SI 4^eannée **Année 2020-2021**
Complexité et Algorithmique

Preuves de NP-complétude

Corrigé partiel de la feuille de travaux dirigés n^{o}3

5 novembre 2020

1. On réduit **Partition** à **Somme minimum des carrés**. Soit donc $A, s(a)\ \forall a \in A$ une instance de **Partition**. Soit $S = \sum_{a \in A} s(a))$. L'instance correspondante de **Somme minimum des carrés** est A, $k = 2$, $J = \frac{S^2}{2}$. Cette restriction de **Somme minimum des carrés** devient alors identique à **Partition**. En effet Si on a une partition ou la somme d'un coté est $\frac{S}{2} - x$ et de l'autre coté $\frac{S}{2} + x$ alors la somme des carrés sera $\frac{S^2}{2} + 2x^2$. Donc si la somme est J c'est qu'on avait une partition. L'autre sens est immédiat.

2. On réduit **Clique** au problème de l'**Isomorphisme de sous-graphes**. Soient donc G et J une instance de **Clique**. L'instance correspondante de **Isomorphisme de sous-graphes** est la donnée de $G_1 = G$ et de $G_2 = K_J$ où K_J représente le graphe complet à J sommets. Par construction, toute instance positive de **Clique** sera une instance positive de **Isomorphisme de sous-graphes** et, réciproquement, étant donné G_1 et K_J, si G_1 contient un sous graphe isomorphe à K_J, il contient une clique de J sommets. Le caractère polynomial de la transformation est immédiat, il suffit de recopier G et de construire un codage adéquat de K_J.
Une autre preuve consisterait de la réduction de **Stable**. En effet si G_2 est un stable alors la réponse est oui ssi G_1 admet un stable de la taille demandée.

3. On réduit **Chaîne Hamiltonienne** au problème de l'**Isomorphisme de sous-graphes partiels**. Soit donc G une instance de **Chaîne Hamiltonienne**. L'instance correspondante de **Isomorphisme de sous-graphes partiels** est la donnée de $G_1 = G$ et G_2 la chaîne à $n = |V(G)|$ sommets. Cette restriction de **Isomorphisme de sous-graphes partiels** devient alors identique à **Chaîne Hamiltonienne**.
Une autre preuve consisterait de la réduction de **Clique**. En effet si G_2 est un graphe complet alors la réponse est oui ssi G_1 admet un clique de la taille demandée. Par contre, comme il s'agit d'isomorphisme de sous-graphe partiel, on ne peut pas utiliser le problème **Stable** car le "partiel" permet de supprimer des arêtes, pour obtenir un stable de n'importe quel taille (qui ne dépasse pas la taille de G_1).

4. On réduit **Chaîne Hamiltonienne** au problème de l'**Arbre couvrant de degré borné**. Soit donc G une instance de **Chaîne Hamiltonienne**. L'instance correspondante de **Arbre couvrant de degré borné** est la donnée de $G = G$ et $k = 2$. Cette restriction de **Arbre couvrant de degré borné** devient alors identique à **Chaîne Hamiltonienne**.

5. On réduit **Partition** à **Ordonnancement de tâches**. Soit donc $A, s(a)\ \forall a \in A$ une instance de **Partition**. L'instance correspondante de **Ordonnancement de tâches** est $t_i = s(a_i)$ pour $1 \leq i \leq |A|$, $n = 2$ et $T = \frac{1}{2} \sum_{a \in A} s(a)$. Cette restriction de **Ordonnancement de tâches** devient alors identique à **Partition**.

6. Il est possible de réduire **3-DM** au **Plus petit ensemble de tests**. Voici l'énoncé de **3-DM**.

Nom : Couplage à trois (3DM)
Instance : Un ensemble $M \subseteq W \times X \times Y$ avec W, X et Y disjoints de même cardinal q.
Question : Existe-t'il $M' \subseteq M$ tel que $|M'| = q$ et dont toutes les coordonnées sont deux à deux distinctes ?

Un exemple

Soit $W=\{A,B,C,D\}$, $X=\{1,2,3,4\}$, $Y=\{a,b,c,d\}$ et $|M|=6$ défini par :

	M_1	M_2
1)	$(A,2,a)$	$(A,2,a)$
2)	$(A,3,b)$	$(A,3,b)$
3)	$(B,1,d)$	$(B,1,d)$
4)	$(B,2,c)$	$(B,2,c)$
5)	$(C,1,a)$	$(C,2,a)$
6)	$(D,4,d)$	$(D,4,d)$

M_1 vérifie la propriété du couplage à trois pour les lignes $2), 4), 5)$ et $6)$ mais pas M_2.

