Lecture 17: Matroids and the Greedy Algorithm

Michael Dinitz

October 26, 2021

601.433/633 Introduction to Algorithms

Introduction

Last time: somewhat greedy algorithm (Prim's), extremely greedy algorithm (Kruskal's)

Question: when does greedy algorithm return optimal solution?

Introduction

Last time: somewhat greedy algorithm (Prim's), extremely greedy algorithm (Kruskal's)

Question: when does greedy algorithm return optimal solution?

Want abstraction that includes MSTs, but also works for many other problems.

Introduction

Last time: somewhat greedy algorithm (Prim's), extremely greedy algorithm (Kruskal's)

Question: when does greedy algorithm return optimal solution?

Want abstraction that includes MSTs, but also works for many other problems.

Weighted Set System:

- Universe $\mathbf{U}$
- Collection $\mathcal{I} \subseteq \mathbf{2^U}$ (so $\mathbf{I} \subseteq \mathbf{U}$ for all $\mathbf{I} \in \mathcal{I}$). Called *independent sets*
- Weights $\mathbf{w} : \mathbf{U} \to \mathbb{R}^+$

Introduction

Last time: somewhat greedy algorithm (Prim's), extremely greedy algorithm (Kruskal's)

Question: when does greedy algorithm return optimal solution?

Want abstraction that includes MSTs, but also works for many other problems.

Weighted Set System:

- Universe $\mathbf{U}$
- Collection $\mathcal{I} \subseteq \mathbf{2^U}$ (so $\mathbf{I} \subseteq \mathbf{U}$ for all $\mathbf{I} \in \mathcal{I}$). Called *independent sets*
- Weights $\mathbf{w} : \mathbf{U} \rightarrow \mathbb{R}^+$

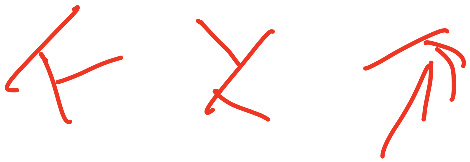
Problem: find *max weight* independent set

MST as Weighted Set System

MST: weighted graph $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{w})$. Find MST.

Set system:

- $\mathbf{U} = \mathbf{E}$
- $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F})$ a forest$\}$



MST as Weighted Set System

MST: weighted graph $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{w})$. Find MST.

Set system:

- $\mathbf{U} = \mathbf{E}$
- $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

What about weights? MST is minimize, but problem we defined is maximize.

MST as Weighted Set System

MST: weighted graph $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{w})$. Find MST.

Set system:

- $\mathbf{U} = \mathbf{E}$
- $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

What about weights? MST is minimize, but problem we defined is maximize.

- Let $\bar{\mathbf{w}} > \mathbf{w}(\mathbf{e})$ for all $\mathbf{e} \in \mathbf{E}$, let $\mathbf{w}'(\mathbf{e}) = \bar{\mathbf{w}} - \mathbf{w}(\mathbf{e})$ for all $\mathbf{e} \in \mathbf{E}$

MST as Weighted Set System

MST: weighted graph $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{w})$. Find MST.

Set system:

- $\mathbf{U} = \mathbf{E}$
- $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

What about weights? MST is minimize, but problem we defined is maximize.

- Let $\bar{\mathbf{w}} > \mathbf{w}(\mathbf{e})$ for all $\mathbf{e} \in \mathbf{E}$, let $\mathbf{w}'(\mathbf{e}) = \bar{\mathbf{w}} - \mathbf{w}(\mathbf{e})$ for all $\mathbf{e} \in \mathbf{E}$

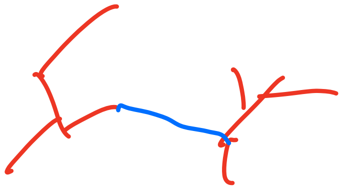
For any spanning tree $\mathbf{T}$:

$$\mathbf{w}'(\mathbf{T}) = \sum_{\mathbf{e} \in \mathbf{T}} \mathbf{w}'(\mathbf{e}) = \sum_{\mathbf{e} \in \mathbf{T}} (\bar{\mathbf{w}} - \mathbf{w}(\mathbf{e})) = (\mathbf{n} - \mathbf{1})\bar{\mathbf{w}} - \sum_{\mathbf{e} \in \mathbf{T}} \mathbf{w}(\mathbf{e}) = (n-1)\bar{w} - w(T)$$

$$= \sum_{e \in T} \bar{w} - \sum_{e \in F} w(e)$$

MST as Weighted Set System

MST: weighted graph $\mathbf{G = (V, E, w)}$. Find MST.



Set system:

- $\mathbf{U = E}$
- $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

What about weights? MST is minimize, but problem we defined is maximize.

