Set-covering problem (清1·1)

13. (a) (4%) Let $X = \{1,2,\ldots,n\}$. For a subset of X, we say that it covers its elements. Given a set $S = \{S_1, S_2, \ldots, S_m\}$ of m subsets of X such that $\bigcup_{i=1}^{m} S_i = X$, the set cover problem is to find the smallest subset T of S whose union is equal to X, that is, $\bigcup_{S_i \in T} S_i = X$. For example, suppose that $X = \{1,2,\ldots,5\}$ and there are the following four subsets in S: $S_1 = \{1,3,4\}$, $S_2 = \{2,5\}$, $S_3 = \{1,5\}$ and $S_4 = \{2,4\}$. Then the optimal set cover is $\{S_1, S_2\}$. Please give a counterexample for the following greedy algorithm (2%): At each stage, the algorithm picks the set that covers the greatest number of remaining elements that are uncovered. Please also write down the greedy and optimal solutions of your counterexample (2%).

(b) (6%) Suppose that each subset S_i in S contains only two elements. Can the set cover problem then be solved in polynomial time? If yes, please also design a polynomial-time algorithm to solve this set cover problem (4%) and analyze its time complexity (2%). If no, please also describe your reason (6%).

(a). greedy algorithm:

令每一輪開始時剩餘未cover之set為B

則每一輪挑 $\max\{|B-S_i|\}$ 之 S_i

counter example: $X=\{1,2,3,4\}$

$S=\{\{1,2\},\{2,3\},\{1,4\},\{3\}\}$

照 greedy choice 可能 greedy cover set 為 $\{\{1,2\},\{3\},\{1,4\}\}$

但 optimal solution 為: $\{\{1,4\},\{2,3\}\}$

GREEDY-SET-COVER$(X, \mathcal{F})$

```
U = X
𝒞 = ∅
while U ≠ ∅
    select an S ∈ 𝓕 that maximizes |S ∩ U|
    U = U − S
    𝒞 = 𝒞 ∪ {S}
return 𝒞
```

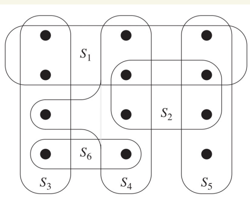


Figure 35.3 An instance $(X, \mathcal{F})$ of the set-covering problem, where X consists of the 12 black points and $\mathcal{F} = \{S_1, S_2, S_3, S_4, S_5, S_6\}$. A minimum-size set cover is $\mathcal{C} = \{S_3, S_4, S_5\}$, with size 3. The greedy algorithm produces a cover of size 4 by selecting either the sets S_1, S_4, S_5, and S_3 or the sets S_1, S_4, S_5, and S_6, in order.

Corollary 35.5
GREEDY-SET-COVER is a polynomial-time $(\ln|X| + 1)$-approximation algorithm.

Proof Use inequality (A.14) and Theorem 35.4. ■

(b). 設 $\forall i = 1,\dots,m$, $S_i \in S$, $|S_i| = 2$

建構一 $G=(V,E)$, $|V|=n$, $\forall S_i \in S$, $S_i=(a,b)$, 則 $(a,b)\in E$, 故 G為一 undirected graph

找 T 滿足 minimum set cover 等價於找 G上之 minimum edge cover $G'=(V,E')$

可利用 Gallai's theorem 知, 求 minimum edge cover 等價於求 G上之 maximum matching

$\because$ $|E'|+|E_m|=|V|$ 即可得 $|E'|$

又 maximum matching 為 poly-time sovlable

Algorithm 如下: minimum_set_cover(X, S)

1. 建構一 $G=(V,E)$ 如上描述 ⇒ $O(|V|+|E|)$
2. 確認 G上有無 isolated vertex, 若有 return "無 set cover" ⇒ $O(|V|+|E|)$
3. G上求 maximum matching ⇒ $O(|V|^4)$ ⇒ Blossom Algo
4. return $|V|-|E|$

另一 Algorithm 如下: step 1. 求得G上之 max matching

step 2. 对無法被 max matching cover 之點再加上边直至所有點都被 cover

Theorem: Set-Covering Problem ∈ NPC

①. Set Covering ∈ NP

給定一T為一 certificate

利用 disjoint set 之 union(x) 即可在 poly time 驗證 $\bigcup_{S_i\in T} S_i = X$

故 Set Covering ∈ NP

②. Vertex Cover $\le_p$ Set Cover

給定一 (G,k) 為 vertex cover 之 instance

欲建構 (S,X,k) 為 set cover 之 instance

使得: G上有size為k之 vertex cover ⇔ S上有一 subset T, $\bigcup_{S_i\in T} S_i = X$ 且 $|T|=k$

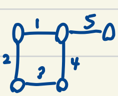
建構方式為: 令 $G=(V,E)$ 而 $|V|=n$, $|E|=m$, 給予每个边一个编号 $1,\dots,m$

則 $X=\{1,2,3,\dots,m\}$, 其中: $i\in X$, 对應至G上的 e_i

而 $\forall v\in V$, 令 v之 incident edge set 為 E_v, 則定義对應之 $S_v=\{j \mid e_j\in E_v\}$

而 $S=\bigcup_{v\in V} S_v$, 而 (S,X,k) 為 Set Cover 上之对應的 instance

Example:

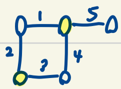


⇔

$X=\{1,2,3,4,5\}$

$S=\{\{1,2\},\{2,3\},\{3,4\},\{1,4,5\},\{5\}\}$

claim: <u>G上有size為k之vertex cover ⇔ S上有一subset T, $\bigcup_{S_i \in T} S_i = X$ 且 $|T| = k$</u>



⇒ 存在 vertex cover V' 且 $|V'|$為k

取其為S之 set cover

Set Covering Problem (台大99)

給定一 universal set U 和 $C = \{C_1, \ldots, C_n\}$, 其中: $C_i \subseteq U$, $\forall i = 1, \ldots, n$

要求 minimum-size 之 C' 其中 $\bigcup_{C_i \in C'} C_i = U$

Algorithm: Set_Cover(U, C)

1. $R = U$
2. $C' = \emptyset$
3. while $R \neq \emptyset$
4. 　　從C中選C_i為$\max\{|R \cap C_i|\}$
5. 　　$C' = C' \cup \{C_i\}$
6. 　　$R = R - \{C_i\}$
7. return C'