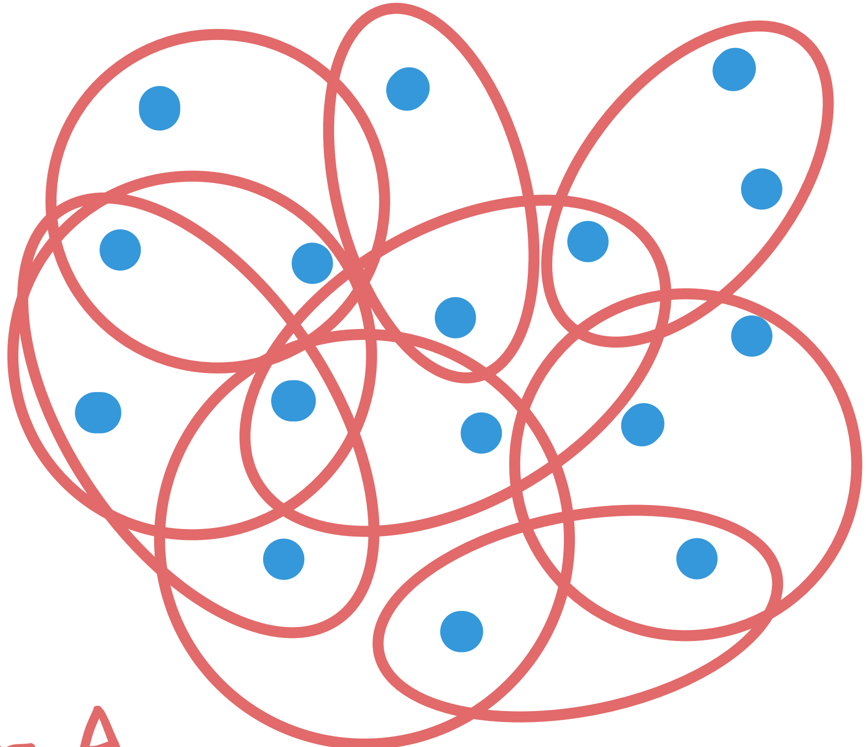
Randomized Greedy

Maximum Coverage



m points $[m] \equiv \{1, \dots, m\}$

n sets $A_1, \dots, A_n \subseteq [m]$

cardinality constraint $k \in \mathbb{N}$

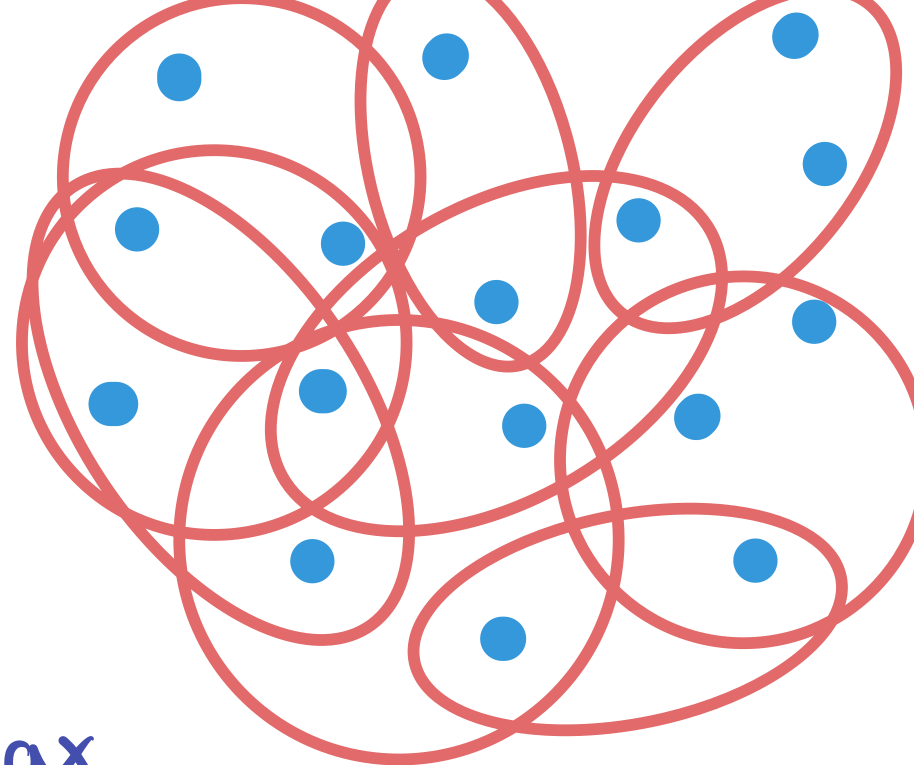
Goal: select k sets $A_{e_1}, A_{e_2}, \dots, A_{e_k}$

