CS7800: Advanced Algorithms

Class 17: Approximation Algorithms II

- Max Coverage
- Vertex and Set Cover

Jonathan Ullman

November 4, 2025

Maximum Coverage

Inputs: Sets $S_1, \dots, S_m \subseteq \underbrace{\{1, \dots, n\}}_{\text{domain}}$

A budget $k \geq 0$

Outputs/Objective: Choose sets $\{A_1, \dots, A_k\} \subseteq \{S_1, \dots, S_m\}$

maximizing $\underbrace{\left|\bigcup_{i=1}^{k} A_i\right|}$ covers as many elements as possible

- Can solve in time $O\left(\binom{m}{k}\right) = O(m^k)$

- Problem is NP-hard to solve exactly

Why?

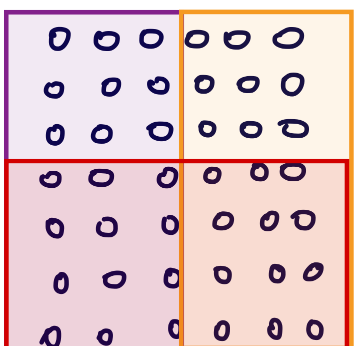
Greedy Max Coverage

For $i = 1, \ldots, k$:

- Let A_i be the set maximizing $|A_1 \cup A_2 \cup \ldots \cup A_i|$

Equivalent to maximizing $|A_i \setminus \underline{(A_1 \cup \ldots \cup A_{i-1})}|$

Bad Example $(k=2)$:



is $1 \times \frac{1}{2}$

is $1 \times \frac{1}{2}$

is $(\frac{1}{2} + \varepsilon) \times 1$

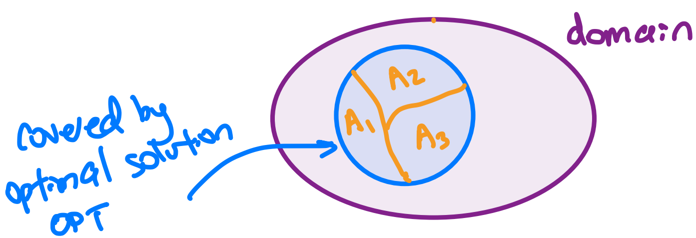
For $k=2$, greedy is at best a $\frac{3}{4}$-approx

Greedy Max Coverage Analysis

<u>Key Claim</u>: Let c_i be the number of elts covered by the first i sets $A_1 \cup \ldots A_i$. Then at iteration i there exists a set that covers at least $\frac{OPT - c_i}{k}$ new elements

$\Longrightarrow$ for every i, $c_i - c_{i-1} \geq \frac{OPT - c_{i-1}}{k}$

<u>Special Case</u>: In the first iteration, greedy chooses A_1 such that $|A_1| \geq \frac{OPT}{k}$



since $A_1, \ldots, A_k$ cover OPT elements, one must cover $\frac{OPT}{k}$ elements

Greedy Max Coverage Analysis

Key Claim: Let c_i be the number of elts covered by the first i sets $A_1 \cup \ldots A_i$. Then at iteration i there exists a set that covers at least $\frac{OPT - c_i}{k}$ new elements

$\Longrightarrow$ for every i, $c_i - c_{i-1} \geq \frac{OPT - c_{i-1}}{k}$

When choosing A_i, you will cover at least $\frac{1}{k}(OPT - \#\text{elts already covered})$

greedy picks sets $A_1 \ldots A_k$

$c_i = |A_1 \cup A_2 \cup \ldots \cup A_i|$

$c_i - c_{i-1} = |A_i \setminus (A_1 \cup \ldots \cup A_{i-1})|$

greedy chooses A_i to maximize this quantity

Greedy Max Coverage Analysis

<u>Key Claim</u>: Let c_i be the number of elts covered by the first i sets $A_1 \cup \dots A_i$. Then at iteration i there exists a set that covers at least $\frac{OPT - c_i}{k}$ new elements

