COL 351:
Analysis and Design of Algorithms

Lecture 37

Some NP Complete Problems

NP Complete

A problem X in NP class is said to be NP-Complete if for each $A \in$ NP, we have $A \leq_P X$

SAT → **3-SAT** → **Vertex-Cover**

Vertex-Cover → **Independent-Set** *(Easy)*

Vertex-Cover → **Dominating-Set** *(Lecture 34)*

Vertex-Cover → **Hitting-Set** *(assignment)*

Vertex-Cover → **Tracking-Set** *(assignment)*

Independent-Set → **Strongly Independent** *(tutorial)*

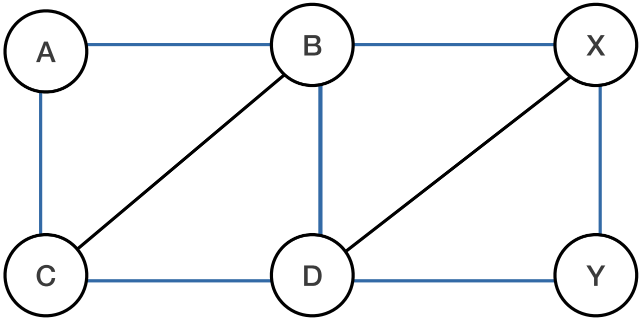
Independent-Set → **GIS** *(tutorial)*

Vertex Cover

Given: A graph $G = (V, E)$ with n vertices.

Def: A subset $W \subseteq V$ such that for each $(a, b) \in E$, one end-point of (a, b) lies in W.

Example:



Optimization Version:
Find a vertex-cover of minimum size.

Decision Version:
Find if there is a vertex-cover of size k.

2-Approximate Vertex Cover

$S = \phi$;

While (E has an uncovered edge):

    $(x, y) \leftarrow$ An arbitrary uncovered edge;

    $S = S \cup \{x, y\}$;

    Mark edges incident to x, y as covered;

Return S;

Claim 1: The set S is a vertex-cover for the input graph $G = (V, E)$.

Claim 2: We have $\dfrac{|S|}{|S_{opt}|}$ is at most 2.

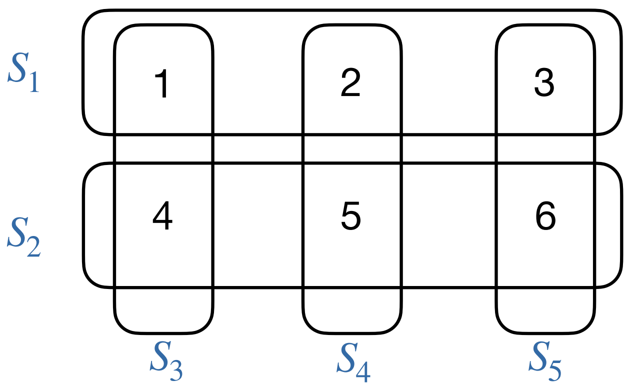
(See Lecture 3)

Set Cover Problem

Given: A universe $U = \{1,\ldots,n\}$ with n elements.
A family $F = \{S_1, \ldots, S_m\}$ of m subsets of U. That is, $S_1, \ldots, S_m \subseteq U$.

Definition: Subsets $S_{j_1}, \ldots, S_{j_k}$ lying in F whose union is U.

Example:



$|U| = 6 \quad m = 5$

Optimization Version:
Find a set-cover of minimum size.

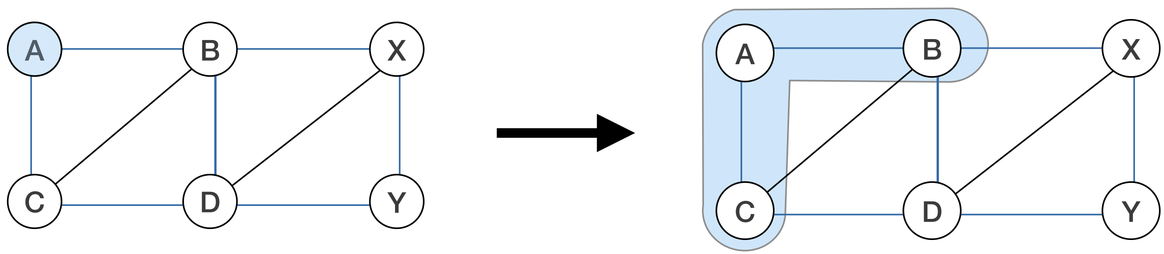
Decision Version:
Find if there is a set-cover of size k.

