

TP2 : Algorithmes pour la couverture par sommets

Dans ce TP nous allons étudier en pratique le problème de la couverture par sommets de coût minimum pour les graphes non orientés. Ce problème a été présenté en cours magistral. On a un graphe $G = (V, E)$ et une fonction de coût $c : V \rightarrow \mathbb{N}$ qui assigne à chaque sommet un coût entier (positif ou nul). Le but est de sélectionner un ensemble S de sommets de G tel que chaque arête a au moins une de ses deux extrémités dans S . Par ailleurs on cherche à minimiser le coût total de la solution, c'est-à-dire minimiser la somme $\sum_{v \in S} c(v)$. Voir la Figure 1 pour un exemple.

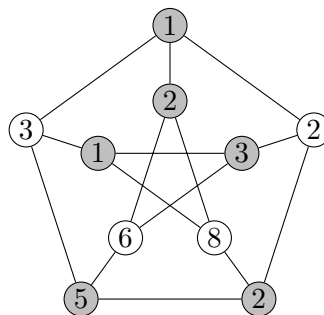


FIGURE 1 – Les sommets grisés forment une couverture par sommets de coût 14.

On va utiliser Sagemath pour coder et comparer différents algorithmes qui calculent une couverture par sommets (voir en bas du document pour quelques exemples de syntaxe Sagemath pour les PL et les graphes). Pour un graphe G , la fonction de coût des sommets, peut être représentée en Python par un dictionnaire avec comme clés, les sommets et comme valeurs, les coûts.

Pour tester vos algorithmes, vous pouvez créer des graphes. Sagemath dispose de fonctions pour cela. Par exemple, `G=graphs.PetersenGraph()` génère le fameux graphe de Petersen de la Figure 1.

`G=graphs.RandomGNP(20,0.2)` génère un graphe aléatoire à 20 sommets où chaque arête est présente avec probabilité $0.2 = 1/5$. Vous pouvez consulter la bibliothèque de création de graphes de Sagemath https://doc.sagemath.org/html/en/reference/graphs/sage/graphs/graph_generators.html pour d'autres exemples.

Vous pouvez générer un dictionnaire de coûts aléatoires c dont les clés sont les sommets du graphe et les valeurs sont les coûts. On peut utiliser la fonction `randrange(1,100)` qui renvoie un entier entre 1 et 100 au hasard.

Si vous avez une liste de sommets S (votre solution par exemple) d'un graphe G vous pouvez afficher le graphe en colorant les sommets de S d'une couleur spécifique, par exemple en rouge :

```
G.plot(vertex_colors={'red':S}).
```

Exercice 1 (Programme linéaire en nombres entiers).

Coder et tester une fonction `CS_PLNE(G,c)` qui prend en entrée un graphe et son dictionnaire de coûts, et renvoie une solution pour la couverture par sommets (sous forme de liste par exemple).

Cette solution sera calculée en définissant le PLNE du problème, défini de la façon suivante (avec une variable x_v pour chaque sommet v). Le graphe est noté $G = (V, E)$ où V est l'ensemble des sommets et E est l'ensemble des arêtes.

minimiser : $\sum_{v \in V} x_v c(v)$ tel que : <table style="display: inline-table; vertical-align: middle; margin-left: 20px;"> <tr> <td>$x_u + x_v$</td> <td>\geq</td> <td>1</td> <td>$\forall uv \in E$</td> </tr> <tr> <td>x_v</td> <td>\leq</td> <td>1</td> <td>$\forall v \in V$</td> </tr> <tr> <td>x_v</td> <td>\geq</td> <td>0</td> <td>$\forall v \in V$</td> </tr> <tr> <td>x_v</td> <td>\in</td> <td>\mathbb{N}</td> <td>$\forall v \in V$</td> </tr> </table>	$x_u + x_v$	\geq	1	$\forall uv \in E$	x_v	\leq	1	$\forall v \in V$	x_v	\geq	0	$\forall v \in V$	x_v	\in	\mathbb{N}	$\forall v \in V$
$x_u + x_v$	\geq	1	$\forall uv \in E$													
x_v	\leq	1	$\forall v \in V$													
x_v	\geq	0	$\forall v \in V$													
x_v	\in	\mathbb{N}	$\forall v \in V$													

Voir en bas du document pour la syntaxe d'un PLNE dans Sagemath.

Notez que cette méthode va calculer la solution optimale. Par contre, cet algorithme prend un peu de temps car résoudre un PLNE n'est pas du tout anodin. Jusqu'à quelle taille de graphes obtient-on la solution en un temps raisonnable (disons, quelques secondes) ?

Exercice 2 (Relaxation linéaire).

On va maintenant calculer une autre solution, mais cette fois on va utiliser la méthode de *l'arrondi de la relaxation linéaire* du PLNE vue en cours. Pour cela on résout d'abord le PL qui a la même formulation que le PLNE de l'exercice précédent, sauf qu'on ne rajoute pas la contrainte " $x_v \in \mathbb{N}$ " (les variables sont non-entières).

Une fois le PL résolu, on va "arrondir" le résultat : pour chaque variable x_v , si $x_v \geq 1/2$ on prend v dans la solution, et si $x_v < 1/2$ on ne prend pas v . Cette méthode renvoie une solution dont le coût est au pire deux fois l'optimum, mais le temps de calcul devrait être bien plus rapide.