La réduction

Soit une instance quelconque de **3-DM** avec W, X et Y disjoints de même cardinal q et $M \subseteq W \times X \times Y$. Les composants de base de notre instance de **3-DM** sont les triplets de M. L'ensemble des pannes possibles est alors constitué de l'union de W, X et Y. On effectue un *remplacement local* en associant à tout $m=(w,x,y)\in M$ une famille $\{w,x,y\}\in C$. On ajoute trois pannes supplémentaires w_0, x_0 et $y_0 \notin W\cup X\cup Y$ ainsi que deux nouveaux tests : $W\cup\{w_0\}$ et $X\cup\{x_0\}$.
L'instance complète de PLUS PETIT ENSEMBLE DE TESTS est :

- $P = W\cup X\cup Y\cup\{w_0,x_0,y_0\}$;
- $C=\{\{w,x,y\} : (w,x,y)\in M\}\cup\{W\cup\{w_0\}, X\cup\{x_0\}\}$
- $J=q+2$.

Cette instance peut être construite en temps polynomial pour une instance donnée de **3-DM**.
On déduit quelques contraintes sur la forme des instances positives de C' :

- C' doit contenir simultanément $W\cup\{w_0\}$ et $X\cup\{x_0\}$ car ce sont les deux seuls tests qui nous permettent de distinguer y_0 à la fois de w_0 et x_0 ;
- Il faut pouvoir distinguer w_0 (resp. x_0 et y_0) des autres éléments de W (resp. X et Y). Pour ce faire, il faut que tous les éléments de W (resp. X et Y) soient couverts par des tests de la forme $c\in C'\setminus\{W\cup\{w_0\}, X\cup\{x_0\}\}$ à prendre dans l'ensemble des familles de triplets. On dispose d'au plus $J-2=q$ tests de ce type.

Comme chacun des tests restants de C contient exactement un élément parmi W, X et Y, et comme W, X et Y sont disjoints et de même cardinal q, il s'ensuit que tout test de la seconde forme de C' doit correspondre à q triplets qui forment un couplage à trois de M. Réciproquement, étant donné un couplage à trois de M, les q tests correspondants de C peuvent être utilisés pour construire la famille souhaitée des $J=q+2$ tests. Ce qui implique que M contient un couplage à trois et donc qu'il existe une sous-famille adéquate de tests.
Ainsi, partant de notre instance positive ci-dessus de **3-DM**, on construit

- $P=\{A,B,C,D,1,2,3,4,a,b,c,d,w_0,x_0,y_0\}$
- $C=\{\{A,B,C,D,w_0\},\{1,2,3,4,x_0\}\}\cup\{\{A,2,a\},\{A,3,b\},\{B,1,d\},\{B,2,c\},\{C,1,a\},\{D,4,d\}\}$
- $J=4+2$

7. SSP est bien dans NP. En effet, il suffit de choisir de manière non-déterministe un sous-ensemble A' et de vérifier en temps polynomial que $\sum_{a\in A'} a = c$ la capacité totale.
On montre maintenant qu'on peut réduire **X3-SAT** à **SSP**. Soit donc x une instance de **X3-SAT** $C_1\wedge C_2\wedge\ldots\wedge C_n$ d'ensemble de variables $U=\{u_1,\ldots,u_m\}$ (n clauses et m variables). On va associer une taille (et donc un élément) à chaque littéral de l'instance de **X3-SAT**, i.e. une taille pour chaque variable et une taille pour la négation de chaque variable. Pour ce faire, on réserve les m digits de poids faible des tailles et on met le digit i de la taille t_j à 1 si t_j correspond à u_i ou à $\overline{u}_i$ et les autre digits de t_j à 0. De cette manière, il est possible d'exprimer la propriété qu'on ne peut pas choisir à la fois u_i et $\overline{u}_i$ en choisissant convenablement la valeur de la capacité, autrement dit si les m digits de poids faible de la capacité sont tous égaux à 1.
Par exemple, pour $U=\{a,b,c,d\}$ et $C_1\wedge C_2\wedge C_3$ avec $C_1=\{a,\overline{b},c\}$, $C_2=\{a,b,d\}$ et $C_3=\{\overline{b},c,d\}$, on construit

les tailles suivantes :

		d	c	b	a
a	t_1	0	0	0	1
$\overline{a}$	t_2	0	0	0	1
b	t_3	0	0	1	0
$\overline{b}$	t_4	0	0	1	0
c	t_5	0	1	0	0
$\overline{c}$	t_6	0	1	0	0
d	t_7	1	0	0	0
$\overline{d}$	t_8	1	0	0	0

