if both $S\cup T$ and $S\cap T$ are in $\mathcal{L}$, then $\mathbb{1}_{\delta^-(T)} \in \operatorname{span}\{\mathbb{1}_{\delta^-(S)} \mid S \in \mathcal{L}\}$, which is a contradiction to the definition of T. On the other hand, if one of $S\cup T$ or $S\cap T$ were not in $\mathcal{L}$, it would be better than T, i.e. if $X \in \{S\cup T, S\cap T\}$ is not in $\mathcal{L}$ then

$$|\{L \in \mathcal{L} \mid L\cap X \neq \emptyset, L\setminus X \neq \emptyset, X\setminus L \neq \emptyset\}| < |\{L \in \mathcal{L} \mid L\cap T \neq \emptyset, L\setminus T \neq \emptyset, T\setminus L \neq \emptyset\}|\,,$$

contradicting again the definition of T. $\square$

[6] This means that adding any other set $S \in \mathcal{F}$ to $\mathcal{L}$ would violate its laminarity.

From the result of **3a** and **3b**, we can conclude that x is the unique solution of a linear system

$$\sum_{e\in\delta^-(S)} x_e = 1, \qquad \forall S \in \mathcal{L}\,,$$

where $\mathcal{L}$ is a laminar family of tight constraints. Next, you will show that x has to be integral. To do so, you can make use of the following theorem.

Theorem 3. *Let y be the unique solution to a linear system $Ay = b$ where A and b have entries in $\{0, 1\}$. Then y is integral if for every subset R of rows of A, there exists a partition of R into two (possibly empty) parts R_1 and R_2 such that the vector*

$$\sum_{i\in R_1} A_i - \sum_{i\in R_2} A_i$$

has all entries in $\{-1, 0, 1\}$. Here A_i denotes the i-th row of A.

3c *(10 pts)* Use Theorem 3 to show that x is in fact integral.

Solution: Let $\mathcal{L}' \subseteq \mathcal{L}$ a subset of the rows of the linear system defining x. Since $\mathcal{L}$ is laminar, so is $\mathcal{L}'$. We partition $\mathcal{L}'$ as follows: we put in $\mathcal{L}_1$ (resp. $\mathcal{L}_2$ the sets of $\mathcal{L}'$ that are contained in an even (resp. odd) number of sets of $\mathcal{L}'$, i.e. $\mathcal{L}_1 = \{S \in \mathcal{L} : |\{T \in \mathcal{L}' : S \subseteq T\}| = 0 \mod 2\}$ and $\mathcal{L}_2 = \{S \in \mathcal{L} : |\{T \in \mathcal{L}' : S \subseteq T\}| = 1 \mod 2\}$. Then, for any $e = (u, v) \in E$, we have

$$\left(\sum_{L\in\mathcal{L}_1} \mathbb{1}_{\delta^-(L)} - \sum_{L\in\mathcal{L}_2} \mathbb{1}_{\delta^-(L)}\right)_e = |\{L \in \mathcal{L}_1 \mid e \in \delta^-(L)\}| - |\{L \in \mathcal{L}_2 \mid e \in \delta^-(L)\}|$$

If there is no set $S \in \mathcal{L}'$ such that $u \notin S$ and $v \in S$, then the right-hand side is 0. Otherwise, let X and Y in $\mathcal{L}'$ be, respectively, the smallest and largest sets to contain v but not u. By laminarity, such sets are unique. Moreover, by laminarity, for any other set $S \in \mathcal{L}'$ such that $u \notin S$ and $v \in S$ we have $X \subseteq S \subseteq Y$. Therefore, the right-hand side above side must evaluate to one of $\{-1, 0, 1\}$. □