Advanced Algorithms | **University of Bremen**
Summer 2025, Lecture 12 | Faculty 3: Math/CS
Lecturer: Alexander Lindermayr | June 23, 2025

Introduction to Matroids

In this and the next lecture, we introduce matroids and learn how they characterize problems that can be solved by a greedy algorithm.

But first, we start by recalling the problem of finding a minimum spanning tree (MST) in a graph. We have seen that the problem can be solved using Kruskal's algorithm, which is a greedy algorithm. Formally, we are given an undirected graph $G = (V, E)$ with edge weights $w : E \to \mathbb{R}$, and we want to find a spanning tree $T \subseteq E$ of G that minimizes the total weight $\sum_{e \in T} w(e)$.

Algorithm 1: Kruskal's algorithms

```
F ← ∅
for e ∈ E sorted ascending by w(e) do
    if F ∪ {e} is acyclic then
        F ← F ∪ {e}
return F
```

Kruskal's Greedy algorithm finds an MST for every weight function w. In particular, it also works for the analogue maximization problem, where we want to find a spanning tree that maximizes the total weight $\sum_{e \in T} w(e)$.

In this lecture, we want to understand for which kind of problems this type of greedy algorithm computes an optimal solution:

- Consider elements greedily one-by-one.
- Add an element to the solution if it maintains feasibility.

A first observation is that there exists a polynomial-time solvable problem for which this algorithm does compute an optimal solution: the maximum bipartite matching problem. Here, greedily adding edges to our matching does not necessarily lead to a maximum matching. This raises the following question:

Which properties guarantee that the Greedy algorithm computes an optimal solution?

To answer this question, we first introduce definitions to abstract many problems in combinatorial optimization.

Definition 1 (Independence System). *Let E be a ground set. A set system $\mathcal{I} \subseteq 2^E$ is called* independence system *if*

(i) $\emptyset \in \mathcal{I}$, *and*

(ii) for all $A \in \mathcal{I}$ *and* $B \subseteq A$, *we have* $B \in \mathcal{I}$.

A set $A \subseteq E$ *is called* independent *if* $A \in \mathcal{I}$, *and* dependent *if* $A \notin \mathcal{I}$. *Minimal dependent sets are called* circuits, *and maximal independent sets are called* bases.

For some set $A \subseteq E$, we call a maximum independent subset of A a *basis* of A.

Using the notion of independence systems, we can reformulate many known combinatorial optimization problems in this language: Given an independence system $(E, \mathcal{I})$ and a weight function $w : E \to \mathbb{R}$, find $I \in \mathcal{I}$ that maximizes / minimizes $w(I) := \sum_{e \in I} w(e)$.

We have seen many examples of such problems:

- **Minimum spanning tree:** $w(e)$ is the weight of edge e, and $\mathcal{I} = \{I \subseteq E \mid I \text{ forest}\}$.
- **Maximum matching:** $w(e) = 1$, and $\mathcal{I} = \{I \subseteq E \mid I \text{ matching}\}$.
- **Knapsack:** $w(e)$ is the value of item e, and $\mathcal{I} = \{I \subseteq E \mid \sum_{e \in I} w(e) \leq B\}$ for some capacity B.
- **Maximum weight independent set:** $w(v)$ is the weight of vertex v, and $\mathcal{I} = \{I \subseteq V \mid I \text{ is independent in } G\}$.

While all of these problems are optimization problems over independence systems, we will see that exactly those can be solved by the greedy algorithm that are a *matroid*.

Definition 2 (Matroid)**.** *An independence system* $(E, \mathcal{I})$ *is called a* matroid *if*

(iii) for all $A, B \in \mathcal{I}$ *with* $|A| < |B|$*, there exists an element* $b \in B \setminus A$ *such that* $A \cup \{b\} \in \mathcal{I}$*.*

Property (iii) is also called *augmentation property*. This property is crucial for the greedy algorithm to work.

Hence, a set system $(E, \mathcal{I})$ is a matroid if it satisfies (i), (ii), and (iii). Note that (ii) implies (i) as long as $\mathcal{I} \neq \emptyset$; we add (i) to rule out that $(E, \emptyset)$ is a matroid.

1 Standard Matroids

1.1 Uniform Matroids

Let E be a universe and let $k \in \mathbb{N}$. Define

$$\mathcal{I} = \{I \subseteq E \mid |I| \leq k\}.$$

Theorem 1. *$\mathcal{M} = (E, \mathcal{I})$ is a matroid (called a* uniform *matroid).*

Proof. Properties (i) and (ii) clearly hold. We now show the augmentation property (iii). Let $A, B \in \mathcal{I}$ with $|A| < |B|$. Since $|A| < |B| \leq k$, there exists at least one element $b \in B \setminus A$. Then

$$|A \cup \{b\}| = |A| + 1 \leq k,$$

so $A \cup \{b\} \in \mathcal{I}$. This verifies the augmentation property. □

1.2 Partition Matroids

Let E be a ground set and let $E_1, E_2, \ldots, E_\ell$ be a partition of E. For fixed integers $k_1, k_2, \ldots, k_\ell$, define

$$\mathcal{I} = \{I \subseteq E \mid |I \cap E_i| \leq k_i \text{ for all } 1 \leq i \leq \ell\}.$$

Theorem 2. *$\mathcal{M} = (E, \mathcal{I})$ is a matroid (called a* partition *matroid).*