De cette manière, tout choix de u_i ou de $\overline{u}_i$ mais pas des deux simultanément apparaîtra dans c.
Ensuite, on ajoute n digits supplémentaires à chacune des tailles qui vont nous permettre d'exprimer que le littéral u_i apparaît dans la clause C_k. On marque le digit $m+k$ de la taille t_j qui correspond au littéral u_i si u_i apparaît dans la clause C_k. En poursuivant l'exemple :

		C_3	C_2	C_1	d	c	b	a
a	t_1	0	1	1	0	0	0	1
$\overline{a}$	t_2	0	0	0	0	0	0	1
b	t_3	0	1	0	0	0	1	0
$\overline{b}$	t_4	1	0	1	0	0	1	0
c	t_5	1	0	1	0	1	0	0
$\overline{c}$	t_6	0	0	0	0	1	0	0
d	t_7	1	1	0	1	0	0	0
$\overline{d}$	t_8	0	0	0	1	0	0	0

Par cette construction, si le littéral u_i est vrai, la ou les clause(s) C_k où u_i apparaît est également à vrai et le $m+k^{\text{e}}$digit de la taille correspondant à u_i est à un. Ainsi, les valuations satisfaisantes de **X3-SAT** correspondent aux ensembles de tailles dont la somme est un entier à $m+n$ digits dont les m digits de poids faible sont tous égaux à 1 et dont les n digits de poids fort sont compris entre 1 et 3. Il faut donc trouver une manière de modifier le problème pour que leur somme sur les n digits de poids fort soit constante.
Pour ce faire, on va ajouter $2n$ tailles supplémentaires appelées *compensatrices* dont le $m+k^{\text{e}}$digit est à 1 si la taille t_j correspond à la clause C_k et 0 sinon.
En poursuivant notre exemple :

		C_3	C_2	C_1	d	c	b	a
a	t_1	0	1	1	0	0	0	1
$\overline{a}$	t_2	0	0	0	0	0	0	1
b	t_3	0	1	0	0	0	1	0
$\overline{b}$	t_4	1	0	1	0	0	1	0
c	t_5	1	0	1	0	1	0	0
$\overline{c}$	t_6	0	0	0	0	1	0	0
d	t_7	1	1	0	1	0	0	0
$\overline{d}$	t_8	0	0	0	1	0	0	0
C_1	t_9	0	0	1	0	0	0	0
C_1	t_{10}	0	0	1	0	0	0	0
C_2	t_{11}	0	1	0	0	0	0	0
C_2	t_{12}	0	1	0	0	0	0	0
C_3	t_{13}	1	0	0	0	0	0	0
C_3	t_{14}	1	0	0	0	0	0	0

Si on a un valuation satisfiable de x, on choisit les littéraux mis à vrai par la valuation et on prend respectivement une ou deux des tailles compensatrices correspondant à C_j pour que la somme des tailles soit $c = \underbrace{3\ldots3}_{m}\ \underbrace{1\ldots1}_{n}$.