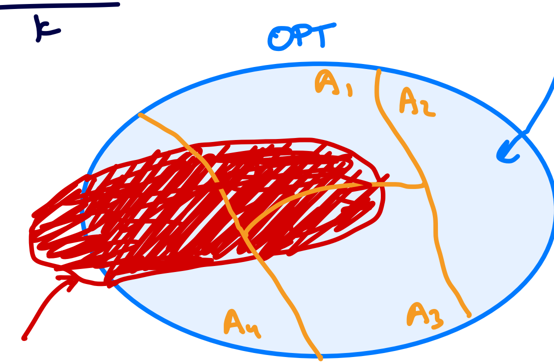
$\Longrightarrow$ for every i, $c_i - c_{i-1} \geq \frac{OPT - c_{i-1}}{k}$

greedy picks sets $A_1 \dots A_k$

$c_i = |A_1 \cup A_2 \cup \dots \cup A_i|$

$c_i - c_{i-1} = |A_i \setminus (A_1 \cup \dots \cup A_{i-1})|$

greedy chooses A_i to maximize this quantity



Greedy Max Coverage Analysis

$$\lim_{k\to\infty}\left(1+\frac{c}{k}\right)^k = e^c$$

Key Clm: In iteration i, $c_i - c_{i-1} \geq \frac{OPT - c_{i-1}}{k}$

value of greedy

$$c_k = c_k - c_{k-1} + c_{k-1}$$

$$\geq \frac{OPT - c_{k-1}}{k} + c_{k-1} = \frac{OPT}{k} + \left(1-\frac{1}{k}\right)\cdot c_{k-1}$$

$$\geq \frac{OPT}{k} + \left(1-\frac{1}{k}\right)\left(\frac{OPT}{k} + \left(1-\frac{1}{k}\right)\cdot c_{k-2}\right)$$

$$= \frac{OPT}{k}\left(1+\left(1-\frac{1}{k}\right)\right) + \left(1-\frac{1}{k}\right)^2\cdot c_{k-2}$$

$$\cdots$$

$$= \frac{OPT}{k}\left(1+\left(1-\frac{1}{k}\right)+\left(1-\frac{1}{k}\right)^2+\ldots+\left(1-\frac{1}{k}\right)^{k-1}\right)$$

$$= \frac{OPT}{k}\cdot\frac{1-\left(1-\frac{1}{k}\right)^k}{1-\left(1-\frac{1}{k}\right)} = \frac{OPT}{\cancel{k}}\cdot\frac{1-\left(1-\frac{1}{k}\right)^k}{\cancel{1/k}} = \left(1-\left(1-\frac{1}{k}\right)^k\right)\cdot OPT$$

$$\geq \left(1-1/e\right)\cdot OPT$$

Greedy Max Coverage

For $i = 1, \ldots, k$:

- Let A_i be the set maximizing $|A_1 \cup A_2 \cup \ldots \cup A_i|$

Equivalent to maximizing $|A_i \setminus (A_1 \cup \ldots \cup A_{i-1})|$

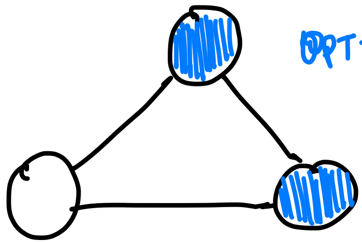
<u>Thm:</u> Greedy MC gives a $(1-1/e)$-approximation in time $O(n^3)$

I didn't think hard about this

Vertex Cover

<u>Input:</u> Given an undirected, unweighted graph $G=(V,E)$

<u>Output/Objective:</u> Find a subset of nodes $S \subseteq V$ such that $\forall (u,v) \in E$ $u \in S$ or $v \in S$ and $|S|$ is as small as possible



