# Lecture 7

Lecturer: Prof. Allan Borodin
Scribe: Ioana Ivan

Independence systems:
Let $M = (E, \mathcal{F})$, where E is a set of elements, $\mathcal{F} \subseteq 2^{|E|}$ and, if $I \in \mathcal{F}$, I is called an independent set.
An independence system satisfies the following properties:
1) $\emptyset \in \mathcal{F}$
2) $S \subseteq T, T \in \mathcal{F} \Rightarrow S \in \mathcal{F}$

A matroid is an independence system that also satisfies:
3) $S, T \in \mathcal{F}, |S| < |T|$, then $\exists x \in T \setminus S$ such that $S \cup \{x\} \in \mathcal{F}$

Admissibility system satisfy propertiy (1) and the following property:
2')$A \neq \emptyset \in \mathcal{F} \Rightarrow \exists x \in A, A \setminus \{x\} \in \mathcal{F}$

A greedoid(introduced by Korte and Lovasz) is an admissibility system that also satisfies (3).

Example of matroids:
- linear independence in a vector system
- $\{S$ : S is an acyclic set of edges in a given graph$\}$
- uniform matroid: $\{S : |S| \leq k\}$

A maximal independent set is called a **basis**.
**Fact**(equivalent definition):
Every basis in a matroid has the same cardinality.

Consider the following two basic problems:
P1. Computing a maximum weight basis in a matroid.
P2. Maximizing a submodular function subject to a matroid constraint.

The natural(standard) greedy algorithm for P1 is optimal.

Sort so that $w_1 \geq w_2 \geq ... \geq w_m$
For $i = 1...m$
    If $S \cup \{e_i\} \in \mathcal{F}$
    Then $S = S \cup \{e_i\}$
End for

The proof of optimality is similar to the proof of correctness for Kruskal's algorithm.

Submodular set functions:
A function f: $2^E \to \mathcal{R}$ (i.e. which takes subsets of E into $\mathcal{R}$ is submodular if:

$f(A \cup B) + f(A \cap B) \leq f(A) + f(B)$
Equivalently, if $A \subseteq B$, then $f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B)$. (decreasing marginal utility).

f is monotone if for all $A \subseteq B$, $f(A) \leq f(B)$.
f is normalized if $f(\emptyset) = 0$.

Applications of monotone submodular function:
Influence in social networks:
Two models for behaviour adoption in social networks:
1. **Threshold model**: nodes become influenced because a sufficiently large weighted proportion of their neighbors have been influenced.
2. **Cascade model**: The influence process unfolds in discrete steps. When a node first becomes influenced in step t, it gets a single chance to influence each of its neighbors, and succeeds with a certain probability(which depends on the two nodes).
Under both models, the total number of influenced nodes is a monotone submodular function.

Greedy algorithms for maximizing submodular functions are studied in:
1. An analysis of approximations for maximizing submodular set functions I by Nemhauser, Wolsey and Fisher
2. An analysis of approximations for maximizing submodular set functions II by Fisher, Nemhauser and Wolsey

Results:
1. The standard greedy algorithm for maximizing a submodular function subject to an uniform matroid is a $\frac{e}{e-1}$ approximation.
2. The standard greedy algorithm for maximizing a submodular function subject to an arbitrary matroid is a 2-approximation.
3. The standard greedy algorithm for maximizing a submodular function subject to k matroids (k matroid intersection problem) is a (k+1)-approximation.
Note that for k=2, the problem can be solved in polynomial time, but it's NP-hard for $k \geq 3$.

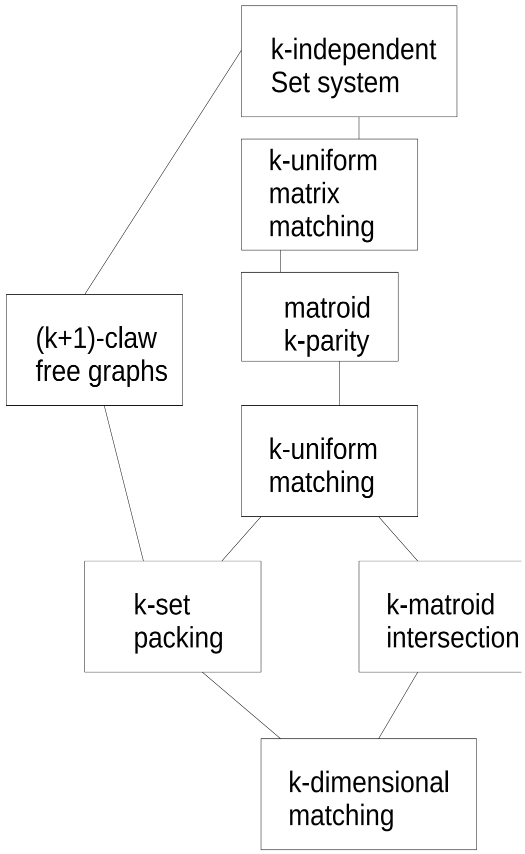
Example: maximum matching in a bipartite graph is the intersection of two matroids (one specifies each vertex on one side has degree at most one, the other one does the same for the other side).