Par exemple, une valuation satisfiable de x est $a\overline{b}\overline{c}\overline{d}$. On choisit donc les tailles $t_1, t_4, t_6, t_8, t_9, t_{11}, t_{12}, t_{13}$ et t_{14} pour que leur somme soit 3331111.
Réciproquement, soit un sous-ensemble dont la somme des tailles soit égal à c. Les m digits de droite assurent que cet ensemble contient une taille correspondant à chaque variable ou à sa négation mais pas les deux simultanément. On prend comme valuation de l'instance de **X3-SAT** correspondante celle qui met chaque littéral à vrai. De plus, comme pour chaque clause C_k, la somme des $(m+k)$-ième digits des tailles vaut 3 dont au plus deux peuvent provenir des tailles compensatrices, il existe au moins une taille qui correspond à un littéral avec un 1 au $(m+k)$-ième digit. Comme le littéral correspondant est vrai, la clause C_k est aussi vraie et nous avons donc bien une valuation satisfiable pour x.
Par exemple, si on choisit l'ensemble de tailles $\{t_2, t_3, t_5, t_7, t_9, t_{10}, t_{11}, t_{14}\}$ dont la somme est 3331111, on obtient une valuation satisfiable de x qui est $\overline{a}bcd$.
Comme la transformation peut être effectuée facilement en temps polynomial, nous avons montré que **X3-SAT**$\propto$**SSP**.
Remarque : la question était de faire la réduction **X3-SAT**$\propto$**SSP**. Si on voulait juste montrer la NP-difficulté, on aurait pu faire la réduction **Partition**$\propto$**SSP**, qui est immédiate. En effet **Partition** est le cas particulier de **SSP**, où la somme à atteindre est exactement la moitié de la somme totale.

8. La difficulté de la preuve est que dans **Score** les deux extrémités de la chaîne sont données. Si ce n'était pas le cas, ce serait très facile, car une chaine simple de $n-1$ arêtes est forcément une chaine hamiltonienne.
On peut utiliser la même réduction que dans la réduction de **Cycle hamiltonien** en **Chaîne hamiltonienne** (voir feuille n° 2), avec en plus en mettant poids 1 à chaque arête. Ainsi en partant d'un graphe G à n sommets dans lequel on pose la question de l'existence d'un cycle hamiltonien, on obtient à un graphe G' à $n+3$ sommets dans lequel on pose la question de l'existence d'une chaine simple de poids au moins $n+2$ (qui ne peut pas être autre chose qu'une chaine hamiltonienne).

9. Une réduction simple consiste à réduire **Partition** à **PP**. En effet, il suffit de doubler la valeur de chaque élément de A et fixer $t=1$. Dans le problème obtenu toutes les valeurs sont pairs, donc la somme des valeurs de tout sous-ensemble est aussi pair ainsi que la différence entre deux sommes. Et comme ce nombre pair doit être borné par 1, il ne peut être autre que 0. Donc la presque partition obtenue est une partition en deux parties égales.

10. Ici encore, on peut réduire **Partition** à **3-partition**. En effet, si on rajoute un élément $a' = \frac{1}{2}\sum_{a\in A} a$, alors chaqun des trois parties doit être de somme égale à a'. Ainsi a' ne peut qu'être dans un ensemble avec d'autres éléments (pour ne pas dépasser la somme demandée) dans lequel et donc la somme des deux autres ensembles est aussi $\frac{1}{2}\sum_{a\in A} a$.

11.

a) Une "prétendue solution" consiste en un routage, qui décrit pour tout couple de sommets la chaîne par laquelle ils communiqueront. Si le routage est donné, alors on peut facilement calculer la charge maximale, comme on peut le faire sur l'exemple.

   On peut remarquer qu'on peut avoir un très grand nombre de routages. En effet, entre deux sommets donnés, il peut y avoir 1 chaîne de longueur 1, $n-2$ chaînes de longueur 2, $(n-2)(n-3)$ de longueur 3, … et $(n-2)!$ chaînes simples de longueur $n-1$ (à condition que le graphe soit complet). Donc plus que $(n-2)!$ routages possibles. Si on considère les routage dans le réseau il y a au moins $(n-2)!^{\frac{n(n-1)}{2}}$ ce qui est une double exponentielle …

b) Comme le problème **RLP** est un cas spécial de **NLP** où le réseau est un anneau (un cycle) la réduction est immédiate, par la transformation identité.

