6.9 Problème de correspondance de Post

Le problème de correspondance de Post (PCP) :

Donnée : Deux suites de mots finies $(u_1, \ldots, u_n)$ et $(v_1, \ldots, v_n)$ de même longueur

Question : existe-t-il un entier k et une suite d'indices $i_1, \ldots, i_k$ tels que

$$u_{i_1} \cdots u_{i_k} = v_{i_1} \cdots v_{i_k}$$

Par exemple : soient

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).

Théorème 6.9.1 *PCP est indécidable.*

On commence par montrer que le problème de correspondance de Post *modifié*, dans lequel on fixe $i_1 = 1$ est indécidable.

Théorème 6.9.2 *Le problème :*

Donnée *deux suites finies de mots* $u_1, \ldots, u_n, v_1, \ldots, v_n$

Question : *existe-t-il un entier* k *et une suite d'indices* $i_1, \ldots, i_k$ *tels que* $i_1 = 1$ *et* $u_{i_1} \cdots u_{i_k} = v_{i_1} \cdots v_{i_k}$ *?*

est indécidable

Pour cela, on réduit le probème suivant :

Donnée : le code d'une machine de Turing M qui efface son ruban en fin de calcul

Question : M s'arrête sur le mot vide

Formellement, une machine qui efface son ruban en fin de calcul est une machine qui contient un état spécial q_e. La machine

— nécrit jamais de blanc dans un état autre que q_e,
— n'écrit \$ qu'en lisant un \$,
— n'entre dans l'état q_e qu'en lisant un blanc, et, dans ce cas, se déplace à droite,
— depuis l'état q_e, elle écrit toujours un blanc et se déplace à gauche, sauf si la lettre lue est \$, auquel cas elle s'arrête.

L'indécidabilité du problème ci-dessus est laissée en exercice.

Avant de montre la réduction générale à PCP modifié, montrons celle-ci sur un exemple.

Exemple 6.9.1 Soit M la machine dont la table est donnée par :

δ	q_0	q_e
\$	$q_0, \$, \rightarrow$	accept
B	$0, q_e, \rightarrow$	$q_e, B, \leftarrow$
0		$q_e, B, \leftarrow$

Sur le mot vide, on obtient le calcul suivant (les configurations sont représentées par des mots wqw' dans lesquels w est la partie du ruban à gauche de la tête de lecture, q est l'état de contrôle et w' est la partie à droite de la tête de lecture, sans les blancs en fin de ruban).

$$q_0\$ \vdash \$q_0 \vdash \$0q_e \vdash \$q_e0 \vdash q_e\$ \vdash \text{accept}$$

Les suites de mots de PCP modifié dans la réduction qui va suivre seront donnés par :

i	u_i	v_i	
1	$\lhd\ q_0\$\star$	$\lhd$	
2	$\$q_0$	$q_0\$$	
3	$0q_e\star$	$q_0\star$	
4	$q_ea\star$	$aq_e\star$	$a \in \Sigma$
5	$q_ea\star$	$aq_eb\star$	$a, b \in \Sigma$
6	$\rhd$	$q_e\$\star\ \rhd$	
7	a	a	$a \in \Sigma$
8	$\star$	$\star$	

On obtient la séquence d'indices de PCP modifié comme suit :

i	1	2	8	7	3	7	4	5	6
u	$\lhd\ q_0\$\star$	$\$q_0$	$\star$	\$	$0q_e\star$	\$	$q_e0\star$	$q_e\$\star$	$\rhd$
v	$\lhd$	$q_0\$$	$\star$	\$	$q_0\star$	\$	$0q_e\star$	$\$q_e0\star$	$q_e\$\star\ \rhd$

Dans le cas général, on construit comme suit l'instance de PCP modifié :

	v_i	u_i	
1.	$\lhd$	$\lhd\ q_0\$w\star$	
2.$*$	a	a	pour $a \in \Sigma$
3.$*$	qa	$a'q'$	si $\delta(q,a) = q', a', \rightarrow$ et $a \neq B$
4.$*$	aqb	$q'ab'$	si $\delta(q,b) = q', b', \leftarrow$ et $a \neq B$
5.$*$	qa	$q'a'$	si $\delta(q,a) = q', a', \downarrow$ et $a \neq B$
6.$*$	$q\star$	$aq'\star$	si $\delta(q,B) = q', a', \rightarrow$
7.$*$	$bq\star$	$q'ba'\star$	si $\delta(q,B) = q', a', \leftarrow$
8.$*$	$q\star$	$q'a\star$	si $\delta(q,B) = q', a, \downarrow$
9.	$q_e\$\star\ \rhd$	$\rhd$	
10.	$\star$	$\star$	
11.$*$	$bq_ea\star$	$q_eb\star$	si $a, b \in \Sigma$, $a, b \neq B$
12.$*$	$bq_e\star$	$q_eb\star$	si $b \in \Sigma, b \neq B$

On utilise la notation $i.*$ pour un ensemble fini d'indices, correspondant aux instances de la condition. Par exemple 2.$*$ est une suite d'indices $2.a$, pour $a \in \Sigma$.