Dominating-Set $\leq_P$ Set-Cover

Instance of Dominating Set: A graph $G = (V = \{v_1, \ldots, v_n\}, E = \{e_1, \ldots, e_m\})$ and a parameter k.

Generating an Instance of Set-Cover

1. Define $U = \{v_1, \ldots, v_n\}$, the parameter k remains same.
2. For each i , define the set $S_i = N(v_i) \cup \{v_i\}$. Thus, $|S_i| = 1 + deg(v_i)$.



Natural approaches: Approximate Set Cover

Approach 1:

Greedily pick set of largest size.

$|U| = n$

$|\text{opt-sol}^n| = 2$

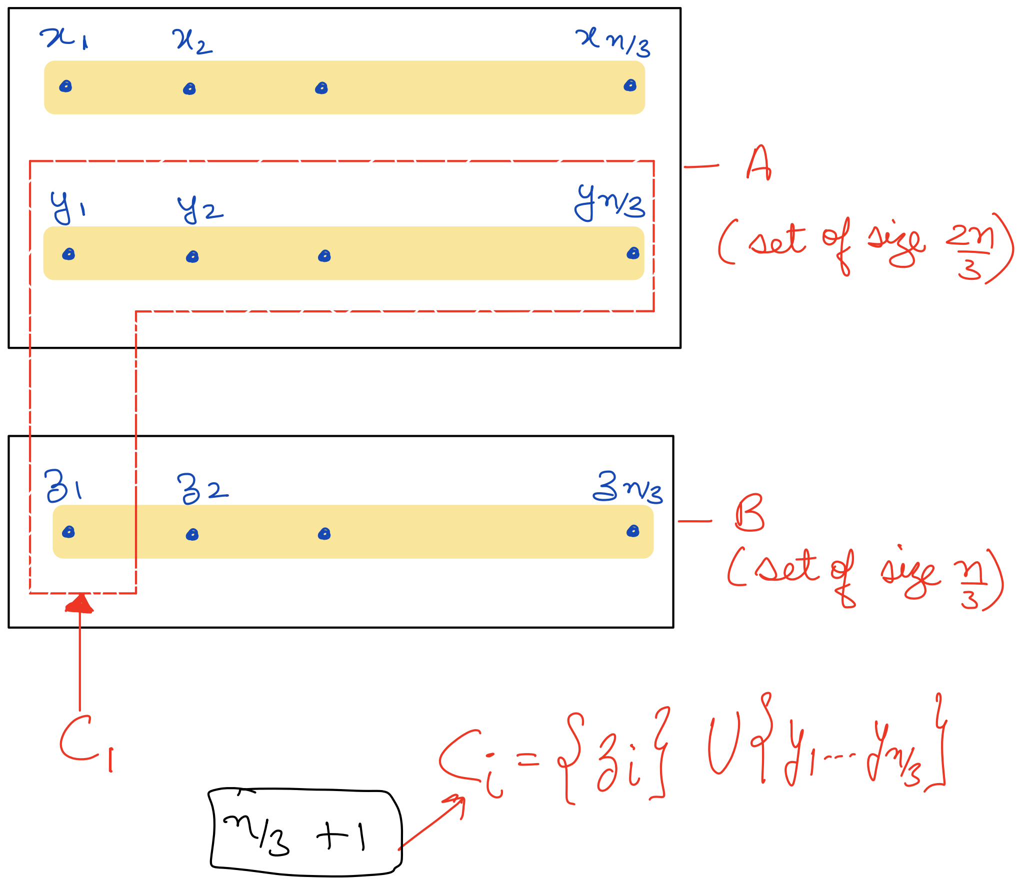
$|\text{greedy sol}^n| = \frac{n}{3} + 1$

So,

approx factor is

$\Omega(n)$ in

worst case



Natural approaches: Approximate Set Cover

Approach 1:

Greedily pick set of largest size. — $\Omega(n)$ approximation bound in worst case

Approach 2:

Greedily pick set containing **largest number of uncovered elements.**

— $O(\log n)$ approximation bound

Approximate Set Cover

$|U| = n$

$\mathcal{F} = \{S_1 \ldots S_m\}$

$\mathcal{A} \subseteq \mathcal{F}$

Approximate-Set-Cover(U, F)

$A \leftarrow \{\}$ /*empty family*/

$X \leftarrow U.$ /*uncovered elements*/

While ($|X| > 0$):

1. Select an $S \in F$ that maximizes $|X \cap S|$.
2. A = Add S to A.
3. $X = X - S$.

Return A.