- Let $\mathbf{\bar{w} > w(e)}$ for all $\mathbf{e \in E}$, let $\mathbf{w'(e) = \bar{w} - w(e)}$ for all $\mathbf{e \in E}$

For any tree $\mathbf{T}$:

$$\mathbf{w'(T) = \sum_{e \in T} w'(e) = \sum_{e \in T} (\bar{w} - w(e)) = (n-1)\bar{w} - \sum_{e \in T} w(e)}$$

So under weights $\mathbf{w'}$, max-weight IS $=$ max-weight forest $=$ max-weight tree $=$ min-weight tree (weights $\mathbf{w}$)

- So finding max-weight forest $=$ finding min spanning tree.

Useful Properties of Forests

Let $\mathbf{U} = \mathbf{E}$ and $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

Useful properties:

1. $\emptyset \in \mathcal{I}$
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$

Useful Properties of Forests

Let $\mathbf{U} = \mathbf{E}$ and $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

Useful properties:

1. $\emptyset \in \mathcal{I}$
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$
3. *Augmentation Property*: If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some edge $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

Useful Properties of Forests

Let $\mathbf{U} = \mathbf{E}$ and $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

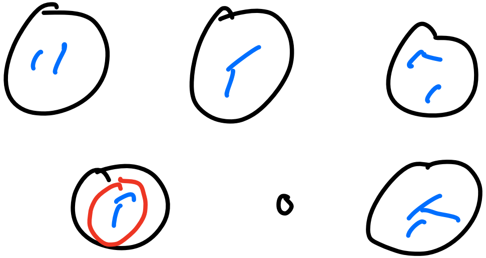
Useful properties:

1. $\emptyset \in \mathcal{I}$
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$
3. *Augmentation Property*: If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some edge $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

Proof Sketch that Forests have Augmentation Property.

Suppose false: no edge in $\mathbf{F_2} \setminus \mathbf{F_1}$ can be added to $\mathbf{F_1}$. Let $\mathbf{c_1} = \#$ components in $\mathbf{F_1}$, $\mathbf{c_2} = \#$ components in $\mathbf{F_2}$

Useful Properties of Forests



Let $\mathbf{U} = \mathbf{E}$ and $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

Useful properties:

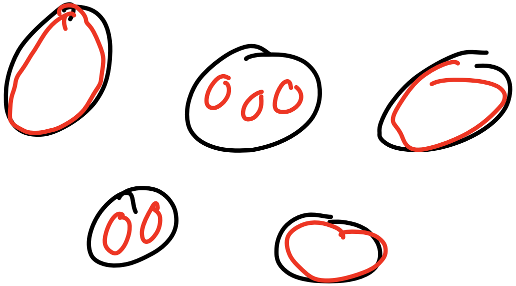
1. $\emptyset \in \mathcal{I}$
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$
3. *Augmentation Property*: If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some edge $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

Proof Sketch that Forests have Augmentation Property.

Suppose false: no edge in $\mathbf{F_2} \setminus \mathbf{F_1}$ can be added to $\mathbf{F_1}$. Let $\mathbf{c_1}$ = # components in $\mathbf{F_1}$, $\mathbf{c_2}$ = # components in $\mathbf{F_2}$

$\Longrightarrow$ every edge of $\mathbf{F_2}$ has both endpoints in same component of $\mathbf{F_1}$

Useful Properties of Forests



Let $\mathbf{U} = \mathbf{E}$ and $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

Useful properties:

1. $\varnothing \in \mathcal{I}$
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$
3. *Augmentation Property*: If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some edge $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

Proof Sketch that Forests have Augmentation Property.

Suppose false: no edge in $\mathbf{F_2} \setminus \mathbf{F_1}$ can be added to $\mathbf{F_1}$. Let $\mathbf{c_1}$ = # components in $\mathbf{F_1}$, $\mathbf{c_2}$ = # components in $\mathbf{F_2}$

$\implies$ every edge of $\mathbf{F_2}$ has both endpoints in same component of $\mathbf{F_1}$

$\implies$ every component of $\mathbf{F_2}$ contained in component of $\mathbf{F_1}$ $\implies$ $\mathbf{c_2} \geq \mathbf{c_1}$

Useful Properties of Forests

Let $\mathbf{U} = \mathbf{E}$ and $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F})$ a forest$\}$

Useful properties:

1. $\emptyset \in \mathcal{I}$
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$
3. *Augmentation Property*: If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some edge $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

Proof Sketch that Forests have Augmentation Property.

