

Rapport de TP 4MMAOD : Génération de patch optimal

“Question 1 : Modélisation par PLNE”

GOUTTEFARDE Léo (groupe 5)
PIELLARD Jérémie (groupe 2)

2 octobre 2015

Modélisation du problème restreint sous forme de PLNE

Données

On cherche à créer un patch P restreint minimal qui transforme un fichier F_1 de n lignes en un fichier F_2 de m lignes.

On note $F_1(i)$, la ligne $i \in 1 \cdots n$ du fichier 1 et $F_2(j)$ la ligne $j \in 1 \cdots m$ du fichier 2.

On définit $L(F_i(j))$ comme le nombre de caractères de la ligne $F_i(j)$ (incluant le caractère de fin de ligne).

Variables

On introduit les variables de coût suivantes :

- $c_s(i, j)$: modélise le coût de substitution de la ligne $F_1(i)$ par la ligne $F_2(j)$, nul si $F_1(i) = F_2(j)$ et égal à $c_s(i, j) = 10 + L(F_2(j))$ sinon.
- $c_a(i, j) = 10 + L(F_2(j))$: modélise le coût d'ajout de la ligne $F_2(j)$ après la ligne $F_1(i)$
- $c_d(i) = 10$: modélise le coût de destruction de la ligne $F_1(i)$ (identique pour tout i , peu utile donc)

On introduit les variables binaires (de valeur 1 ou 0) suivantes :

- $s(i, j)$: modélise la substitution (ou non) de la ligne $F_1(i)$ par la ligne $F_2(j)$
- $d(i)$: modélise la destruction (ou non) de la ligne $F_1(i)$
- $a(i, j)$: modélise l'ajout (ou non) de la ligne $F_2(j)$ après la ligne $F_1(i)$

Objectif à minimiser

On obtient la fonction objectif suivante :

$$\min_{s, d, a} \sum_{i, j} c_s(i, j) \cdot s(i, j) + \sum_{i, j} c_a(i, j) \cdot a(i, j) + 10 \sum_i d(i)$$

Contraintes

Pour l'ajout et la destruction, on peut d'abord distinguer plusieurs cas :

Cas 1 : Nombre minimal de destructions ($n > m$)

Si $n > m$, on aura minimum $n - m$ destructions à effectuer, soit un coût minimal de $10(n - m)$.

Cas 2 : Nombre minimal d'ajouts ($n < m$)

Si $n < m$, on aura minimum $m - n$ ajouts à effectuer, soit un coût minimal de

$$10(m - n) + \sum_{j=0}^m a(j) \cdot L(F_2(j)).$$

Cas 3 : Même nombre de lignes dans chaque fichier ($n = m$)

Si $n = m$, il n'y a pas forcément d'ajout ni de destruction de ligne.

Finalement, il ne peut pas y avoir à la fois une substitution et une destruction sur une même ligne i de F_1 .