TD 7 : Approximation pour le problème de couverture

Correction

1 Algorithmes naïfs

Question 1 : **Correction** : Pour un r fixe, chaque sommet de L est connecté à au plus un sommet de R_i avec $i \in [1, r]$, donc le degré maximal est r.

Le nombre de sommets de R est $\sum_{i=1}^{r} \lfloor \frac{r}{i} \rfloor \leq r \sum_{i=1}^{r} \frac{1}{i}$, donc r fois le r-ième nombre harmonique.

Une autre approximation du nombre de sommets est $\sum_{i=1}^{r} \lfloor \frac{r}{i} \rfloor \leq \sum_{i=1}^{r} i = \frac{r(r+1)}{2}$.

La couverture minimale des sommets est L car :

— comme G_r est biparti, $\forall e \in E, e \cap L \neq \emptyset$ et $e \cap R \neq \emptyset$.

— L est minimal car, si on prend les arcs entre L et R_1, il y a r qui n'ont pas de sommet en commun, donc la couverture est de taille au moins $r = |L|$.

Question 2 : **Correction** : L'algorithme `naïf` peut prendre à chaque tour une arête $\{u, v\}$ avec $u \in L$ et $v \in R$ et ajouter v à C. Comme chaque sommet de R est connecté à L, on obtiendra à la fin $C = R$.

Comme il s'agit d'un problème de minimisation, la garantie de performance est $\frac{C(sol)}{C^*} = \frac{rH_r}{r} = \sum_{i=1}^{r} \frac{1}{r} \leq \ln(r) - \ln(2) + 1 \in \Omega(\log(r))$

Question 3 : **Correction** : L'algorithme `moinsnaïf` commence par le sommet le sommet u de R_r qui a un degré maximal ($= r$) et l'ajoute à C. En supprimant les arêtes adjacentes de u on diminue le degré des sommets de L de 1, donc il auront un degré d'au plus $r-1$. On continue donc en choisissant les sommets de $R_{r-1}, R_{r-2}, \ldots, R_1$. On démontre qu'après avoir retiré $R_r, \ldots R_k$, le degré maximal des noeuds restant en (V, E') est $k-1$.

Question 4 : **Correction** : Le problème de décision sous-jacent est : "Etant donné un poids P, décider s'il existe une couverture C tel que $\sum_{v \in C} \leq P$".

La réduction à 3-SAT avec n variable $X = \{x_1, \ldots, x_n\}$ construit un graphe $G = (V, E)$ tel que :

— pour tout x_i, V contient un sommet x_i et $\neg x_i$, i.e., un sommet pour chaque littéral,

— $\{x_i, \neg x_i\} \in E$,

— chaque clause $\ell_i \vee \ell_j \vee \ell_k$ est codée par une clique de taille 3 $\{\ell_i, \ell_j\} \in E$, $\{\ell_j, \ell_k\} \in E$, $\{\ell_k, \ell_i\} \in E$.

Le poids de chaque sommet/littéral est de 1.

On fixe le poids $P = n + k$ avec n est le nombre de variables et k le nombre de clauses. Une couverture C du graphe G doit inclure forcément un sommet par littéral pour couvrir les arêtes $\{x_i, \neg x_i\}$, Le codage des clauses demande deux littéraux pour couvrir 3 arêtes de la clique. Donc au plus $n + k$ sommets. La couverture nous donne la valuation (plus précisément un ensemble de valuations) des littéraux.

Question 5 : **Correction** : Soit le poids $w(v_i) = n - i$; l'algorithme prendra v_n, puis $v_{n-1}, \ldots$ jusqu'à v_1. La couverture trouvé est de taille n.

2 Couvertures dans les graphes bipartis

Question 6 : <u>**Correction**</u> : Dans $G' = (V', E')$, tous les poids des arcs de G sont les mêmes, M.