maximizing the union $|A_{e_1} \cup A_{e_2} \cup \dots \cup A_{e_k}|$

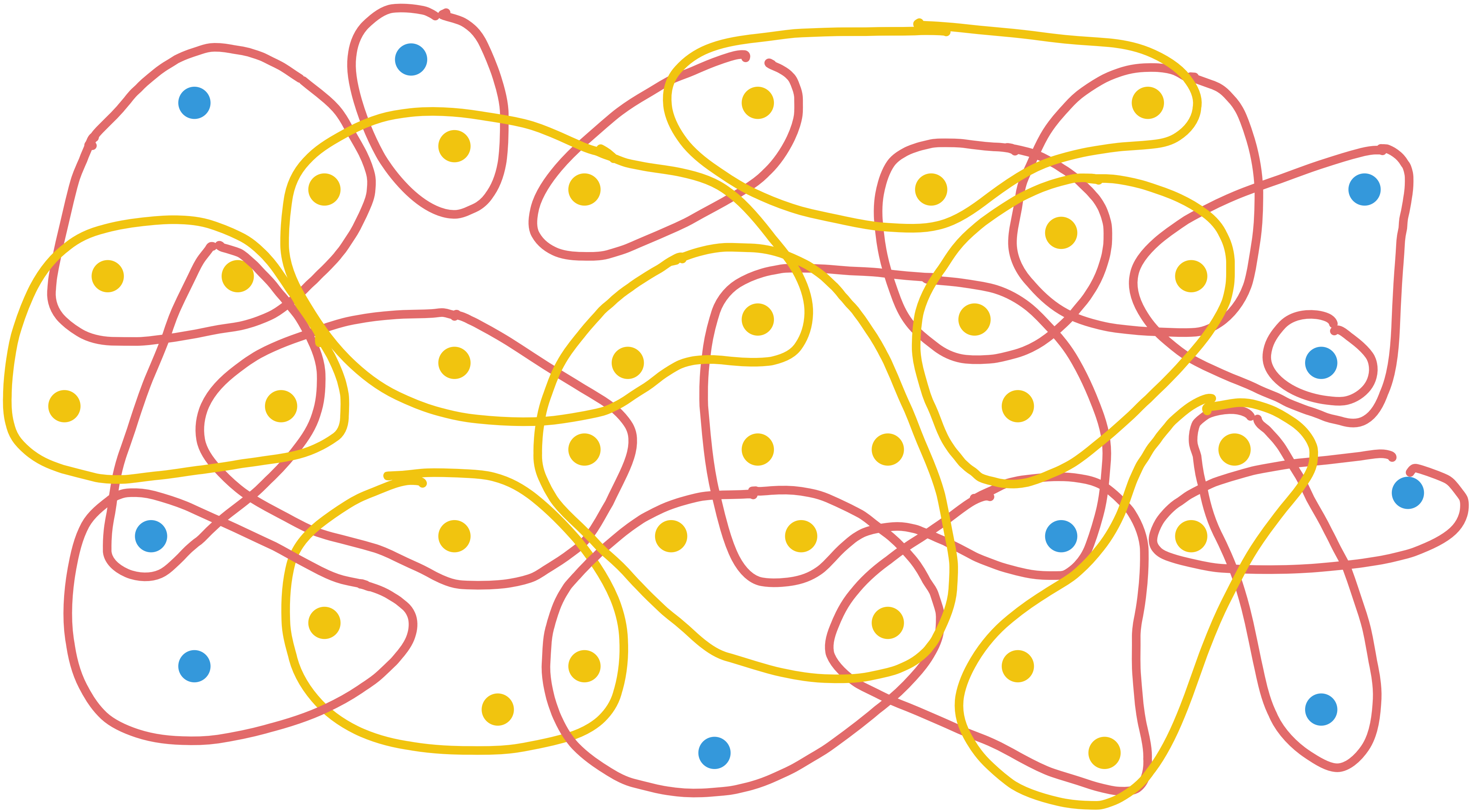
NP-Hard via set cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $A_1, \dots, A_n \subseteq [m]$