The natural(standard) greedy algorithm for P2:
$S = \emptyset$
While $\exists x : S \cup \{x\} \in \mathcal{F}$:
    Let $u* = \max_{e \in E} f(S \cup \{e\}) - f(S)$
    % Note we assume the algorithm has access to an "incremental oracle" which returns $u*$
    % If $u*$ is an $\alpha$-approximation, we call it an "$\alpha$-approximative incremental oracle"
    If $S \cup \{u*\} \in \mathcal{F}$
        $S = S \cup \{u*\}$
    $E = E \setminus \{u*\}$
End While

Results:
1. If the standard greedy algorithm for maximizing a submodular function subject to an uniform matroid

has access to an $\alpha$-approximative incremental oracle, it's a $\frac{e^{\frac{1}{\alpha}}}{e^{\frac{1}{\alpha}}-1}$ approximation(and this bound is tight).
2. If the standard greedy algorithm for maximizing a submodular function subject to k matroids has access to an $\alpha$-approximative incremental oracle, it's a $(\alpha + 1)k-$approximation.

The greedy algorithm is optimal in an independent system iff the system is a matroid.
The greedy algorithm is optimal in an admissibility system iff the system is a greedoid.




k-independent set systems(E, $\mathcal{F}$) - as defined and first studied by Jenkins
$\forall Y \subseteq E$:
$\frac{max_{S \subseteq Y, S \in \mathcal{F}} |S|}{min_{S \subseteq Y, S \in \mathcal{F}} |S|} \leq k$
In a matroid k = 1 (all basis have the same size).

The standard greedy algorithm for maximum weight $I \in \mathcal{F}$(I is a k-independent set) is a k-approximation.
The standard greedy algorithm for maximizing a submodular function subject to independence in a k-independence set system constraint is a (k+1)-approximation.

Using local search, we can get better approximations for linear and submodular function maximization in many frameworks;
Results by Halldorsson:
1. $\frac{k+1}{2}$ approximation for unweighted MIS on (k+1)-claw free graphs.
$\frac{k}{2} + \epsilon$ using t-local search.
t-local search for weighted k-set packing is exactly a $(k - 1 + \frac{1}{t})$-approximation.

2. For unweighted versions of the matroid problems (up to k-uniform matroid matching) we can get a $\frac{k}{2}+\epsilon$ approximation.

Berman:
1. We can get a $\frac{k+1}{2}+\epsilon$ approximation for weighted MIS on (k+1)-claw free graphs.

Proof that Halldorsson's algorithm for unweighted MIS is a $\frac{k+1}{2}$ approximation:
While $\exists u,v : |S' = S \cup \{u,v\} \setminus N(u) \setminus N(v)| > |S|$ or
$\exists u : |S' = S \cup \{u\} \setminus N(u)| > |S|$
$S = S'$
End While

Let A be the independent set produced by the greedy algorithm, and B be a maximum independent set.

Let $A' = A \setminus (A \cap B)$,
$B' = B \setminus (A \cap B)$.

Let $B_1 = \{v \in B' | v$ has exactly one neighbor in A$\}$
$B_2 = \{v \in B' | v$ has at least two neighbors in A$\}$
$A_1 = \{u \in A' | u$ has at exactly one neighbor in $B_1\}$
Note that
$|B_1| + 2 \cdot |B_2| \leq k|A|$(since the graph is (k+1)-claw free)

By the defintion of $B_1$, we have $|B_1| \geq A_1$. If $|B_1| > |A_1|$, then, by pigeon hole property, there exists $x \in A_1$ such that at least two $y_i$ in $B_1$ are adjacent to x, and therefore A is not 2-local optimum.
So we must have $|B_1| = |A_1|$

Therefore:
$2 \cdot |B_1| + 2 \cdot |B_2| \leq (k+1)|A|$
So, $2|B| \leq (k+1)|A|$
This shows that the algorithm is a $\frac{k+1}{2}$ approximation.

Chandra and Halldorsson also show a $\frac{2(k+1)}{3}$-approximation for weighted MIS.
They use a greedy algorithm to approximate a max weight independent set. Then they use local search:
While there exists claw C such that $w(S \cup C \setminus N(C,S)) > w(S)$:
$S = S \cup C \setminus N(C,S)$
End While

**Berman** uses a non-oblivious local search algorithm:
While there exists claw C such that $w^2(S \cup C \setminus N(C,S)) > w^2(S)$:
$S = S \cup C \setminus N(C,S)$
End While