No of uncovered elements in S.

Goal : $O(\log n)$ - approx. sol^n

Approximate Set Cover

Approximate-Set-Cover(U, F)

$A \leftarrow \{\}$ /*empty family*/

$X \leftarrow U.$ /*uncovered elements*/

While ($|X| > 0$):

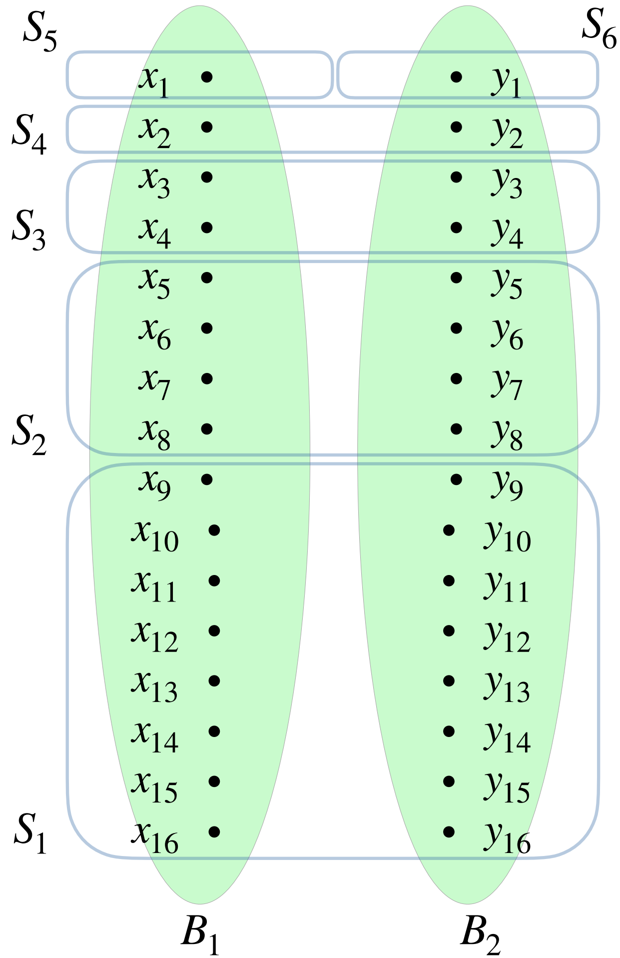
1. Select an $S \in F$ that maximizes $|X \cap S|$.
2. A = Add S to A.
3. $X = X - S$.

Return A.

$A = \{ S_1, S_2, S_3, S_4, S_5, S_6 \}$

In general,

approx-factor is $\Omega(\log n)$ in worst case.



Finding Approximate Set Cover

Approximate-Set-Cover(U, F)

$A \leftarrow \{\}$ /*empty family*/

$X \leftarrow U$. /*uncovered elements*/

While ($|X| > 0$):

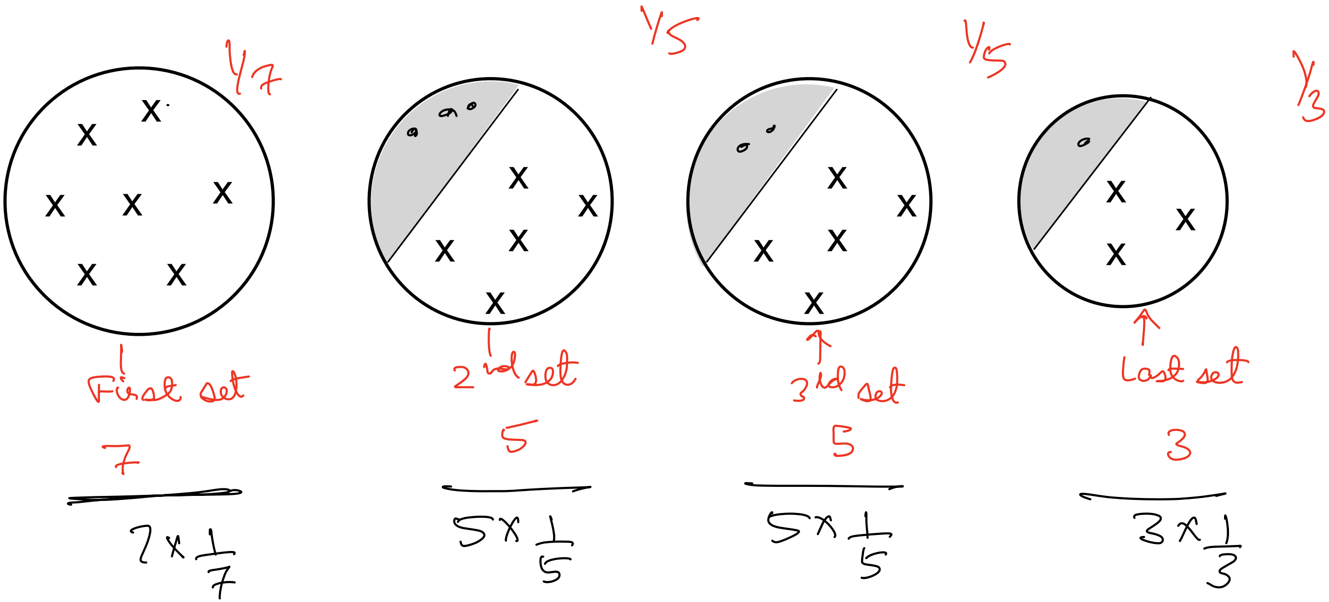
1. Select an $S \in F$ that maximizes $|X \cap S|$.
2. A = Add S to A.
4. 3. $X = X - S$.
3. 4. For $w \in \underbrace{X \cap S}$: set $c(w) = \frac{1}{|X \cap S|}$ ← cost of element