Proof. Properties (i) and (ii) clearly hold. We now show the augmentation property (iii). Let $A, B \in \mathcal{I}$ with $|A| < |B|$. Since

$$\sum_{i=1}^{\ell} |A \cap E_i| < \sum_{i=1}^{\ell} |B \cap E_i|,$$

there exists at least one index j with

$$|A \cap E_j| < |B \cap E_j|.$$

Choose any $b \in (B \cap E_j) \setminus A$. Then,

$$|(A \cup \{b\}) \cap E_j| = |A \cap E_j| + 1 \leq |B \cap E_j| \leq k_j,$$

and for all $i \neq j$,

$$|(A \cup \{b\}) \cap E_i| = |A \cap E_i| \leq k_i.$$

Thus, $A \cup \{b\} \in \mathcal{I}$, proving the augmentation property. □

1.3 Linear Matroids

Let F be a field and let $A \in F^{m \times n}$ be a matrix whose columns are indexed by a ground set $E = \{1, \ldots, n\}$. Define

$$\mathcal{I} = \{I \subseteq E \mid \text{the columns of } A_I \text{ are linearly independent}\},$$

where A_I denotes the submatrix of A consisting of the columns indexed by I.

Theorem 3. *$\mathcal{M} = (E, \mathcal{I})$ is a matroid (called a* linear *matroid).*

1.4 Graphic Matroids

Let $G = (V, E)$ be a graph. Define

$$\mathcal{I} = \{I \subseteq E \mid I \text{ is acyclic}\} .$$

Theorem 4. *$\mathcal{M} = (E, \mathcal{I})$ is a matroid (called a* graphic *matroid).*

In particular, this shows that the problem of finding a MST in a graph is a matroid optimization problem.

Proof. Properties (i) and (ii) clearly hold, because removing edges of an acyclic graph cannot create cycles. We now show the augmentation property (iii).

Let $A, B \in \mathcal{I}$ with $|A| < |B|$. Both A and B are forests in $G = (V, E)$. Let $k(F)$ denote the number of connected components of a forest F on the vertex set V. The number of edges in such a forest is $|V| - k(F)$. Thus, $|A| = |V| - k(A)$ and $|B| = |V| - k(B)$. Since $|A| < |B|$, it follows that $|V| - k(A) < |V| - k(B)$, which implies $k(A) > k(B)$. Now, assume for contradiction that for every edge $e \in B \setminus A$, the set $A \cup \{e\}$ is cyclic. This means that the endpoints of e are already connected in the forest (V, A). If this holds for all $e \in B \setminus A$, then every edge in B connects vertices that lie within the same connected component of (V, A). Consequently, each connected component of (V, B) must be a subgraph of some connected component of (V, A). This implies that $k(B) \geq k(A)$, a contradiction to $k(A) > k(B)$. Therefore, there must exist an edge $b \in B \setminus A$ such that its endpoints lie in different connected components of (V, A). Adding such an edge b to A results in $A \cup \{b\}$ being acyclic. Thus, $A \cup \{b\} \in \mathcal{I}$, which proves the augmentation property. $\square$

1.5 Matching Matroids

Let $G = (V, E)$ be a graph. Define

$$\mathcal{I} = \{I \subseteq V \mid \text{there exists a matching in } G \text{ that covers } I\}.$$

Theorem 5. *$\mathcal{M} = (V, \mathcal{I})$ is a matroid (called a* matching *matroid).*

Proof. Properties (i) and (ii) clearly hold, because any subset of a matching is also a matching. We now show the augmentation property (iii).

Let $A, B \in \mathcal{I}$ with $|A| < |B|$, and let M_A, M_B be matchings covering A and B respectively. If there is a $b \in B \setminus A$ that is covered by M_A, then $A \cup \{b\} \in \mathcal{I}$ and we are done. Otherwise, the symmetric difference $M_A \triangle M_B$ contains alternating paths between edges of M_A and M_B that start in $b \in B \setminus A$. Since $|B \setminus A| > |A \setminus B|$, there must exists an augmenting path P that ends in $B \setminus A$. Thus, $A \cup \{b\}$ is covered by $M_A \Delta P$, proving the augmentation property. $\square$

As mentioned above, the independence system $(E, \mathcal{I})$ where $G = (V, E)$ where each $I \in \mathcal{I}$ is a matching in G is **not** a matroid. (Proof as an exercise.)

2 The Rank of a Matroid

We first show that the size of every base of a matroid is the same.

Lemma 1. *Every base of a matroid has the same size.*

Proof. Let B_1 and B_2 be two bases of the matroid, and suppose for contradiction that $|B_1| < |B_2|$. By the augmentation property, there is some element $b \in B_2 \setminus B_1$ such that $B_1 \cup \{b\} \in \mathcal{I}$. But this contradicts the maximality of B_1. Hence $|B_1| = |B_2|$. $\square$

The size of a base of a matroid is thus also called the *rank* of the matroid.

Definition 3. *Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid. The* rank function *$r : 2^E \to \mathbb{N}_{\geq 0}$ associated to $\mathcal{M}$ is defined as*

$$r(S) := \max_{I \subseteq S s.t. I \in \mathcal{I}} |I|$$

for each $S \subseteq E$. We call $r(E)$ the rank *of $\mathcal{M}$.*

Informally, the rank of a set $S \subseteq E$ is the size of the largest independent subset of S.