Sujet L3 - Partition en cliques et couverture de sommets - Corrigé

Un *graphe* G est un couple (V, E) où V est un ensemble de sommets (cercles) et E est un ensemble d'arêtes (trait entre deux cercles).

Une *clique* est un graphe dans lequel chaque paire de sommets est liée par une arête. Une clique C est dite *maximale* dans le graphe G (au sens de l'inclusion) s'il n'existe pas un sommet u de G tel que u est voisin de tous les sommets de C dans G.

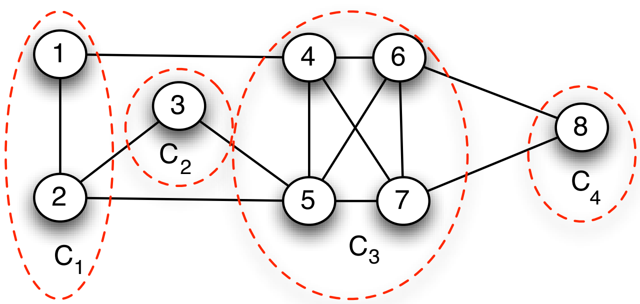


FIGURE 1 – Partition en cliques : les cliques $\{1, 2\}$ et $\{2, 3, 5\}$ sont des cliques maximales. L'ensemble $C = \{C_1, C_2, C_3, C_4\}$ est une partition en cliques maximales.

Une *couverture de sommets* S d'un graphe G est un sous-ensemble de sommets tel que toute arête de G est incidente à au moins un sommet de S. Dans la suite, nous allons nous concentrer sur la couverture de sommets de plus petite cardinalité (en terme du nombre de sommets).

Par la suite, nous noterons S^* une couverture de sommets de plus petite cardinalité. Nous allons noter OPT le nombre de sommets de la solution S^* optimale de la couverture de sommets de G.

Question 1. Dans une clique C de n sommets, donner une couverture de sommets de plus petite cardinalité.

Toute couverture de sommets S de C contient $n - 1$ sommets de cette clique.

Question 2. Considérons l'algorithme $\mathcal{A}$ pour le problème de la couverture de sommets :
Entrée : Un graphe G de n sommets, une numérotation[1] L des sommets de G de 1 à n
Sortie : Un sous-ensemble S des sommets

1. $S \leftarrow \emptyset$
2. pour chaque sommet $u = 1, 2, \cdots, n$ faire

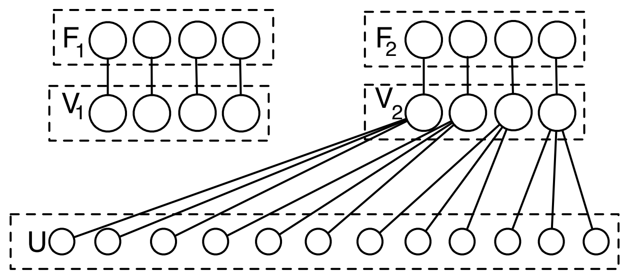
   si u a un voisin v dans G tel que $u < v$, alors $S \leftarrow S \cup \{u\}$
3. Retourner S

1. Une *numérotation des sommets* $L : V \to [1, \ldots, n]$ est une fonction injective qui donne des numéros distincts à chaque sommet de 1 à n.

Question 2.1 Montrer que S est une couverture de sommets.

Il suffit de raisonner par contradiction. Si S n'est pas une couverture de sommets, alors il existe une arête (u, v) non couverte par S. Sans perte de généralité, supposons que $u < v$. Lorsque le sommet u est traité par l'algorithme, comme u a un voisin v dans G tel que $u < v$, u est dans S et (u, v) est une arête couverte dans S. Ceci mène à une contradiction.

Question 2.2 Considérons le graphe suivant : L'ensemble des sommets est décomposé en 5 sous-ensembles : V_1, V_2, F_1, F_2 de 4 sommets et U de 12 sommets.
Tous les sommets de V_1 sont voisins de tous les sommets U (par soucis de lisibilité, ces arêtes ne sont pas dessinées dans la figure). Chaque sommet de V_1 (resp. V_2) a un seul voisin de F_1 (resp. F_2) et réciproquement. Chaque sommet de V_2 a trois voisins dans U et chaque sommet de U a un seul voisin dans V_2.



(a) Donner une couverture de sommets S^* de plus petite cardinalité.

(b) Donner la numérotation des sommets de telle façon que la couverture de sommets S retournée par l'algorithme $\mathcal{A}$ soit la plus grande possible.

(c) Soit $k > 2$ une constante. Construire un graphe tel que le nombre de sommets de la couverture de sommets S retournée par l'algorithme $\mathcal{A}$ soit au moins k fois plus grand que celui de la couverture de sommets S^* de plus petite cardinalité : $|S| > k\ OPT$.

1. $S^* = V_1 \cup V_2$ et $|S^*| = 8$
2. On numérote les sommets de la manière suivante : les sommets de F_2 de 1 à 4, les sommets de V_2 de 5 à 8, les sommets de F_1 de 9 à 12, les sommets de V_1 avec les numéros 21, 22, 23 et 28 et les sommets de U avec les numéros 24, 13, 14, 25, 15, 16, 26, 17, 18, 27, 19 et 20 de gauche à droite. On vérifie qu'alors tous les sommets, sauf celui de numéro 28, sont couverts par l'algorithme, ce qui est optimal puisqu'on ne peut pas couvrir tous les sommets.
3. On peut prendre un graphe en forme d'étoile composé de $k + 1$ sommets.