Suppose false: no edge in $\mathbf{F_2} \setminus \mathbf{F_1}$ can be added to $\mathbf{F_1}$. Let $\mathbf{c_1} = \#$ components in $\mathbf{F_1}$, $\mathbf{c_2} = \#$ components in $\mathbf{F_2}$

$\implies$ every edge of $\mathbf{F_2}$ has both endpoints in same component of $\mathbf{F_1}$

$\implies$ every component of $\mathbf{F_2}$ contained in component of $\mathbf{F_1}$ $\implies$ $\mathbf{c_2} \geq \mathbf{c_1}$

But $\mathbf{c_2} = \mathbf{n} - |\mathbf{F_2}| < \mathbf{n} - |\mathbf{F_1}| = \mathbf{c_1}$.

Useful Properties of Forests

Let $\mathbf{U} = \mathbf{E}$ and $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{E} : (\mathbf{V}, \mathbf{F}) \text{ a forest}\}$

Useful properties:

1. $\emptyset \in \mathcal{I}$
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$
3. *Augmentation Property*: If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some edge $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

Proof Sketch that Forests have Augmentation Property.

Suppose false: no edge in $\mathbf{F_2} \setminus \mathbf{F_1}$ can be added to $\mathbf{F_1}$. Let $\mathbf{c_1}$ = # components in $\mathbf{F_1}$, $\mathbf{c_2}$ = # components in $\mathbf{F_2}$

$\implies$ every edge of $\mathbf{F_2}$ has both endpoints in same component of $\mathbf{F_1}$

$\implies$ every component of $\mathbf{F_2}$ contained in component of $\mathbf{F_1}$ $\implies$ $\mathbf{c_2} \geq \mathbf{c_1}$

But $\mathbf{c_2} = \mathbf{n} - |\mathbf{F_2}| < \mathbf{n} - |\mathbf{F_1}| = \mathbf{c_1}$.

Contradiction. □

Matroids

Definition

$(\mathbf{U}, \mathcal{I})$ is a *matroid* if the following three properties hold:

1. $\emptyset \in \mathcal{I}$,
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$, and
3. If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some element $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

Matroids

Definition

$(\mathbf{U}, \mathcal{I})$ is a *matroid* if the following three properties hold:

1. $\emptyset \in \mathcal{I}$,
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$, and
3. If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some element $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

$(\mathbf{U}, \mathcal{I})$ is a *hereditary set system* if the first two properties hold.

Matroids

Definition

$(\mathbf{U}, \mathcal{I})$ is a *matroid* if the following three properties hold:

1. $\emptyset \in \mathcal{I}$,
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$, and
3. If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some element $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

$(\mathbf{U}, \mathcal{I})$ is a *hereditary set system* if the first two properties hold.

Matroid theory: super interesting area of combinatorics! Surprising amount of structure.

Matroids

Definition

$(\mathbf{U}, \mathcal{I})$ is a *matroid* if the following three properties hold:

1. $\emptyset \in \mathcal{I}$,
2. If $\mathbf{F} \in \mathcal{I}$ and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}' \in \mathcal{I}$, and
3. If $\mathbf{F_1} \in \mathcal{I}$ and $\mathbf{F_2} \in \mathcal{I}$ with $|\mathbf{F_2}| > |\mathbf{F_1}|$, then there is some element $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e}\} \in \mathcal{I}$.

$(\mathbf{U}, \mathcal{I})$ is a *hereditary set system* if the first two properties hold.

Matroid theory: super interesting area of combinatorics! Surprising amount of structure.

Warmup: In any matroid, the maximal independent sets (called bases) have the same size (called the rank of the matroid).

Examples of Matroids

- Forests in graphs

Examples of Matroids

- Forests in graphs
- Linearly independent vectors in vector space
  - $\mathbf{U}$ a finite set of vectors in $\mathbb{R}^{\mathbf{d}}$
  - $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{U} : \mathbf{F} \text{ linearly independent}\}$

Examples of Matroids

- Forests in graphs
- Linearly independent vectors in vector space
  - $\mathbf{U}$ a finite set of vectors in $\mathbb{R}^{\mathbf{d}}$
  - $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{U} : \mathbf{F}$ linearly independent$\}$
  - $\varnothing$ linearly independent
  - If $\mathbf{F}$ linearly independent and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}'$ linearly independent
  - Augmentation: if $\mathbf{F_1}$ linearly independent, $\mathbf{F_2}$ linearly independent, and $|\mathbf{F_2}| > |\mathbf{F_1}| \implies \dim(\text{span}(\mathbf{F_1})) = |\mathbf{F_1}| < |\mathbf{F_2}| = \dim(\text{span}(\mathbf{F_2}))$

