Set Cover : $S_1 \ldots S_n$, $U = \bigcup_{i=1}^{n} S_i$

find $\min I \subseteq [1, n]$ s.t. $\bigcup_{i \in I} S_i = U$

I.P.	L.P.
$\min \sum_{i=1}^{n} x_i$	$\min \sum_{i=1}^{n} x_i$
$\sum_{i \mid z \in S_i} x_i \geq 1, \quad \forall z \in U$	$\sum_{i \mid z \in S_i} x_i \geq 1, \quad \forall z \in U$
$x_i \geq 0, \quad x_i \in \mathbb{N} \cup \{0\}$	$x_i \geq 0$

Rand Algo.

1. $I = \phi$
2. $x^* = $ opt of L.P.
3. For $i = 1$ to n : add i to "I" w.p. x_i^*
4. Return I.

* $Exp(|I|) = \sum_{i=1}^{n} x_i^* = LP_{opt} \leq IP_{opt} = $ min-set-cover size

Lemma : For any $z \in U$. Then Prob(z is covered by Rand algo) $\geq 1 - \frac{1}{e}$

Ques : How to ensure that all elements covered with good prob.

Ans : Repeat $\log(|U|)$ times

Rand Algo New:

1. $I = \emptyset$
2. $x^* = $ opt of L.P.
3. Repeat $(\log|U| + \alpha)$ times:
   $\underbrace{\text{For } i=1 \text{ to } n: \quad \text{add } i \text{ to "}I\text{" w.p. } x_i^*}$
4. Return I.

<u>Lemma</u>: Prob($\exists\, z \in U$ which isn't covered by new algo) $\leq e^{-\alpha}$

Proof: Take $z \in U$. Prob(z isn't covered) $= (1/e)^{\log|U| + \alpha}$

$$\text{Prob}(\exists\, z \in U \text{ which isn't covered}) \leq |U| \cdot e^{-(\log|U| + \alpha)} = e^{-\alpha}.$$

(*) $\text{Exp}(|I|) \leq (\log|U| + \alpha) * \underbrace{\sum_{i=1}^{n} x_i}$

$\leq (\log|U| + \alpha) \cdot$ min set cover size

By Markov's Inequality, $\text{Prob}(R \geq c\,E[R]) \leq \frac{1}{c}$ if $R \geq 0$

$$\text{Prob}\left(|I| \geq 10 \underbrace{(\log|U| + \alpha)^*}_{\text{min set cover size}}\right) \leq \text{Prob}(|I| \geq 10\,E[|I|]) \leq 0.1$$

Our stretch $= O(\log |U|)$ — Best stretch know for set cover.