Question 3. Donner un algorithme polynomial qui étant donnés un graphe G et un sommet u, retourne une clique maximale contenant u dans G. Donner la complexité de l'algorithme.

Clique_Max(G,u) ; Entrée : un graphe G, un sommet u
Sortie : Un ensemble de sommets C_u
Variables locales : b un booléen

1. $C_u \leftarrow \{u\}$; *// O(1) opérations*
2. Pour tout sommet v faire *// la boucle est exécutée $O(|V|)$ fois*

(a) $b \leftarrow True$; *// O(1) opérations*

(b) Pour chaque sommet x de C_u *// la boucle est exécutée $O(|V|)$ fois au pire*

si x n'est pas voisin de v alors $b \leftarrow False$; *// O(1) opérations*

(c) Si $b == False$ alors $C_u \leftarrow C_u \cup \{v\}$; *// O(1) opérations*

3. Retourner C_u ; *// O(1) opérations*

Complexité : $O(|V|^2)$ opérations

Nous appelons une *partition en cliques maximales* du graphe $G = (V, E)$, une partition des sommets V en sous-ensembles disjoints $\mathcal{C} = \{C_1, \ldots, C_k\}$ tels que

1. $\bigcup_{i=1}^{k} C_i = V$;
2. pour tout couple (i, j) tel que $1 \leq i < j \leq k$, nous avons $C_i \cap C_j = \emptyset$;
3. chaque sous-graphe induit par C_i dans G est une clique maximale ;

Question 4. Donner un algorithme polynomial qui retourne une partition en cliques maximales de G.

Entrée : un graphe G
Sortie : Un ensemble de partition en cliques $\mathcal{C}$
Variables locales : Q, C deux ensembles de sommets

1. $Q \leftarrow V$ *// $O(|V|)$ opérations au pire*
2. $C \leftarrow \emptyset$; *// O(1) opérations*
3. Tant que $Q \neq \emptyset$ faire *// la boucle est exécutée $O(|V|)$ fois au pire*
   (a) Choisir un sommet u de Q ; *// O(1) opérations*
   (b) $C \leftarrow Clique\_Max(G, u)$; *// $O(|V|^2)$ opérations au pire*
   (c) $Q \leftarrow Q \backslash C$; *// $O(|V|^2)$ opérations au pire*
   (d) $\mathcal{C} \leftarrow \mathcal{C} \cup \{C\}$ *// O(1) opérations*
4. Retourner $\mathcal{C}$; *// O(1) opérations*

Complexité : $O(|V|^3)$ opérations

Une clique qui ne contient qu'un seul sommet est appelée une *clique triviale*. Dans la figure 1, les cliques C_4 et C_2 sont triviales.

Soit $\mathcal{C} = \{C_1, \ldots, C_k\}$ une partition en cliques maximales du graphe G. Considérons C' l'ensemble de sommets qui sont dans une clique non triviale de la partition C. Dans la figure 1, C' vaut $\{1, 2, 4, 5, 6, 7\}$.

Question 5. Montrer que C' est une couverture de sommets.

Soit e une arête de G. Deux cas se posent :

1. Si e est une arête dont au moins une de ces extrémités sont dans des cliques non-triviales, alors e est couvert par un sommet de C'.

2. Si e est une arête dont les deux de ces extrémités sont dans des cliques triviales, alors, e formerait une clique non-trivial. Ceci amène à une contradiction avec la partition en cliques maximales du graphe G.

Nous ne pouvons prouver que $|C'| < 2OPT$. En effet, il suffit de considérer le graphe ayant uniquement une arête (u, v). Sa partition en cliques maximales se réduit en une clique composée de l'arête (u, v). Alors $C' = \{u, v\}$ et une solution optimale contient simplement un seul noeud. Donc nous avons $|C'| = 2OPT$ dans ce cas.

Question 6. Montrer que $|C'| \leq 2OPT$ si G possède au moins une arête. *Indication : Trouver une relation entre OPT et le nombre de cliques non-triviales.*

Rappellons que pour chaque clique C non triviale de $\mathcal{C}$, toutes couvertures de sommets contient au moins $|C| - 1$ sommets.

Soit c le nombre de cliques non-triviales. Nous avons l'inéquation suivante ;

$$OPT \geq |C'| - c \tag{1}$$

Nous pouvons réécrire l'équation (précédente) (1) comme suit $OPT + c \geq |C'|$.

Comme chaque clique possède au moins deux sommets, une solution optimale à au moins un sommet dans une clique non triviale. Ainsi, nous avons $OPT \geq c$. Ce qui nous permet de déduire que $2OPT \geq |C'|$.

Question 7. Trouver un graphe tel que $|C'| = 2OPT$?

Il suffit de considérer le graphe ayant uniquement une arête (u, v). Sa partition en cliques maximales se réduit en une clique composée de l'arête (u, v). Alors $C' = \{u, v\}$ et une solution optimale contient simplement un seul noeud. Donc nous avons $|C'| = 2OPT$ dans ce cas.

Question 8. Qu'en déduisez-vous des trois questions précédentes ?

L'algorithme proposé est une 2-approximation du problème de couverture de sommets, et la borne 2 est optimale pour cet algorithme.