Examples of Matroids

- Forests in graphs
- Linearly independent vectors in vector space
  - $\mathbf{U}$ a finite set of vectors in $\mathbb{R}^{\mathbf{d}}$
  - $\mathcal{I} = \{\mathbf{F} \subseteq \mathbf{U} : \mathbf{F}$ linearly independent$\}$
  - $\varnothing$ linearly independent
  - If $\mathbf{F}$ linearly independent and $\mathbf{F}' \subseteq \mathbf{F}$, then $\mathbf{F}'$ linearly independent
  - Augmentation: if $\mathbf{F_1}$ linearly independent, $\mathbf{F_2}$ linearly independent, and $|\mathbf{F_2}| > |\mathbf{F_1}| \implies \dim(\text{span}(\mathbf{F_1})) = |\mathbf{F_1}| < |\mathbf{F_2}| = \dim(\text{span}(\mathbf{F_2}))$

Matroids: generalize both graph theory and linear algebra!

- Originally invented by Whitney as an attempt to generalize the concept of "linear independence"

Representation

To do algorithms with matroids, need to figure out how they're represented.

Representation

To do algorithms with matroids, need to figure out how they're represented.

Option 1: list all independent sets

Representation

To do algorithms with matroids, need to figure out how they're represented.

Option 1: list all independent sets

- Too many of them!

Representation

To do algorithms with matroids, need to figure out how they're represented.

Option 1: list all independent sets

- Too many of them!

What did we need for MST (Kruskal)?

Representation

To do algorithms with matroids, need to figure out how they're represented.

Option 1: list all independent sets

- Too many of them!

What did we need for MST (Kruskal)?

Independence Oracle: algorithm which take $\mathbf{F} \subseteq \mathbf{U}$, returns YES if $\mathbf{F} \in \mathcal{I}$, NO if $\mathbf{F} \notin \mathcal{I}$

Representation

To do algorithms with matroids, need to figure out how they're represented.

Option 1: list all independent sets

- Too many of them!

What did we need for MST (Kruskal)?

Independence Oracle: algorithm which take $\mathbf{F} \subseteq \mathbf{U}$, returns YES if $\mathbf{F} \in \mathcal{I}$, NO if $\mathbf{F} \notin \mathcal{I}$

For MST: "does **F** have any cycles"? Independence oracle: DFS/BFS

Representation

To do algorithms with matroids, need to figure out how they're represented.

Option 1: list all independent sets

- Too many of them!

What did we need for MST (Kruskal)?

Independence Oracle: algorithm which take $\mathbf{F} \subseteq \mathbf{U}$, returns YES if $\mathbf{F} \in \mathcal{I}$, NO if $\mathbf{F} \notin \mathcal{I}$

For MST: "does **F** have any cycles"? Independence oracle: DFS/BFS

We'll assume we have independence oracle.

Greedy Algorithm

Kruskal, generalized to matroids (and max weight)!

Greedy Algorithm

Kruskal, generalized to matroids (and max weight)!

$\mathbf{F} = \varnothing$
Sort $\mathbf{U}$ by weight (largest to smallest)
For each $\mathbf{u} \in \mathbf{U}$ in sorted order $\{$
    If $\mathbf{F} \cup \{\mathbf{u}\} \in \mathcal{I}$, add $\mathbf{u}$ to $\mathbf{F}$
$\}$
Return $\mathbf{F}$

Correctness

Theorem

Let $\mathbf{F}$ *be independent set returned by greedy. Then* $\mathbf{w(F) \geq w(F')}$ *for all* $\mathbf{F'} \in \mathcal{I}$.

Correctness

Theorem

Let $\mathbf{F}$ *be independent set returned by greedy. Then* $\mathbf{w(F) \geq w(F')}$ *for all* $\mathbf{F'} \in \mathcal{I}$.

- $\mathbf{F = \{f_1, f_2, \ldots, f_r\}}$, where $\mathbf{w(f_i) \geq w(f_{i+1})}$ for all $\mathbf{i}$ (order added by greedy)
- $\mathbf{F' = \{e_1, e_2, \ldots, e_r\}}$ where $\mathbf{w(e_i) \geq w(e_{i+1})}$ for all $\mathbf{i}$

Correctness

Theorem

Let $\mathbf{F}$ *be independent set returned by greedy. Then* $\mathbf{w(F) \geq w(F')}$ *for all* $\mathbf{F'} \in \mathcal{I}$.

- $\mathbf{F = \{f_1, f_2, \ldots, f_r\}}$, where $\mathbf{w(f_i) \geq w(f_{i+1})}$ for all $\mathbf{i}$ (order added by greedy)
- $\mathbf{F' = \{e_1, e_2, \ldots, e_r\}}$ where $\mathbf{w(e_i) \geq w(e_{i+1})}$ for all $\mathbf{i}$

Claim: $\mathbf{w(f_i) \geq w(e_i)}$ for all $\mathbf{i}$.