(Point 1). Supposons par absurde qu'il existe $\{u, v\} \in E$ avec $u \in S$ et $v \in T$. Soit $S' = S \cup \{v\}$ et $T' = T \setminus \{v\}$. Le poids de la coupe (S', T') est

$$\begin{aligned}
P(S', T') &= \sum_{u' \in S', v' \in T', (u', v') \in E'} w'(u', v') \\
&= \sum_{u' \in S, v' \in T', (u', v') \in E'} w'(u', v') + \overbrace{\sum_{v' \in T', (v, v') \in E'} w'(v, v')}^{= w'(v,t) = w(v)} \\
&= \sum_{u' \in S, v' \in T' \cup \{v\}, (u', v') \in E'} w'(u', v') - \sum_{u' \in S, (u', v) \in E'} w'(u', v) + w(v) \\
&= P(S, T) - \overbrace{\sum_{u' \in S, (u', v) \in E'} w'(u', v)}^{\geq M > w(v)} + w(v) \\
&< P(S, T) \text{ contradiction}
\end{aligned}$$

(Point 2) Soient $L' = L \cap T$ et $R' = R \cap S$. Montrons que $R' \cup L'$ est une couverture, càd pour tout $(u, v) \in E \subset L \times R$, $u \in R' \cup L'$ ou $v \in R' \cup L'$.

— Cas $u \in S$ alors $v \in S$ (cf. Point 1), donc $v \in R'$.

— Cas $u \notin S$ alors $u \in T$, donc $u \in L'$.

Quel est le coût le la couverture $C = R' \cup L'$? Pour une coupe minimale (S, T) de G', $L = \underbrace{(L \cap T)}_{L'} \uplus \underbrace{(L \cap S)}_{L''}$ et $R = \underbrace{(R \cap T)}_{R''} \uplus \underbrace{(R \cap S)}_{R'}$ avec absence d'arc entre L'' et R'' (d'après le Point 1), en plus de l'absence d'arc entre L' et L'' respectivement R' et R'' (car G biparti) et d'arc entre L' et R' (d'après définition de E'). Donc (un dessin peut aider) :

$$\begin{aligned}
P(S, T) &= \sum_{u \in S, v \in T, (u, v) \in E'} w'(u, v) \\
&= \sum_{v \in (T \cap L) = L', (s, v) \in E'} w'(s, v) + \sum_{u \in (S \cap R) = R', (u, t) \in E'} w'(u, t) + \sum_{u \in (S \cap L) = L'', v \in (T \cap R) = R'', (u, v) \in E'} w'(u, v) \\
&= \sum_{v \in L'} w(v) + \sum_{u \in R'} w(u) = \text{poids de la couverture } R' \cup L'
\end{aligned}$$

Supposons par absurde qu'il existe une couverture C' de G telle que $w(C') < w(R' \cup L')$. Soient $S' = \{s\} \cup (C' \cap R) \cup L \setminus (L \cap C')$ et $T' = \{t\} \cup (L \cup R) \setminus S'$. Alors (S', T') est une s-t coupe de G' dont le poids est :

$$\begin{aligned}
P(S', T') &= \sum_{u \in S', v \in T', (u, v) \in E'} w'(u, v) \\
&= \sum_{v \in (C' \cap L), (s, v) \in E'} w'(s, v) + \sum_{u \in (C' \cap R), (u, t) \in E'} w'(u, t) + \overbrace{\sum_{u \in L \setminus C', v \in R \setminus C', (u, v) \in E'} w'(u, v)}_{=0} \\
&= w(C') < w(R' \cup L') = P(S, T)
\end{aligned}$$

(pas d'arc (u, v) t.q. $u \notin C'$ et $v \notin C'$ car C' couverture), contradiction avec (S, T) une s-t coupe de poids minimal.

Question 7 : **Correction** : Le lien est : flot maximal $=$ s-t coupe de poids minimal. On donne l'intuition et puis l'algorithme.

Soit G' un graphe orienté construit à partir du graphe biparti G comme indiqué dans l'énoncé et (S, T) une s-t coupe de G' de poids $P(S, T)$.

Propriété 1 : Pour tout flot f et pour toute s-t coupe (S, T), $|f| \leq P(S, T)$.