IP $\longrightarrow$ LP $\longrightarrow$ Dual of LP & Greedy algo

rand algo

Primal

$$\min c^T x \quad \text{s.t. } Ax \geq b, \quad x \geq 0$$

$\longrightarrow$

Dual

$$\max y^T b \quad \text{s.t. } y^T A \leq c, \quad y \geq 0$$

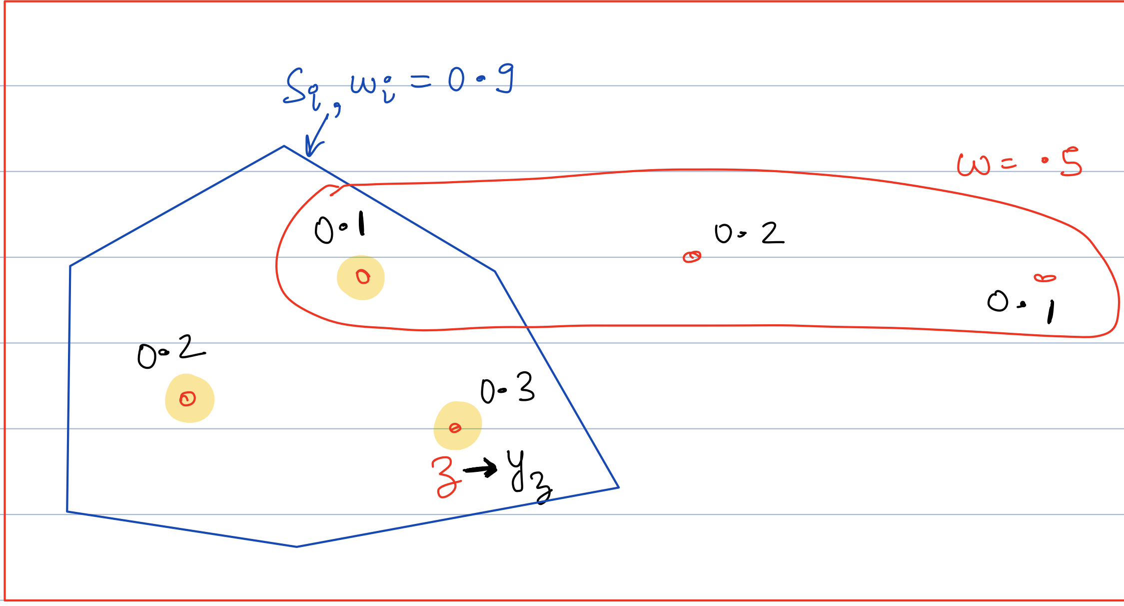
$m = |U|$

$n =$ no of sets

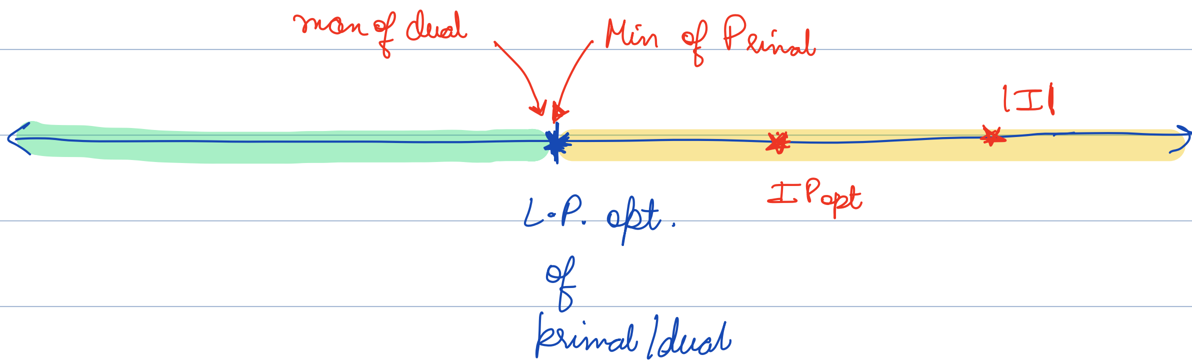
$$\begin{array}{cc} & \begin{array}{ccc} x_1 & x_i & x_n \end{array} \\ \text{For } z \in U \quad y_z & \begin{bmatrix} & & \\ & 1 & \\ & & \end{bmatrix}_{m \times n} \end{array}$$

$(z, i) = 1$ iff $z \in S_i$

L.P. $\min \sum_{i=1}^{n} w_i x_i$	**Dual** $\max \sum_{z \in U} y_z$
$\boxed{\sum_{i \mid z \in S_i} x_i \geq 1}, \quad \forall z \in U$	$\boxed{\sum_{z \mid z \in S_i} y_z \leq w_i}, \quad \forall i \leq n$
$\lvert U \rvert$ inequalities $\quad x_i \geq 0$	$y_z \geq 0 \quad \forall z \in U$



$$0.1 + 0.2 + 0.3 \leq 0.9$$



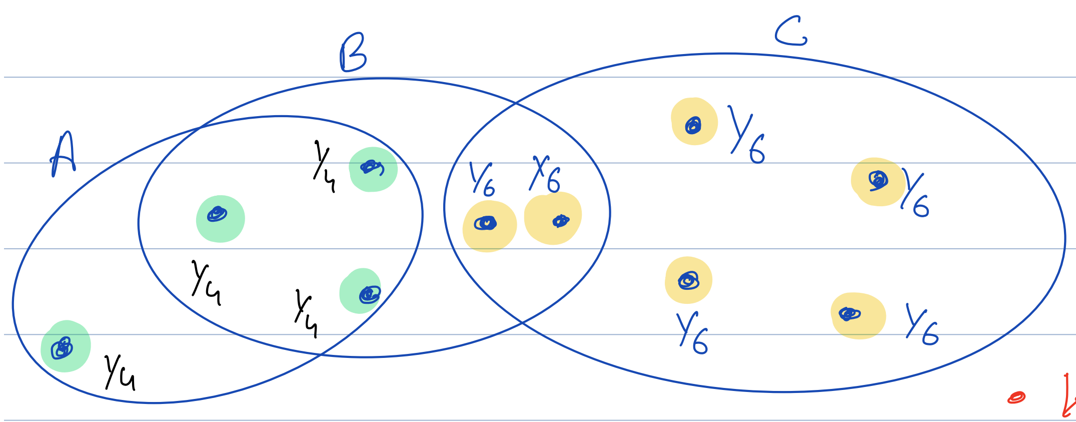
Greedy Algo. (for unweighted case)

1. $I \leftarrow \phi$
2. While $\exists$ an uncovered element.
   - $i \leftarrow$ index of set S_i with largest no. of uncovered elements. /// (cost effec. of set picked)
   - $I = I \cup \{i\}$
3. Return I

$z_1 z_2 \cdots z_m \leftarrow$ ordering in which element of U are covered.

For each S_i in our sol^n, $i \in I$ //

COST-EFFECTIVENESS $(S_i) \equiv$ No of uncovered elements that were covered by set S_i



$C \leftarrow 6 \quad A \leftarrow 4$

- What is sol^n by greedy algo?
- What is cost-effectiveness of sets in our sol^n?

- For $j = 1$ to m $\quad z_j \leftarrow j^{th}$ element in U covered by our sol^n

$c_j \leftarrow$ cost effectiveness of set that first covered z_j.

H.W.1 $\quad |I| = \sum_{j=1}^{m} 1/c_j$

H.W.2

$$\forall j \in [1, m], \left(\underbrace{0 \cdots 0}_{j-1} \quad \underbrace{1/c_j \cdots 1/c_j}_{m-(j-1)} \right) \leftarrow \text{feasible for Dual}$$