Correctness

Theorem

Let $\mathbf{F}$ *be independent set returned by greedy. Then* $\mathbf{w(F) \geq w(F')}$ *for all* $\mathbf{F'} \in \mathcal{I}$.

- $\mathbf{F = \{f_1, f_2, \dots, f_r\}}$, where $\mathbf{w(f_i) \geq w(f_{i+1})}$ for all $\mathbf{i}$ (order added by greedy)
- $\mathbf{F' = \{e_1, e_2, \dots, e_r\}}$ where $\mathbf{w(e_i) \geq w(e_{i+1})}$ for all $\mathbf{i}$

Claim: $\mathbf{w(f_i) \geq w(e_i)}$ for all $\mathbf{i}$.

Proof: Suppose false, let $\mathbf{j}$ smallest integer such that $\mathbf{w(f_j) < w(e_j)}$.

Correctness

Theorem

Let $\mathbf{F}$ *be independent set returned by greedy. Then* $\mathbf{w(F) \geq w(F')}$ *for all* $\mathbf{F'} \in \mathcal{I}$.

- $\mathbf{F} = \{\mathbf{f_1, f_2, \ldots, f_r}\}$, where $\mathbf{w(f_i) \geq w(f_{i+1})}$ for all $\mathbf{i}$ (order added by greedy)
- $\mathbf{F'} = \{\mathbf{e_1, e_2, \ldots, e_r}\}$ where $\mathbf{w(e_i) \geq w(e_{i+1})}$ for all $\mathbf{i}$

Claim: $\mathbf{w(f_i) \geq w(e_i)}$ for all $\mathbf{i}$.

Proof: Suppose false, let $\mathbf{j}$ smallest integer such that $\mathbf{w(f_j) < w(e_j)}$.

Let $\mathbf{F_1} = \{\mathbf{f_1, \ldots, f_{j-1}}\}$ and let $\mathbf{F_2} = \{\mathbf{e_1, \ldots, e_j}\}$

Correctness

Theorem

Let $\mathbf{F}$ be independent set returned by greedy. Then $\mathbf{w(F) \geq w(F')}$ for all $\mathbf{F'} \in \mathcal{I}$.

- $\mathbf{F} = \{\mathbf{f_1, f_2, \dots, f_r}\}$, where $\mathbf{w(f_i) \geq w(f_{i+1})}$ for all $\mathbf{i}$ (order added by greedy)
- $\mathbf{F'} = \{\mathbf{e_1, e_2, \dots, e_r}\}$ where $\mathbf{w(e_i) \geq w(e_{i+1})}$ for all $\mathbf{i}$

Claim: $\mathbf{w(f_i) \geq w(e_i)}$ for all $\mathbf{i}$.

Proof: Suppose false, let $\mathbf{j}$ smallest integer such that $\mathbf{w(f_j) < w(e_j)}$.

Let $\mathbf{F_1} = \{\mathbf{f_1, \dots, f_{j-1}}\}$ and let $\mathbf{F_2} = \{\mathbf{e_1, \dots, e_j}\}$

$|\mathbf{F_2}| > |\mathbf{F_1}|$, so by augmentation there is some $\mathbf{e_z} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e_z}\} \in \mathcal{I}$.

Correctness

Theorem

Let $\mathbf{F}$ *be independent set returned by greedy. Then* $\mathbf{w(F) \geq w(F')}$ *for all* $\mathbf{F'} \in \mathcal{I}$.

- $\mathbf{F = \{f_1, f_2, \dots, f_r\}}$, where $\mathbf{w(f_i) \geq w(f_{i+1})}$ for all $\mathbf{i}$ (order added by greedy)
- $\mathbf{F' = \{e_1, e_2, \dots, e_r\}}$ where $\mathbf{w(e_i) \geq w(e_{i+1})}$ for all $\mathbf{i}$

Claim: $\mathbf{w(f_i) \geq w(e_i)}$ for all $\mathbf{i}$.

Proof: Suppose false, let $\mathbf{j}$ smallest integer such that $\mathbf{w(f_j) < w(e_j)}$.

Let $\mathbf{F_1 = \{f_1, \dots, f_{j-1}\}}$ and let $\mathbf{F_2 = \{e_1, \dots, e_j\}}$

$|\mathbf{F_2}| > |\mathbf{F_1}|$, so by augmentation there is some $\mathbf{e_z} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e_z}\} \in \mathcal{I}$.

$$\mathbf{w(e_z) \geq w(e_j) > w(f_j)}$$

Correctness

Theorem

Let $\mathbf{F}$ *be independent set returned by greedy. Then* $\mathbf{w(F) \geq w(F')}$ *for all* $\mathbf{F'} \in \mathcal{I}$.