Preuve : Notons $f(S, T) = \sum_{u \in S, v \in T, (u,v) \in E'} f(u, v)$. Alors :

$$
\begin{aligned}
P(S,T) \geq f(S,T) = & & \text{car } f(u,v) \leq w'(u,v) \\
&= \sum_{v \in T \cap L} f(s,v) + \sum_{u \in S \setminus \{s\}, v \in T} f(u,v) & \text{car pas d'arc } (s,v), v \in R \\
&= \sum_{v \in T \cap L} f(s,v) + \sum_{u \in S \cap L, v \in T \cap R} f(u,v) + \sum_{u \in S \cap R} f(u,t) & \text{car pas d'autres arcs} \\
&\geq \sum_{v \in T \cap L} f(s,v) + \sum_{v \in S \cap L} f(s,v) & \text{par conservation de flots} \\
&= \sum_{v \in (T \cap L) \cup (S \cap L)} f(s,v) = |f|
\end{aligned}
$$

Il nous reste à montrer que

Propriété 2 : Si f est un flot maximal alors $|f| = P(S, T)$ pour une coupe s-t de coût minimal.

Preuve : Remarquons que si démontre qu'il existe une s-t coupe (S, T) ayant la propriété $|f| = P(S, T)$ alors elle est de coût minimal par la propriété 1.

Pour obtenir la coupe (S, T), remarquons qu'un chemin entre s et t se décompose comme suit :

$$s \frac{w'_1}{f_1} u \in L \frac{w'_2}{f_2} v \in R \frac{w'_3}{f_3} t$$

Pour que le flot soit maximal, alors :

$$f_1 = w'_1 \quad \bigvee \quad f_2 = w'_2 \quad \bigvee \quad f_3 = w'_3 \tag{1}$$

sinon on pourrait augmenter les flots individuels et avoir un meilleur flot.

Mais $f_2 = w'_2$ est impossible par définition de w'_2 qui est $w'_2 = 1 + \sum_{v \in R} w(v)$.

On définit :

$$S = \{s\} \cup \{u \in L \mid w'(s,u) > f(s,u)\} \cup \{v \in R \mid (u,v) \in E, w'(s,u) > f(s,u)\} \tag{2}$$
$$T = V' \setminus S \tag{3}$$

Intuitivement, T inclut les sommets u de L tels que $f(s, u) = w'(s, u)$. Alors :

$$
\begin{aligned}
P(S,T) &= \sum_{u \in S, v \in T, (u,v) \in E'} w'(u,v) \\
&= \sum_{v \in T \cap L} w'(s,v) + \sum_{u \in S \cap R} w'(u,t) & \text{par } (2) - (3) \\
&= \sum_{v \in T \cap L} f(s,v) + \sum_{v \in S \cap R} f(v,t) & \text{par } (1) \\
&= \sum_{v \in T \cap L} f(s,v) + + \sum_{u \in S \cap L} f(s,u) & \text{conservation de flot} \\
&= |f|
\end{aligned}
$$

Pour calculer le flot maximal, on a l'algorithme de Ford-Fulkerson, d'où le résultat demandé.

3 Algorithme d'approximation

Question 8 : **Correction** : Soit C une double couverture de poids minimal. Supposons par absurde qu'il existe $v \in V$ avec $C(v) \geq 3$. On définit C' à partir de C tel que $C'(v) = 2$; C' satisfait encore la conditions $\forall \{u, v\} \in E \; C'(u) + C'(v) = C(u) + 2 \geq 2$. On a obtenu une double coupe de poids plus petit, contradiction avec C de poids minimal.

Question 9 : **Correction** : Soit C une double couverture de G.

Pour tout $\{u, v\} \in E$, $C(u) + C(v) \geq 2$ donc $C(u) > 0$ ou $C(v) > 0$ et donc u ou v sont dans $Supp(C)$.

Question 10 : **Correction** : Soit C' une couverture de G de poids minimal p'. On définit la double couverture de C'' de G telle que pour tout $u \in C'$, $C''(u) = 2$. Alors C'' est une double couverture de G et son poids est $2p'$. Comme C est une double couverture de poids minimal, alors $p \leq 2p'$.