

TD 7: Approximation pour le problème de couverture

1 Couverture des sommets pondérés

On s'intéresse au problème de la couverture des sommets de poids minimal. On se donne un graphe $G = (V, E)$ et une pondération $w : V \mapsto \mathbb{N}^*$. Il s'agit de trouver une couverture de sommets C t.q. $\sum_{v \in C} w(v)$ soit minimal. On notera le graphe pondéré $G = (V, E, w)$.

Question 1 : Montrer que le problème de décision sous-jacent est NP-complet (par réduction de 3-SAT).

Soit l'algorithme suivant de calcul de couverture de sommets pour un graphe pondéré $G = (V, E, w)$.

- On trie les sommets par poids croissant, on initialise le sous-ensemble C à vide.
- On itère la procédure suivante: On cherche dans la liste triée le premier sommet u t.q. il existe une arête $\{u, v\}$ avec $\{u, v\} \cap C = \emptyset$. S'il n'en existe pas alors on renvoie C qui est la couverture recherchée. Sinon on ajoute u à C et on continue.

Soit $G = (V, E)$ avec $V = \{v_0, v_1, \dots, v_n\}$ et $E = \{\{v_0, v_i\} \mid 1 \leq i \leq n\}$ (autrement dit une "étoile" de centre v_0).

Question 2 : Proposer une pondération des sommets de ce graphe telle que l'algorithme renvoie une couverture de taille n .

2 Set Cover problem

The set cover problem generalizes vertex cover and many other covering problems. We are given a universe U with $|U| = n$ and a collection of subsets $\mathcal{S} = \{S_1, S_2, \dots, S_m\}$ where $S_i \subseteq U$. The goal is to find a minimum cardinality subcollection $\mathcal{C} \subseteq \mathcal{S}$ such that the selected sets cover all elements, that is, $\bigcup_{S \in \mathcal{C}} S = U$. You may assume that there is at least one such covering.

Question 3 : Show that the vertex cover problem is a special case of the set cover problem.

Consider the following algorithm:

Greedy Set Cover Algorithm:

1. Initialize $\mathcal{C} = \emptyset$ and $R = U$.
2. While $R \neq \emptyset$:
 - (a) Select $S_i \in \mathcal{S}$ that maximizes $|S_i \cap R|$.
 - (b) Add S_i to \mathcal{C} .
 - (c) Update $R \leftarrow R \setminus S_i$.

3. Return \mathcal{C} .

We will use a so-called “charging” argument to assign costs to new elements as they are added. Consider an iteration when the greedy algorithm selects a set S covering t new elements. We charge a cost of $1/t$ to each of these newly covered elements, so the total cost assigned in this iteration is exactly 1. Since the algorithm selects exactly $|\mathcal{C}|$ sets, and each set is “charged” a total cost of 1, the overall cost equals $|\mathcal{C}|$.

Question 4: Let OPT denote the optimal answer. Let k denote the number of uncovered elements remaining before the above iteration. Give a bound on the cost $1/t$ assigned to the newly covered elements in terms of OPT and k .

Question 5: Conclude that the algorithm is an H_n -approximation, where $H_n = \sum_{i=1}^n \frac{1}{i}$ is the n -th harmonic number.

Question 6: Define the *frequency* of an element in U as the number of sets in \mathcal{S} containing it. Show that one can now design an f -approximation algorithm for the set cover problem.

3 Couverture dans un graphe biparti

On suppose dans cette partie que le graphe pondéré $G = (V, E, w)$ est biparti avec $V = L \sqcup R$.

On construit un graphe **orienté** $G' = (V', E')$ dont les arcs sont pondérés par une fonction w' de la façon suivante:

- $V' = V \sqcup \{s, t\}$. Le sommet s est la source et le sommet t est le puits.
- $E' = \{(s, u) \mid u \in L\} \cup \{(v, t) \mid v \in R\} \cup \{(u, v) \mid u \in L \wedge v \in R \wedge \{u, v\} \in E\}$.
- $\forall u \in L \ w'(s, u) = w(u) \wedge \forall v \in R \ w'(v, t) = w(v)$
 $\wedge \forall \{u, v\} \in E \ w'(u, v) = M$ avec $M = 1 + \sum_{v \in R} w(v)$.

Une s - t coupe de G' est une partition (S, T) de V' avec $s \in S$ et $t \in T$. Le poids d'une s - t coupe est défini par $P(S, T) = \sum_{u \in S, v \in T, (u, v) \in E'} w'(u, v)$.

Question 7: Soit une s - t coupe de G' , (S, T) , de poids minimal. Montrer que:

1. $\forall u \in L \ \forall v \in R \ (\{u, v\} \in E \wedge u \in S) \Rightarrow v \in S$.
2. Soient $L' = L \cap T$ et $R' = R \cap S$. Alors $R' \cup L'$ est une couverture de poids minimal de G .

On rappelle qu'un flot dans le graphe orienté G' est une fonction $f : E' \mapsto \mathbb{N}$ telle que $f(u, v) \leq w'(u, v)$ pour tout $(u, v) \in E'$ et $\sum_{\{u|(u,v) \in E'\}} f(u, v) = \sum_{\{u|(v,u) \in E'\}} f(v, u)$ pour tout sommet $v \in V' \setminus \{s, t\}$. La valeur d'un flot, notée $|f|$, est définie par $\sum_{\{v|(s,v) \in E'\}} f(s, v)$.

Question 8: Montrer le lien entre le calcul d'un flot maximal et une s - t coupe de poids minimal. En déduire un algorithme en temps polynomial pour la couverture de sommets de poids minimal dans un graphe biparti.