- $\mathbf{F} = \{\mathbf{f_1, f_2, \ldots, f_r}\}$, where $\mathbf{w(f_i) \geq w(f_{i+1})}$ for all $\mathbf{i}$ (order added by greedy)
- $\mathbf{F'} = \{\mathbf{e_1, e_2, \ldots, e_r}\}$ where $\mathbf{w(e_i) \geq w(e_{i+1})}$ for all $\mathbf{i}$

Claim: $\mathbf{w(f_i) \geq w(e_i)}$ for all $\mathbf{i}$.

Proof: Suppose false, let $\mathbf{j}$ smallest integer such that $\mathbf{w(f_j) < w(e_j)}$.

Let $\mathbf{F_1} = \{\mathbf{f_1, \ldots, f_{j-1}}\}$ and let $\mathbf{F_2} = \{\mathbf{e_1, \ldots, e_j}\}$

$|\mathbf{F_2}| > |\mathbf{F_1}|$, so by augmentation there is some $\mathbf{e_z} \in \mathbf{F_2} \setminus \mathbf{F_1}$ such that $\mathbf{F_1} \cup \{\mathbf{e_z}\} \in \mathcal{I}$.

$$\mathbf{w(e_z) \geq w(e_j) > w(f_j)}$$

Contradiction! Greedy would add $\mathbf{e_z}$ next, not $\mathbf{f_j}$.

Converse

So greedy works on matroids. Amazing fact: if greedy works, set system is a matroid!

Converse

So greedy works on matroids. Amazing fact: if greedy works, set system is a matroid!

Theorem

Let $(\mathbf{U}, \mathcal{I})$ be an hereditary set system. If for every weighting $\mathbf{w} : \mathbf{U} \to \mathbb{R}_{\geq 0}$ the greedy algorithm returns a maximum weight independent set, then $(\mathbf{U}, \mathcal{I})$ is a matroid.

Converse

So greedy works on matroids. Amazing fact: if greedy works, set system is a matroid!

Theorem

Let $(\mathbf{U}, \mathcal{I})$ *be an hereditary set system. If for every weighting* $\mathbf{w} : \mathbf{U} \to \mathbb{R}_{\geq 0}$ *the greedy algorithm returns a maximum weight independent set, then* $(\mathbf{U}, \mathcal{I})$ *is a matroid.*

So for hereditary set systems, matroids exactly characterize when the greedy algorithm works!

Proof

Contradiction. Suppose false $\implies$ $(\mathbf{U}, \mathcal{I})$ hereditary but not matroid.

Proof

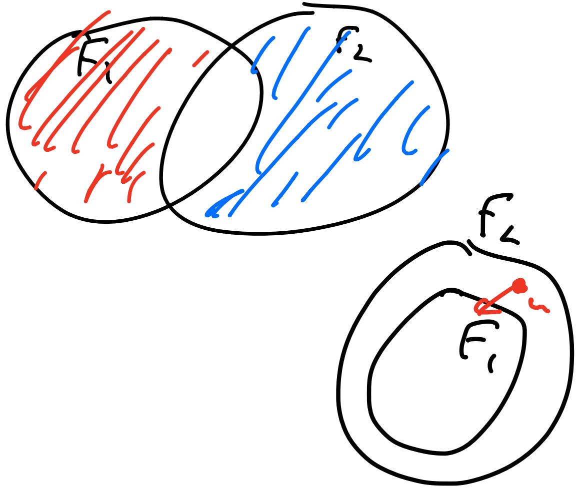
Contradiction. Suppose false $\implies$ $(\mathbf{U}, \mathcal{I})$ hereditary but not matroid.

$\implies \exists \mathbf{F_1}, \mathbf{F_2} \in \mathcal{I}$ such that $|\mathbf{F_1}| < |\mathbf{F_2}|$ but $\mathbf{F_1} \cup \{\mathbf{e}\} \notin \mathcal{I}$ for all $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$

Proof

Contradiction. Suppose false $\implies$ $(\mathbf{U}, \mathcal{I})$ hereditary but not matroid.