Set Cover (next time)	Max Coverage
minimize # sets to cover all points	maximize # points covered w/ $\leq k$ sets (packing constraint)
NP-Hard (via SAT) $\overset{\text{(why?)}}{\Longrightarrow}$	NP-Hard
LP + randomized rounding	greedy algorithm



greedy algorithm repeatedly selects the set covering the most new points

how good is the greedy algorithm?
(approximation ratio)

Set function $f: 2^N \to \mathbb{R}_{\geq 0}$ s.t.

- **Normalized:** $f(\emptyset) = 0$
- **Nonnegative:** $f(S) \geq 0 \quad \forall S$
- **Monotone:** $f(S) \leq f(T)$ for $S \subseteq T$
- **Submodular:** for $A, B \subseteq N$,

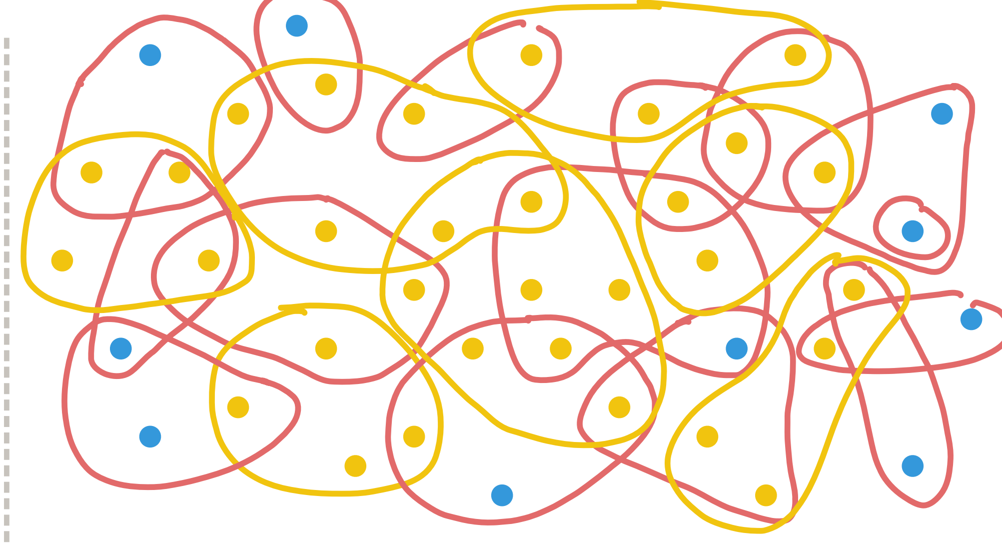
$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$$

equivalently:

"decreasing marginal values"

for $S \subseteq T \subseteq N$, $e \in N$,

$$\underbrace{f(S+e) - f(S)}_{\stackrel{\text{def}}{=} f(e|S)} \geq \underbrace{f(T+e) - f(T)}_{\stackrel{\text{def}}{=} f(e|T)}$$