Coder et tester une fonction `CS_arrondiPL(G,c)` qui prend un graphe et un dictionnaire de coûts en entrée, et qui renvoie une solution obtenue par cette méthode.

Remarque. Bien souvent, si la solution est grande, cette méthode a tendance à renvoyer tous les sommets. Sur certains graphes (notamment les graphes bipartis) le PL classique renvoie déjà une solution à valeurs entières, et donc optimale. Ces deux cas ne se présentent pas sur l'exemple de la figure suivante.

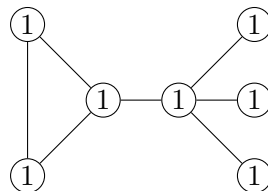


FIGURE 2 – Un graphe intéressant pour tester la relaxation linéaire.

Exercice 3 (Algorithme glouton).

On va maintenant utiliser une des heuristiques gloutonnes vues en cours, qui sélectionne à chaque étape le sommet qui couvrira le plus de nouvelles arêtes (en pondérant par son coût) :

- $S = \emptyset$
- Tant que toutes les arêtes ne sont pas couvertes :
 - Calculer le "score" de chaque sommet v : c'est le ratio "nombre d'arêtes non couvertes par S qui seraient couvertes par v " / $c(v)$. On pourra le stocker, par exemple, dans un dictionnaire.
 - Choisir un sommet v qui maximise le score
 - $S = S \cup \{v\}$

- Renvoyer S

Coder et tester cette heuristique dans une fonction `CS_greedy(G,c)` qui renvoie la liste des sommets de la solution obtenue.

Exercice 4 (Comparaison des algos).

On peut maintenant comparer la qualité et la performance en temps des différents algorithmes. Pour mesurer le temps de calcul dans une feuille de calcul Sagemath, on peut utiliser :

`timeit.eval("mafonction()",number=1)` (l'argument "number" est le nombre de fois que le même code sera exécuté : si on l'exécute plus qu'une fois il nous affichera le meilleur temps obtenu).

Testez vos différentes fonctions sur une série de graphes de plus en plus grands (par exemple 10,20,50,100,200 sommets...). Pour chaque graphe, lancer chaque fonction implémentée, notez le temps et la qualité de la solution, puis remplissez un tableau qui ressemblera à celui-ci :

type de graphe	# sommets	algo PLNE		algo relax linéaire		algo glouton	
		coût solution	temps de calcul	coût	temps	coût	temps
aléatoire	50	2.5	60	3.5	30	4.3	20

Exemples de syntaxe Sagemath

Sagemath est très pratique pour manipuler des graphes, et s'interface avec des solveurs de PL. Pour le lancer, écrire la commande `sage`, puis, dans la console de sage, écrire `notebook()`. Cela lance une interface (basée sur `jupyter notebook`) dans votre navigateur. Vous pouvez ensuite créer une nouvelle feuille de calcul en cliquant sur `New worksheet`. Ensuite on code en Python.

On peut définir et résoudre un PL dans Sagemath de la façon suivante :

```
p = MixedIntegerLinearProgram(maximization=False)           #crée un PL de minimisation
vars = p.new_variable(integer=True, nonnegative=True)       #vars est un dictionnaire qui
                                                            #contiendra les variables du PL
                                                            #ici ce sont des variables
                                                            #entières et non-négatives

p.set_objective( sum(3*vars[i] for i in [0,1]) + vars[2] )   #fonction objectif
p.add_constraint( vars[0] + 2*vars[1] >= 4 )                #contraintes
p.add_constraint( 5*vars[2] - vars[1] >= 8 )
opt = p.solve()                                              #résoudre le PL
print(opt)                                                  #afficher la valeur optimale
print(p.get_values(vars))                                   #afficher la valeur
                                                            #des variables
```

Cela correspond au programme linéaire suivant :

minimiser	$3vars[0]$	+	$3vars[1]$	+	$vars[2]$	
tel que	$vars[0]$	+	$2vars[1]$			≥ 4
			$-vars[1]$	+	$5vars[2]$	≥ 8
	$vars[0]$					≥ 0
			$vars[1]$			≥ 0
					$vars[2]$	≥ 0
	$vars[0]$					$\in \mathbb{N}$
			$vars[1]$			$\in \mathbb{N}$
					$vars[2]$	$\in \mathbb{N}$

Pour créer et manipuler un graphe non orienté dans Sagemath on peut écrire :¹

```
G = Graph(3)                                                #Crée un graphe orienté à 3 sommets (nommés 0,1,2 par défaut)
G.add_vertex(name="toto")                                   #ajoute un sommet nommé "toto"
G.add_edge(1,2)                                             #Ajoute l'arête 1-2
for e in G.edges():                                         #boucle sur la liste des arcs de G
    print(e[0],e[1])                                       #affiche l'origine, la destination de l'arête e
for v in G.vertices():                                     #boucle sur la liste des sommets de G
    print(G.neighbors(v))                                  #affiche la liste de tous les voisins de v
    print(G.edges_incident([v]))                          #affiche la liste des arêtes qui touchent v
```

1. Voir <https://doc.sagemath.org/html/en/reference/graphs/index.html> pour plus de fonctions, si nécessaire.