$\implies \exists \mathbf{F_1}, \mathbf{F_2} \in \mathcal{I}$ such that $|\mathbf{F_1}| < |\mathbf{F_2}|$ but $\mathbf{F_1} \cup \{\mathbf{e}\} \notin \mathcal{I}$ for all $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$



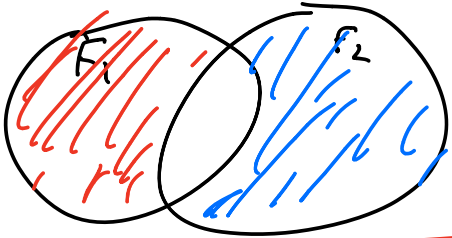
Easy facts:

1. $|\mathbf{F_2} \setminus \mathbf{F_1}| > |\mathbf{F_1} \setminus \mathbf{F_2}|$
2. $|\mathbf{F_2} \setminus \mathbf{F_1}| \geq \mathbf{1}$
3. $|\mathbf{F_1} \setminus \mathbf{F_2}| \geq \mathbf{1}$ (hereditary)

Proof

Contradiction. Suppose false $\implies$ $(\mathbf{U}, \mathcal{I})$ hereditary but not matroid.

$\implies \exists \mathbf{F_1}, \mathbf{F_2} \in \mathcal{I}$ such that $|\mathbf{F_1}| < |\mathbf{F_2}|$ but $\mathbf{F_1} \cup \{\mathbf{e}\} \notin \mathcal{I}$ for all $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$



Easy facts:

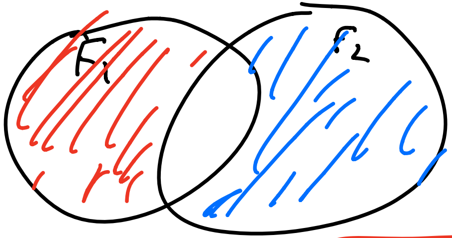
1. $|\mathbf{F_2} \setminus \mathbf{F_1}| > |\mathbf{F_1} \setminus \mathbf{F_2}|$
2. $|\mathbf{F_2} \setminus \mathbf{F_1}| \geq \mathbf{1}$
3. $|\mathbf{F_1} \setminus \mathbf{F_2}| \geq \mathbf{1}$ (hereditary)

$\implies \exists \epsilon > \mathbf{0}$ such that $\mathbf{0} < (\mathbf{1} + \epsilon)|\mathbf{F_1} \setminus \mathbf{F_2}| < |\mathbf{F_2} \setminus \mathbf{F_1}|$

Proof

Contradiction. Suppose false $\implies$ $(\mathbf{U}, \mathcal{I})$ hereditary but not matroid.

$\implies \exists \mathbf{F_1}, \mathbf{F_2} \in \mathcal{I}$ such that $|\mathbf{F_1}| < |\mathbf{F_2}|$ but $\mathbf{F_1} \cup \{\mathbf{e}\} \notin \mathcal{I}$ for all $\mathbf{e} \in \mathbf{F_2} \setminus \mathbf{F_1}$



Easy facts:

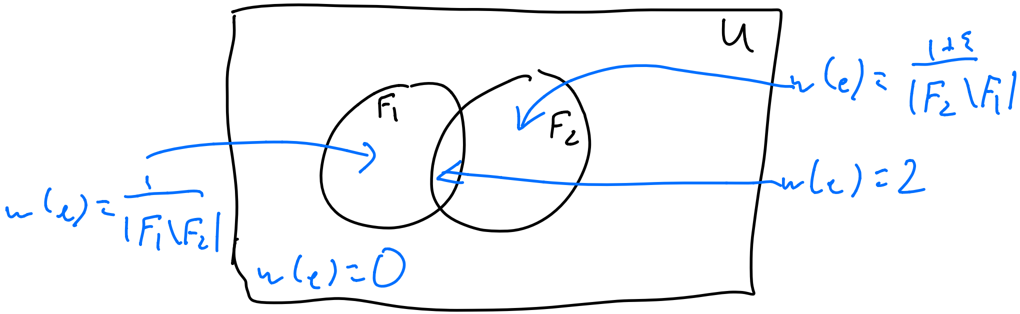
1. $|\mathbf{F_2} \setminus \mathbf{F_1}| > |\mathbf{F_1} \setminus \mathbf{F_2}|$
2. $|\mathbf{F_2} \setminus \mathbf{F_1}| \geq \mathbf{1}$
3. $|\mathbf{F_1} \setminus \mathbf{F_2}| \geq \mathbf{1}$ (hereditary)

$\implies \exists \epsilon > \mathbf{0}$ such that $\mathbf{0} < (\mathbf{1} + \epsilon)|\mathbf{F_1} \setminus \mathbf{F_2}| < |\mathbf{F_2} \setminus \mathbf{F_1}|$

$$\implies \frac{\mathbf{1}}{|\mathbf{F_1} \setminus \mathbf{F_2}|} > \frac{\mathbf{1} + \epsilon}{|\mathbf{F_2} \setminus \mathbf{F_1}|}$$

Proof (cont'd)

Use fact that $\frac{1}{|F_1 \setminus F_2|} > \frac{1+\epsilon}{|F_2 \setminus F_1|}$ to define weights.



