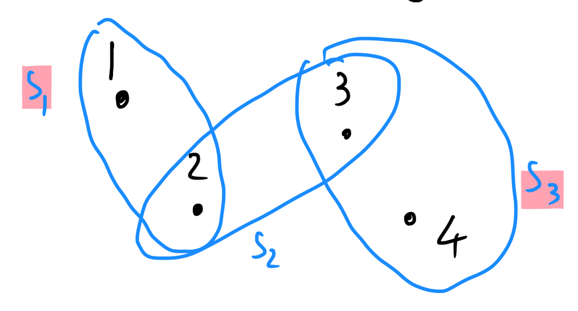
SET COVER

INPUT: $U=[n]=\{1,2,\dots,n\}$, $S_1, S_2, \dots, S_m \subseteq U$

Q: WHAT IS THE SIZE OF A SMALLEST $T \subseteq [m]$, S.T., $\bigcup_{j\in T} S_j = [n]$?



$n=4$
$m=3$

$$IP_{sc}\begin{cases} \min \sum_{j=1}^{m} x_j \\ \sum_{\substack{j\in[m]\\ i\in S_j}} x_j \geq 1 & \forall i \in [n] \\ x_j \in \{0,1\} & \forall j \in [m] \end{cases}$$

($x_j = 1$ IF S_j IS IN THE SOLUTION)

$\Downarrow$

$$LP_{sc}\begin{cases} \min \sum_{j=1}^{m} x_j \\ \sum_{\substack{j\in[m]\\ i\in S_j}} x_j \geq 1 & \forall i \in [n] \\ 0 \leq x_j \leq 1 & \forall j \in [m] \end{cases}$$

LP^*_{sc} IS THE VALUE OF AN OPTIMAL SOLUTION OF LP_{sc}.

OBS.: $LP^* \leq IP^*$

P: THE LP IS A "RELAXATION" TO THE IP (THE CONSTRAINTS OF THE LP ARE LESS STRINGENT THAN THOSE OF THE IP). THUS EACH IP SOLUTION IS ALSO A LP SOLUTION. □

("RANDOMIZED ROUNDING" TECHNIQUE)

$A \leftarrow \emptyset$
LET x^* BE AN OPTIMAL LP SOLUTION
FOR $K=1\dots\lceil 2 \ln n\rceil$ // PHASES
  FOR $j=1\dots m$ // SETS
    FLIP AN IND. COIN WITH HEAD PROB. x^*_j
    IF THE COIN COMES UP HEADS:
      $A \leftarrow A \cup \{S_j\}$
RETURN A

L1: CONSIDER THE GENERIC ITERATION K OF THE OUTER LOOP. LET p_i BE THE PROBABILITY THAT, IN THIS ITERATION K, AT LEAST ONE OF THE SETS CONTAINING i GETS ADDED TO A. THEN, $p_i \geq 1 - \frac{1}{e} \approx 0.63\dots$

$1-x \leq e^{-x}$

P: $\Pr\{i$ IS NOT COVERED IN ITERATION $K\}$

$$= \prod_{\substack{j\in[m]\\ i\in S_j}} (1-x^*_j) \leq \prod_{\substack{j\in[m]\\ i\in S_j}} e^{-x^*_j} = e^{-\sum_{j\in[m], i\in S_j} x^*_j} \leq e^{-1}$$

THE x^*_j'S FORM A FEASIBLE SOL. TO THE LP.

$\Pr\{i$ IS COVERED IN ITERATION $K\} \geq 1-\frac{1}{e}$. □

L2: LET $i\in[n]$. $\Pr\{i$ IS NOT COVERED BY ANY SET IN $A\} \leq \frac{1}{n^2}$.

P: $\Pr\{i$ IS NOT COVERED BY ANY SET IN $A\}$ =

$$= \prod_{K=1}^{\lceil 2\ln n\rceil} \Pr\{i \text{ IS NOT COVERED IN ITERATION } K\}$$

$$\leq \prod_{K=1}^{\lceil 2\ln n\rceil} \frac{1}{e} = e^{-\lceil 2 \ln n\rceil} \leq e^{-2\ln n} = n^{-2}.$$ □

L3: $\Pr\{A$ IS A SET COVER$\} \geq 1 - \frac{1}{n}$.

P: $\Pr\{A$ IS NOT A SET COVER$\} = \Pr\{\exists i : A$ DOES NOT COVER $i\}$

UNION BOUND $\rightarrow \leq \sum_{i=1}^{n} \Pr\{A$ DOES NOT COVER $i\}$

L2 $\rightarrow \leq \sum_{i=1}^{n} \frac{1}{n^2} = \frac{n}{n^2} = \frac{1}{n}$. □

SE A È UN SC, ALLORA $|A| \geq IP^* \geq LP^*$.