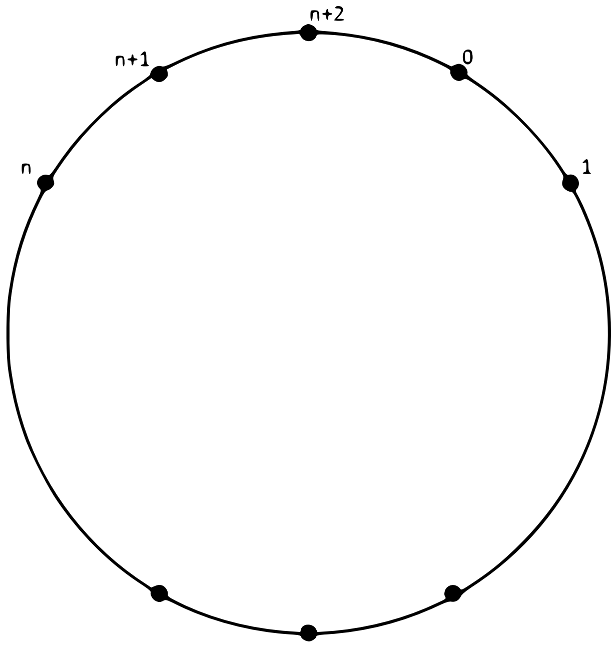
c) La différence principale entre les deux problèmes est que cette fois-ci le nombre de routages n'est que $2^{\frac{n(n-1)}{2}}$, puisque pour chaque sommet il faut décider si on va vers la gauche ou vers la droite. Ainsi la taille d'une "prétendue solution" est $\frac{n(n-1)}{2}$ au lieu de $\frac{n(n-1)}{2}(n-2)$, ce qui rend la génération non-déterministe plus simple.

d) La différence entre le problème **RLP** et le problème **SSRLP** consiste en la matrice des demandes. Alors que dans le premier on peut avoir n'importe quelle matrice symétrique avec des 0 sur la diagonale, dans le second on a une matrice avec des 0 partout, sauf sur une ligne i et la colonne i. Donc le second est un cas particulier du premier et à ce titre la réduction est assurée.

e) La transformation est donnée dans l'énoncé : Soient $a_1,\ a_2,\ \ldots,\ a_n$ les données du problème **Partition**. Soit $S = \frac{1}{2}\sum_{i=1}^{n} a_i$. Les sommets de l'anneau seront numérotés 0, 1, …, n+2. Les communications sont entre le sommet n+2 et les autres sommets. La valeur de la demande entre le sommet 0 et le sommet i est comme suit :

i	0	1	2	…	n	n+1
valeur	S	a_1	a_2	…	a_n	S

La valeur de C est fixée à $2S$. La transformation se fait en temps polynomial.



Pour les routage on a le choix entre deux solutions : pour communiquer entre le sommet i et le sommet n+2 on peut soit passer à gauche (par n+1 ou à droite par 0).

OUI ⇒ OUI : Soient A_1 et A_2 la partition de A. Le routage proposé :
— pour $1 \leq i \leq n$ le routage est par la gauche si $a_i \in A_1$ et par la droite sinon
— pour les nœuds 0 et n+2, le routage est par l'arête directe
Il suffit de vérifier que la charge maximum est au plus C. On fait la vérification arête par arête :
— par l'arête (n+2,0) passent les S messages de 0 et les S messages qui passent à droite, en tout $2S$.
— par l'arête (0,1) passent les messages qui passent à droite, donc S.
— par l'arête (1,2) passent au pire tous les message de la partition, en tout $2S$.
— même situation pour toutes les arêtes (i,i+1) avec $2 \leq i \leq n-1$.
— par l'arête (n,n+1) passent les messages qui passent à gauche, donc S.
— par l'arête (n+1,n+2) passent les S messages de n+1 et les S messages qui passent à gauche, en tout $2S$.
Ainsi le maximum est $2S$.

OUI ⇐ OUI : La quantité totale des communications de n+2 est $4S$ et ceci doit passer par deux arêtes avec une charge maximale chacune de $2S$. Ainsi la charge des arêtes (0,n+2) et (n+1,n+2) doit être $2S$ pour chacun. On peut se poser la question par quelle route communique n+2 avec 0 (et avec n+1). Si on suppose que cette communication se fait par le "tour du monde", on arrive à une contradiction. En effet, cela veut dire que par l'arête (0,n+2) passent d'autres communications en quantité $2S$ et elles arrivent toutes en passant par l'arête (0,1). Donc sur cette arête la charge est $3S$ ($2S$ en direction de 0 et S en direction de 1). Donc n+2 communique avec 0 (et avec n+1) par l'arête directe. Mais dans ce cas qui sont les autres communicants sur l'arête (0,n+2) ? Certains sommets i avec $1 \leq i \leq n$. La somme des communications de ces sommets est exactement S, donc on a une partition qui consiste en P_1 les sommets qui communiquent à gauche et P_2 ceux qui communiquent à droite.

f) Comme nous avons montré que **SSRLP** est dans NP et il est NP-difficile on peut conclure qu'il est NP-complet. De même pour **RLP** et aussi pour **NLP**.