Corrigé Exercice sur PCP

vraisemblablement rédigé par Marie van den Bogaard

Exercice 1 (PCP)
Montrer que le Problème de Correspondance de Post reste indécidable lorsque tous les mots des deux séquences ont pour longueur au plus 2. (Qu'en est-il si tous les mots ont pour longueur 2 ?)

Solution :

On réduit *PCP* à *PCP avec seulement des mots de longueur 2 au plus.* Soit $(u_1, \ldots, u_n), (v_1, \ldots, v_n)$ une instance de PCP.

Pour chaque mot u_i de longueur supérieure ou égale à 2, $u_i = a_i^1 \cdots a_i^{k_i}$, on remplace la paire (u_i, v_i) par les paires

$$(u_i^1, v_i^1) = (a_i^1, v_i\alpha_1), (u_i^2, v_i^2) = (a_i^2, \overline{\alpha_1}\alpha_2), \ldots (u_i^{k_i}, v_i^{k_i}) = (a_i^{k_i}, \overline{\alpha_{k_i-1}}),$$
$$(u_i^{k_i+1}, v_i^{k_i+1}) = (\alpha_1\overline{\alpha_1}, \epsilon), \ldots (u_i^{2k_i-1}, v_i^{2k_i-1}) = (\alpha_{k_i-1}\overline{\alpha_{k_i-1}}, \epsilon)$$

où les α_i, $\overline{\alpha_i}$ sont des nouveaux symboles, distincts deux à deux.

Après cette transformation, les mots u_i^j ont une longueur au plus deux et les seuls mots de longueur 2 sont de la forme $\alpha\overline{\alpha}$, le mot v_i^j correspondant étant ϵ. Montrons que la nouvelle instance de PCP a une solution ssi la première instance a une solution.

Si $u_{i_1} \cdots u_{i_k} = v_{i_1} \cdots v_{i_k}$ alors, $s_{i_1} \cdots s_{i_k} = \phi(t_{i_1} \cdots t_{i_k})$ où $s_j = u_j^1 \cdots u_j^{k_j}$, $t_j = v_j^1 \cdots v_j^{k_j}$ et ϕ est le morphisme qui envoie tous les mots $\alpha\overline{\alpha}$ sur ϵ. On obtient donc $t_{i_1} \cdots t_{i_k}$ par insertion dans $s_{i_1} \cdots s_{i_k}$ de mots $\alpha\overline{\alpha}$: $t_{i_1} \cdots t_{i_k} = s'_{i_1} \cdots s'_{i_k}$ où $s'_j = \theta_j^1 u_j^1 \cdots \theta_j^{k_j} u_j^{k_j} \theta_j^{k_j+1}$ et θ_j^i est un produit de mots de la forme $\alpha\overline{\alpha}$. On obtient alors une solution pour la nouvelle instance de PCP.

Réciproquement, si la nouvelle instance de PCP a une solution $s_{i_1} \cdots s_{i_k} = t_{i_1} \cdots t_{i_k}$, alors, comme toutes les lettres α sont suivies de $\overline{\alpha}$ dans les mots s_j, si $t_m = v_i\alpha_1$ alors il existe $p \geq 1$ tel que $t_m \cdots t_{m+p} = v\alpha_1 \cdot \overline{\alpha_1}\alpha_2 \cdots \overline{\alpha_r}$ et $\phi(s_m \cdots s_{m+p}) = u_i$. Il en résulte que $\phi(s_{i_1} \cdots s_{i_k}) = u_{j_1} \cdots u_{j_l} = \phi(t_{i_1} \cdots t_{i_k}) = t_{j_1} \cdots t_{j_l}$ et on a ainsi une solution de l'instance originale de PCP.

On procède ensuite de même pour chaque v_i de longueur strictement supérieure à 2. On obtient alors deux suites de mots, chacun d'eux étant de longueur au plus 2. Par symétrie, le raisonnement ci-dessus s'applique et cette nouvelle instance de PCP a une solution ssi l'instance de départ a une solution.

Exemple de codage :

i	1	2	3	4
u_i	a	b	ca	abc
v_i	ab	ca	a	c

Cette instance de PCP a une solution (12314).

Codage :

i	1	2	3	4	5	6	7	8	9	10
s_i	a	b	c	a	$\alpha_1\overline{\alpha_1}$	a	b	c	$\alpha_2\overline{\alpha_2}$	$\alpha_3\overline{\alpha_3}$
t_i	ab	ca	$a\alpha_1$	$\overline{\alpha_1}$	ϵ	$c\alpha_2$	$\overline{\alpha_2}\alpha_3$	$\overline{\alpha_3}$	ϵ	ϵ

Avec pour solution (1234156789[10])

Si tous les mots sont de longueur 2, le problème devient décidable. En fait, il n'a de solution que si, pour au moins un indice i, $u_i = v_i$.