OPT = 2

NP-hard to solve exactly

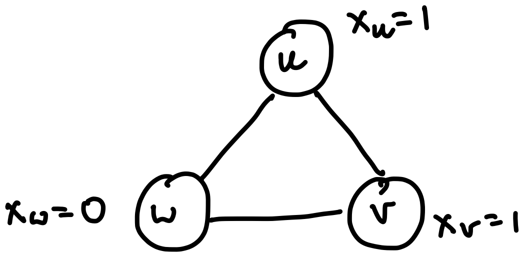
LP / ILP for Vertex Cover

ILP

$$\min \sum_{u \in V} x_u$$

$$x_u + x_v \geq 1 \quad \text{for } (u,v) \in E$$

$$x_u \in \{0,1\} \quad \text{for } u \in V$$



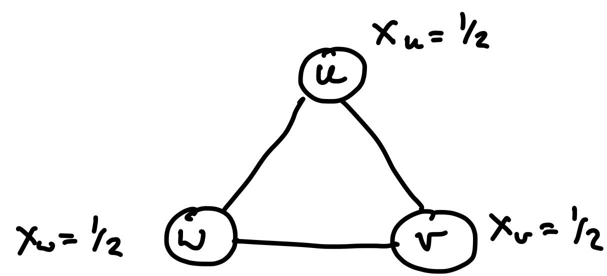
$OPT_{IP} = 2$

LP

$$\min \sum_{u \in V} x_u$$

$$x_u + x_v \geq 1 \quad \text{for } (u,v) \in E$$

$$0 \leq x_u \leq 1 \quad \text{for } u \in V$$



$OPT_{LP} = 3/2$

LP / ILP for Vertex Cover

ILP

$$\min \sum_{u \in V} x_u$$

$$x_u + x_v \geq 1 \quad \text{for } (u,v) \in E$$

$$x_u \in \{0,1\} \quad \text{for } u \in V$$

LP

$$\min \sum_{u \in V} x_u$$

$$x_u + x_v \geq 1 \quad \text{for } (u,v) \in E$$

$$0 \leq x_u \leq 1 \quad \text{for } u \in V$$

LP Rounding:
- Let x^*_{LP} be the optimal LP solution
- Turn it into x_{IP} that is feasible for the ILP and also $OBJ(x_{IP}) \leq c \cdot OBJ(x^*_{LP})$
- $OPT_{IP} \geq OPT_{LP} = OBJ(x^*_{LP}) \geq \frac{1}{c} \cdot OBJ(x_{IP})$

LP / ILP for Vertex Cover

ILP

$$\min \sum_{u \in V} x_u$$
$$x_u + x_v \geq 1 \quad \text{for } (u,v) \in E$$
$$x_u \in \{0,1\} \quad \text{for } u \in V$$

LP

$$\min \sum_{u \in V} x_u$$
$$x_u + x_v \geq 1 \quad \text{for } (u,v) \in E$$
$$0 \leq x_u \leq 1 \quad \text{for } u \in V$$

- Let x^* be the optimal LP solution
- Let x be as follows: $x_u = \begin{cases} 1 & \text{if } x^*_u \geq 1/2 \\ 0 & \text{if } x^*_u < 1/2 \end{cases}$

① x is feasible for ILP

② $\sum_{u \in V} x_u \leq 2 \cdot \sum_{u \in V} x^*_u$

$$OPT_{IP} \geq OPT_{LP} = OBJ(x^*) \geq \tfrac{1}{2} \cdot OBJ(x)$$
$$\longrightarrow OBJ(x) \leq 2 \cdot OPT_{IP}$$

LP Relaxation for Set Cover

Domain $\{1, \dots, m\}$

Sets $S_1, \dots, S_n \subseteq \{1, \dots, m\}$

Generalization of Vertex Cover

ILP

$$\min_x \sum_{i=1}^{n} x_i$$

$$\sum_{i : j \in S_i} x_i \geq 1 \quad \text{for } j \in \{1, \dots, m\}$$

$$x_i \in \{0, 1\}$$

Approx Alg

1. Solve LP relaxation to get x^*
2. For $i = 1, \dots, n$:

   set $x_i = 1$ with probability x_i^*
3. Repeat r times and take the union of the covers

If $x_1 + x_2 + \dots + x_t \geq 1$

Suppose $x_1 = x_2 = \dots = x_t = 1/t$

Claim: For every domain elt j

$$\Pr(\exists\, i,\ j \in S_i \text{ and } x_i = 1) \geq 1 - 1/e$$