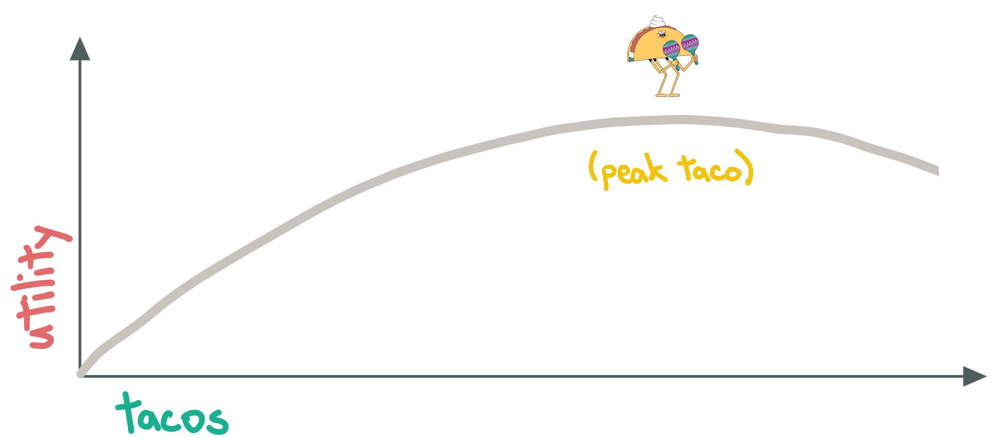


m points $[m] \equiv \{1, \dots, m\}$

n sets $A_1, \dots, A_n \subseteq [m]$

$S \subseteq [n]$ interpreted as collection of sets (via indices)

$$f(S) \stackrel{\text{def}}{=} \begin{pmatrix} \text{\# points covered by} \\ \text{sets (w/ index in) } S \end{pmatrix}$$



greedy$(f, \mathcal{N}, k)$

1. $S_0 \leftarrow \emptyset$.
2. For $i = 1, \ldots, k$:
   A. Let $e_i \in \mathcal{N}$ maximize $f(e_i \mid S_{i-1})$
   B. $S_i \leftarrow S_i + e$.
3. Return S_k.

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$ s.t.

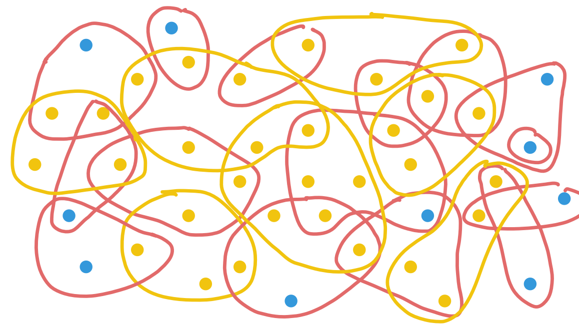
- Normalized: $f(\emptyset) = 0$
- Nonnegative: $f(S) \geq 0 \quad \forall S$
- Monotone: $f(S) \leq f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$,

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,

$$\underbrace{f(S+e) - f(S)}_{\overset{\text{def}}{=} f(e \mid S)} \geq \underbrace{f(T+e) - f(T)}_{\overset{\text{def}}{=} f(e \mid T)}$$



m points $[m] = \{1, \ldots, m\}$

n sets $A_1, \ldots, A_n \subseteq [m]$

$S \subseteq [n]$ interpreted as collection of sets (via indices)

$$f(S) \overset{\text{def}}{=} \begin{pmatrix} \text{\# points covered by} \\ \text{sets (w/ index in) } S \end{pmatrix}$$

Lemma $f(e_i \mid S_{i-1}) \geq \frac{1}{k}\left(\mathrm{OPT} - f(S_{i-1})\right)$

Lemma $f(e_i \mid S_{i-1}) \geq \frac{1}{k}(OPT - f(S_{i-1}))$

Set function $f: 2^N \to \mathbb{R}_{\geq 0}$ s.t.

- **Normalized:** $f(\emptyset) = 0$
- **Nonnegative:** $f(S) \geq 0 \quad \forall S$
- **Monotone:** $f(S) \leq f(T)$ for $S \subseteq T$
- **Submodular:** for $A, B \subseteq \mathcal{N}$,

$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,

$$\underbrace{f(S+e) - f(S)}_{\overset{\text{def}}{=} f(e \mid S)} \geq \underbrace{f(T+e) - f(T)}_{\overset{\text{def}}{=} f(e \mid T)}$$

greedy(f, $\mathcal{N}$, k)

1. $S_0 \leftarrow \emptyset$.
2. For $i = 1, \ldots, k$:
   - A. Let $e_i \in \mathcal{N}$ maximize $f(e_i \mid S_{i-1})$
   - B. $S_i \leftarrow S_i + e$.
3. Return S_k.

Lemma $f(e_i \mid S_{i-1}) \geq \frac{1}{k}(OPT - f(S_{i-1}))$

Theorem $f(S_k) \geq (1-\frac{1}{e})OPT$

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$ s.t.