L4: $E[|A|] \leq \lceil 2\ln n\rceil\, LP^* \leq \lceil 2 \ln n\rceil\, OPT_{sc}$. $(= \lceil 2\ln n\rceil\, IP^*)$

P: LET US FIX ONE ITERATION K OF THE OUTER LOOP. LET A_K BE THE CLASS OF SETS ADDED TO A IN ITERATION K.

THEN,

LIN. OF EXPECT.

$$E[|A_K|] = \sum_{j=1}^{m} \Pr\{S_j \in A_K\} = \sum_{j=1}^{m} x^*_j = LP^* \leq OPT_{sc}$$

$A = A_1 \cup A_2 \cup \dots \cup A_K \cup \dots \cup A_{\lceil 2\ln n\rceil}$

$|A| \leq \sum_{i=1}^{\lceil 2\ln n\rceil} |A_i|$

$$E[|A|] \leq \sum_{i=1}^{\lceil 2 \ln n\rceil} E[|A_i|] = \sum_{i=1}^{\lceil 2\ln n\rceil} LP^* = \lceil 2\ln n\rceil\, LP^* \leq \lceil 2\ln n\rceil\, OPT_{sc}.$$ □

$\exists$ INSTANCE $\quad$ $\forall$ INSTANCES

$$\frac{1}{4\ln 2}\ln n \leq \frac{OPT_{sc}}{LP^*_{sc}} \leq \lceil 2\ln n\rceil$$

$\binom{S}{K} = \{T \mid T\subseteq S \wedge |T| = K\}$

$[q] = \{1,2,\dots,q\}$

IG LOWER BOUND

UNIVERSE SET: $E = \{e_A \mid A \in \binom{[q]}{q/2}\}$, FOR SOME EVEN $q \geq 2$.

$n = |E| = \binom{q}{q/2} = \Theta\left(\frac{2^q}{\sqrt{q}}\right) \Rightarrow q \geq \lg n - \Theta(\lg\lg n)$

$\forall i \in [q]$: $S_i = \{e_A \mid e_A \in E \wedge i \in A\}$

$\mathcal{C} = \{S_i \mid i \in [q]\}$

$|\mathcal{C}| = q \stackrel{\Delta}{=} m$

$q=4$
$E=\{e_{12}, e_{13}, e_{14}, e_{23}, e_{24}, e_{34}\}$
$S_1 = \{e_{12}, e_{13}, e_{14}\}$ $S_2 = \{e_{12}, e_{23}, e_{24}\}$ $S_3 = \{e_{13}, e_{23}, e_{34}\}$ $S_4 = \{e_{14}, e_{24}, e_{34}\}$

$$LP_{sc}\begin{cases} \min \sum_{j=1}^{m} x_j \\ \sum_{\substack{j\in[m]\\ i\in S_j}} x_j \geq 1 & \forall i \in [n] \\ 0 \leq x_j \leq 1 & \forall j \in [m] \end{cases}$$

L: $LP^*_{sc} \leq 2$

P: CONSIDER THE LP SOL: $x_1 = x_2 = x_3 = \dots = x_q = \frac{2}{q}$ $(q=m)$

THIS SOLUTION HAS A VALUE OF 2: $\sum_{i=1}^{q} x_i = q \cdot \frac{2}{q} = 2$

THE GENERIC LP CONSTRAINT IS $\sum_{\substack{S_j\\ e_A\in S_j}} x_j \geq 1$ $\quad(\forall e_A \in E)$

$$\sum_{\substack{S_j\\ e_A\in S_j}} x_j = \sum_{\substack{S_j\\ e_A\in S_j}} \frac{2}{q} = |A|\cdot\frac{2}{q} = \frac{q}{2}\cdot\frac{2}{q} = 1 \geq 1.$$ ✓

THUS, THE SOL. IS FEASIBLE. □

L: $OPT_{sc} \geq \frac{1}{2}\lg_2 n - O(\lg\lg n)$

P: SUPPOSE THAT THE SETS $S_{i_1}, S_{i_2}, \dots, S_{i_K}$ FOR $K \leq \frac{q}{2}$ COVER EACH ELEMENT OF THE INSTANCE.

CONSIDER $T = [q] - \{i_1,\dots,i_K\}$.

THEN, $\exists A \subseteq T$ S.T. $|A| = \frac{q}{2}$ $\left(|T| \geq q - K \geq \frac{q}{2}\right)$.

THEN, $e_A \in E$.

ELEMENT e_A IS NOT COVERED BY $S_{i_1}, \dots, S_{i_K}$ BECAUSE $A \cap \{i_1,\dots,i_K\} = \emptyset$.

THUS, $S_{i_1},\dots,S_{i_K}$ IS NOT A SET COVER IF $K \leq \frac{q}{2}$.

IT FOLLOWS THAT THE MINIMUM SET COVER CONTAINS $\frac{q}{2}+1$ SETS. □

$$IG = \frac{OPT_{sc}}{LP^*_{sc}} \geq \frac{\frac{1}{2}\lg_2 n - O(\lg\lg n)}{2} \approx \frac{1}{4}\lg_2 n$$

ON SOME INSTANCES.