Greedy:

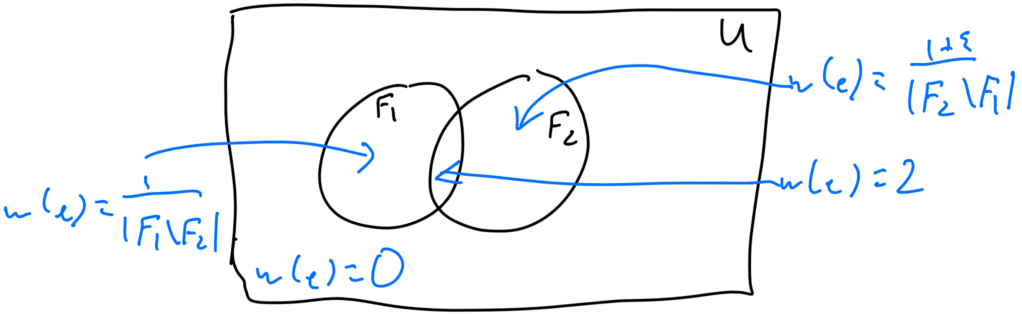
- Adds all of $F_1 \cap F_2$
- Adds all of $F_1 \setminus F_2$
- Can't add any of $F_2 \setminus F_1$

$$\mathbf{w}(\text{greedy}) = 2|F_1 \cap F_2| + |F_1 \setminus F_2| \frac{1}{|F_1 \setminus F_2|}$$

$$= 2|F_1 \cap F_2| + 1$$

Proof (cont'd)

Use fact that $\frac{1}{|F_1 \setminus F_2|} > \frac{1+\epsilon}{|F_2 \setminus F_1|}$ to define weights.



Greedy:

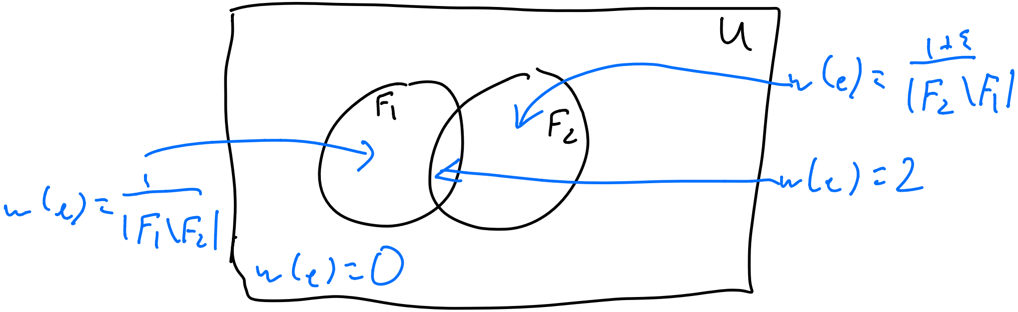
- Adds all of $\mathbf{F_1 \cap F_2}$
- Adds all of $\mathbf{F_1 \setminus F_2}$
- Can't add any of $\mathbf{F_2 \setminus F_1}$

$$\mathbf{w}(\text{greedy}) = 2|F_1 \cap F_2| + |F_1 \setminus F_2| \frac{1}{|F_1 \setminus F_2|} = 2|F_1 \cap F_2| + 1$$

$$\mathbf{w}(F_2) = 2|F_1 \cap F_2| + |F_2 \setminus F_1| \frac{1+\epsilon}{|F_2 \setminus F_1|} = 2|F_1 \cap F_2| + 1 + \epsilon$$

Proof (cont'd)

Use fact that $\frac{1}{|F_1 \setminus F_2|} > \frac{1+\epsilon}{|F_2 \setminus F_1|}$ to define weights.



Greedy:

- Adds all of $F_1 \cap F_2$
- Adds all of $F_1 \setminus F_2$
- Can't add any of $F_2 \setminus F_1$

$$\mathbf{w}(\text{greedy}) = 2|F_1 \cap F_2| + |F_1 \setminus F_2| \frac{1}{|F_1 \setminus F_2|} = 2|F_1 \cap F_2| + 1$$

$$\mathbf{w}(F_2) = 2|F_1 \cap F_2| + |F_2 \setminus F_1| \frac{1+\epsilon}{|F_2 \setminus F_1|} = 2|F_1 \cap F_2| + 1 + \epsilon$$

Greedy not optimal: contradiction!