- Normalized: $f(\emptyset)=0$
- Nonnegative: $f(S) \geq 0 \quad \forall S$
- Monotone: $f(S) \leq f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$,

$$f(A)+f(B) \geq f(A \cup B) + f(A \cap B)$$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,

$$\underbrace{f(S+e)-f(S)}_{\overset{\text{def}}{=} f(e|S)} \geq \underbrace{f(T+e)-f(T)}_{\overset{\text{def}}{=} f(e|T)}$$

greedy$(f, \mathcal{N}, k)$

1. $S_0 \leftarrow \emptyset$.
2. For $i = 1, \ldots, k$:
   - A. Let $e_i \in \mathcal{N}$ maximize $f(e_i \mid S_{i-1})$
   - B. $S_i \leftarrow S_i + e$.
3. Return S_k.

Monotone Submodular f

Lemma $f(e_i | S_{i-1}) \geq \frac{1}{k}(OPT - f(S_{i-1}))$

Theorem $f(S_k) \geq (1 - \frac{1}{e}) OPT$

Suppose f was not monotone
does greedy work?

$$f(S^* | S_i) = f(S^* \cup S_i) - f(S_i)$$

$$\overset{\text{monotonicity}}{\geq} f(S^*) - f(S_i)$$

f not monotone: $f(S^* \cup S_i) \not\geq f(S^*)$

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$ s.t.

- Normalized: $f(\emptyset) = 0$
- Nonnegative: $f(S) \geq 0 \quad \forall S$
- Monotone: $f(S) \leq f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$,

$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,

$$\underbrace{f(S+e) - f(S)}_{\overset{\text{def}}{=} f(e|S)} \geq \underbrace{f(T+e) - f(T)}_{\overset{\text{def}}{=} f(e|T)}$$

<u>greedy(f, $\mathcal{N}$, k)</u>

1. $S_0 \leftarrow \emptyset$.
2. For $i = 1, \ldots, k$:
   A. Let $e_i \in \mathcal{N}$ maximize $f(e_i \mid S_{i-1})$
   B. $S_i \leftarrow S_i + e$.
3. Return S_k.

Monotone Submodular f

Lemma $f(e_i | S_{i-1}) \geq \frac{1}{k}(OPT - f(S_{i-1}))$

Theorem $f(S_k) \geq (1-\frac{1}{e})OPT$

Suppose f was not monotone

Lemma $R \subseteq \mathcal{N}$ random, $P[e \in R] \leq p \;\; \forall e$
then
$$E[f(R)] \geq (1-p) f(\emptyset)$$

e.g. if S_i random, $P[e \in S_i] \leq 1/2 \;\; \forall e$

$$E[f(S^* \cup S_i)] \geq \frac{1}{2} f(S^*)$$

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$ s.t.

- Normalized: $f(\emptyset) = 0$
- Nonnegative: $f(S) \geq 0 \;\; \forall S$
- Monotone: $f(S) \leq f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$,
  $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,
$$\underbrace{f(S+e) - f(S)}_{\overset{\text{def}}{=} f(e|S)} \geq \underbrace{f(T+e) - f(T)}_{\overset{\text{def}}{=} f(e|T)}$$

---

greedy$(f, \mathcal{N}, k)$

1. $S_0 \leftarrow \emptyset$.
2. For $i = 1, \ldots, k$:
   A. Let $e_i \in \mathcal{N}$ maximize $f(e_i \mid S_{i-1})$
   B. $S_i \leftarrow S_i + e$.
3. Return S_k.

Suppose f was not monotone

Lemma $R \subseteq \mathcal{N}$ random, $P[e \in R] \leq p \ \forall e$ then $E[f(R)] \geq (1-p) f(\emptyset)$

Continuous extensions of f

$F : [0,1]^{\mathcal{N}} \to \mathbb{R}$ extends $f : 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$

if $F(\mathbb{1}_S) = f(S) \ \forall S \subseteq \mathcal{N}$

Convex closure F^-

$F^-(p) \stackrel{\text{def}}{=} \inf \{ E[f(R)] \text{ s.t. } P[e \in R] = p \ \forall e \}$

Lovász:

$F^-(p) = E[f(S_t)]$ w/ $S_t = \{ e : p_e \geq t \}$,

$t \in [0,1]$ unif. random

Set function $f : 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$ s.t.