Return A.

uncovered elements that are getting covered by set S.

Example

$|U| = 20$



Total sum = 4

= |A|

$|A| = 4$

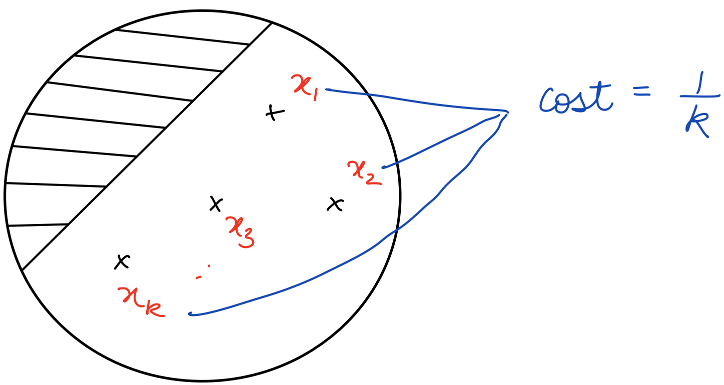
Observations

Claim 1: $$\sum_{\substack{w \in S \,:\, w \text{ was} \\ \text{covered by } s}} c(w) = 1.$$

Proof

Suppose S covered k elements, $x_1 \; x_2 \; \ldots \; x_k$

Then $c(x_i) = \frac{1}{k}$, $i = 1$ to k



Observations

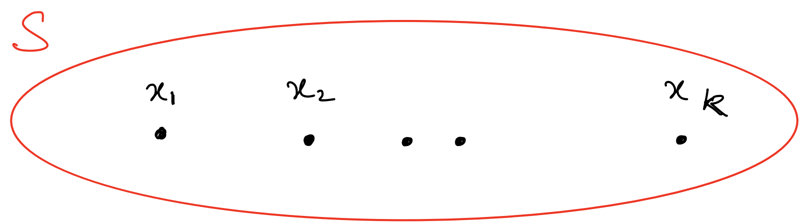
Claim 2: $\sum_{w \in S} c(w) \leq \log n.$

Proof

Suppose $S = \{x_1 \ldots x_k\}$,

and the order in which

elements are covered is also $x_1, \ldots x_k$.



$c(x_1) \leq \frac{1}{k}$ bcoz when x_1 is covered there is a set (S) with k uncovered elements.

$c(x_2) \leq \frac{1}{k-1}$ bcoz when x_2 is covered there is a set (S) with $(k-1)$ uncovered elements.

$c(x_k) \leq 1$

$\therefore$ sum total cost $\leq 1 + \frac{1}{2} + \ldots + \frac{1}{k} \leq \log n$

Observations

Claim 3: $|A| \leq |A_{opt}| \times \log n.$

Proof

By Claim 1 and Claim 2

By Claim 1, $|A| = \sum_{w \in U} cost(w)$

Now RHS $\leq \sum_{S \in A_{opt}} \sum_{w \in S} cost(w) \leq |A_{opt}| * \log n$

(due to claim 2)

Thus, $|A| \leq |A_{opt}| * \log(n)$