- Normalized: $f(\emptyset) = 0$
- Nonnegative: $f(S) \geq 0 \ \forall S$
- Monotone: $f(S) \leq f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$, $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,

$$\underbrace{f(S+e) - f(S)}_{\stackrel{\text{def}}{=} f(e|S)} \geq \underbrace{f(T+e) - f(T)}_{\stackrel{\text{def}}{=} f(e|T)}$$

greedy($f, \mathcal{N}, k$)

1. $S_0 \leftarrow \emptyset$.
2. For $i = 1, \ldots, k$:
   A. Let $e_i \in \mathcal{N}$ maximize $f(e_i \mid S_{i-1})$
   B. $S_i \leftarrow S_i + e$.
3. Return S_k.

Convex closure F^-

$$F^-(p) \stackrel{\text{def}}{=} \inf\{E[f(R)] \text{ s.t. } P[e \in R] = p \ \forall e\}$$

Lovász: $F^-(p) = E[f(S_t)]$

w/ $S_t = \{e : p_e \geq t\}$, $t \in [0,1]$ unif. random

proof:

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$ s.t.

- Normalized: $f(\emptyset) = 0$
- Nonnegative: $f(S) \geq 0 \quad \forall S$
- Monotone: $f(S) \leq f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$,

$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,

$$\underbrace{f(S+e) - f(S)}_{\stackrel{\text{def}}{=} f(e|S)} \geq \underbrace{f(T+e) - f(T)}_{\stackrel{\text{def}}{=} f(e|T)}$$

greedy$(f, \mathcal{N}, k)$

1. $S_0 \leftarrow \emptyset$.
2. For $i = 1, \ldots, k$:
   - A. Let $e_i \in \mathcal{N}$ maximize $f(e_i \mid S_{i-1})$
   - B. $S_i \leftarrow S_i + e$.
3. Return S_k.

Convex closure F^-

$$F^-(p) \stackrel{\text{def}}{=} \inf\{E[f(R)] \text{ s.t. } P[e\in R]=p \;\forall e\}$$

Lovász: $F^-(p) = E[f(S_t)]$

w/ $S_t = \{e : p_e \geq t\}$, $t\in[0,1]$ unif. random

Suppose f was not monotone

Lemma $R\subseteq \mathcal{N}$ random, $P[e\in R]\leq p$ $\forall e$

then $E[f(R)] \geq (1-p) f(\emptyset)$

proof

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$ s.t.

- Normalized: $f(\emptyset)=0$
- Nonnegative: $f(S)\geq 0$ $\forall S$
- Monotone: $f(S)\leq f(T)$ for $S\subseteq T$
- Submodular: for $A,B\subseteq \mathcal{N}$,

$f(A)+f(B)\geq f(A\cup B)+f(A\cap B)$

"decreasing marginal values"

for $S\subseteq T\subseteq \mathcal{N}$, $e\in\mathcal{N}$,

$$\underbrace{f(S+e)-f(S)}_{\stackrel{\text{def}}{=} f(e|S)} \geq \underbrace{f(T+e)-f(T)}_{\stackrel{\text{def}}{=} f(e|T)}$$

greedy$(f,\mathcal{N},k)$

1. $S_0 \leftarrow \emptyset$.
2. For $i=1,\dots,k$:
   - A. Let $e_i\in\mathcal{N}$ maximize $f(e_i \mid S_{i-1})$
   - B. $S_i \leftarrow S_i + e$.
3. Return S_k.

Suppose f was not monotone

Lemma $R \subseteq \mathcal{N}$ random, $P[e \in R] \leq p$ $\forall e$

then $E[f(R)] \geq (1-p) f(\emptyset)$

`randomized-greedy(`f`,`k`)`

1. $S_0 \leftarrow \emptyset$
2. For $i = 1, \dots, k$:
   - A. Let $R_i \subseteq \mathcal{N}$ be the k elements e w/ max $f(e \mid S_{i-1})$.
   - B. Sample $e_i \sim R$ uniformly at random.
   - C. $S_i \leftarrow S_{i-1} + e_i$.
3. Return S_k.

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$ s.t.

- Normalized: $f(\emptyset) = 0$
- Nonnegative: $f(S) \geq 0$ $\forall S$
- Monotone: $f(S) \leq f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$,

$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,

$$\underbrace{f(S+e) - f(S)}_{\overset{\text{def}}{=} f(e|S)} \geq \underbrace{f(T+e) - f(T)}_{\overset{\text{def}}{=} f(e|T)}$$

`greedy(`f`,`$\mathcal{N}$`,`k`)`

1. $S_0 \leftarrow \emptyset$.
2. For $i = 1, \dots, k$:
   - A. Let $e_i \in \mathcal{N}$ maximize $f(e_i \mid S_{i-1})$
   - B. $S_i \leftarrow S_i + e$.
3. Return S_k.

Lemma $R \subseteq \mathcal{N}$ random, $P[e \in R] \le p$ $\forall e$
then $E[f(R)] \ge (1-p) f(\emptyset)$

randomized-greedy(f, k)

1. $S_0 \leftarrow \emptyset$
2. For $i = 1, \ldots, k$:
   A. Let $R_i \subseteq \mathcal{N}$ be the k elements e w/ max $f(e \mid S_{i-1})$.
   B. Sample $e_i \sim R$ uniformly at random.
   C. $S_i \leftarrow S_{i-1} + e_i$.
3. Return S_k.

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\ge 0}$ s.t.
- Normalized: $f(\emptyset) = 0$
- Nonnegative: $f(S) \ge 0$ $\forall S$
- Monotone: $f(S) \le f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$,
  $f(A) + f(B) \ge f(A \cup B) + f(A \cap B)$

"decreasing marginal values"
for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,
$$\underbrace{f(S+e) - f(S)}_{\stackrel{\text{def}}{=} f(e \mid S)} \ge \underbrace{f(T+e) - f(T)}_{\stackrel{\text{def}}{=} f(e \mid T)}$$

Lemma $R \subseteq \mathcal{N}$ random, $P[e \in R] \le p$ $\forall e$
then $E[f(R)] \ge (1-p) f(\emptyset)$

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\ge 0}$ s.t.

- Normalized: $f(\emptyset) = 0$
- Nonnegative: $f(S) \ge 0$ $\forall S$
- Monotone: $f(S) \le f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$,

$f(A) + f(B) \ge f(A \cup B) + f(A \cap B)$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,

$$\underbrace{f(S+e) - f(S)}_{\overset{\text{def}}{=} f(e|S)} \ge \underbrace{f(T+e) - f(T)}_{\overset{\text{def}}{=} f(e|T)}$$

randomized-greedy(f, k)

1. $S_0 \leftarrow \emptyset$
2. For $i = 1, \ldots, k$:
   A. Let $R_i \subseteq \mathcal{N}$ be the k elements e w/ max $f(e \mid S_{i-1})$.
   B. Sample $e_i \sim R$ uniformly at random.
   C. $S_i \leftarrow S_{i-1} + e_i$.
3. Return S_k.

Lemma $R \subseteq \mathcal{N}$ random, $P[e \in R] \leq p$ $\forall e$
then $E[f(R)] \geq (1-p) f(\emptyset)$

`randomized-greedy`(f, k)

1. $S_0 \leftarrow \emptyset$
2. For $i = 1, \ldots, k$:
   A. Let $R_i \subseteq \mathcal{N}$ be the k elements e w/ max $f(e \mid S_{i-1})$.
   B. Sample $e_i \sim R$ uniformly at random.
   C. $S_i \leftarrow S_{i-1} + e_i$.
3. Return S_k.

Set function $f: 2^{\mathcal{N}} \to \mathbb{R}_{\geq 0}$ s.t.

- Normalized: $f(\emptyset) = 0$
- Nonnegative: $f(S) \geq 0$ $\forall S$
- Monotone: $f(S) \leq f(T)$ for $S \subseteq T$
- Submodular: for $A, B \subseteq \mathcal{N}$,
  $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$

"decreasing marginal values"

for $S \subseteq T \subseteq \mathcal{N}$, $e \in \mathcal{N}$,

$$\underbrace{f(S+e) - f(S)}_{\overset{\text{def}}{=} f(e \mid S)} \geq \underbrace{f(T+e) - f(T)}_{\overset{\text{def}}{=} f(e \mid T)}$$