Montrons que PCP modifié a une solution ssi M s'arrête :

Si M s'arrête Soit

$$\gamma_0 = q_0\$ \vdash \cdots \vdash \gamma_i = w_i q_i w_i' \vdash \cdots \vdash \gamma_n = q_e\$$$

la séquence de configurations de M sur la donnée ϵ. On montre, par récurrence sur m, qu'il existe une séquence d'indices $i_1, \ldots, i_{k_m}$ telle que $i_1 = 1$ et

$$u_{i_1} \cdots u_{i_{k_m}} =\lhd \gamma_0 \star \cdots \gamma_{m+1} \star \qquad v_{i_1} \cdots v_{i_{k_m}} =\lhd \cdots \gamma_m \star$$

Si $m = 0$, $k_1 = 1$ et $u_{i_1} =\lhd \gamma_0 \star$, $v_{i_1} =\lhd$.

Pour la récurrence, soit $\gamma_{m+1} = w_{m+1} q_{m+1} w'_{m+1}$.

Si $w_{m+1}, w'_{m+1} \neq \epsilon$ et $q \neq q_e$, $\gamma_{m+1} = x_{m+1} a_{m+1} q_{m+1} b_{m+1} y_{m+1}$ et $\gamma_{m+2} = x_{m+1} u y_{m+1}$ avec $a_{m+1} q_{m+1} b_{m+1} \vdash_M u$.

On construit alors une séquence d'indices $k_m + 1, \ldots, k_{m+1}$ comme suit :

— $2.\alpha_1, \ldots, 2.\alpha_p$ si $x_{m+1} = \alpha_1 \cdots \alpha_p$

— l'indice qui correspond à la transition $a_{m+1} q_{m+1} b_{m+1} \vdash_M u$

— $2.\beta_1, \ldots, 2.\beta_r$ si $y_{m+1} = \beta_1 \cdots \beta_r$

— 10

Par construction des séquences u_i, v_i, on obtient bien

$$u_{i_1} \cdots u_{i_{k_{m+1}}} =\lhd \gamma_0 \star \cdots \gamma_{m+2} \star \qquad v_{i_1} \cdots v_{i_{k_{m+1}}} =\lhd \cdots \gamma_{m+1} \star$$

Cette construction s'étend au cas où $w_{m+1} = \epsilon$: il suffit de remplacer a_{m+1} par ϵ.

Dans le cas où $w'_{m+1} = \epsilon$ (et $q_{m+1} \neq q_e$), on procède de même, en remplaçant l'indice $3.z, 4.z, 5.z$ par un indice $6.z, 7.z, 8.z$.

Si maintenant $q = q_e$, dans le cas où $w'_{m+1} \neq \epsilon, \$$, on utilise la séquence d'indices 2^*8 et, si $w'_{m+1} = \epsilon$, la séquence d'indices 2^*7. Enfin, si $w'_{m+1} = \$$, il suffit de choisir $i_{k_m+1} = 9$.

Réciproquement, si PCP modifié a une solution Alors montrons que M s'arrête sur w.

Soit $u_{i_1} \cdots u_{i_m} = v_{i_1} \cdots v_{i_m}$. On montre, par récurrence sur $k \geq 2$ qu'il existe un p, et des mots t, t' tels que $u_{i_1} \cdots u_{i_k} =\lhd \gamma_0 \star \cdots \gamma_p \star t$, $v_{i_1} \cdots v_{i_k} =\lhd \gamma_0 \star \cdots \gamma_{p-1} \star t'$ et

— $\gamma_0 \vdash_M \cdots \vdash_M \gamma_p$

— t' est un préfixe de γ_p et ou bien $t = t'$ ou bien $t' \vdash_M t$

Si $k = 2$, $u_{i_1} = u_1 =\lhd q_0\$\star$ et $v_{i_1} = v_1 =\lhd$. Comme $i_1, \ldots, i_m$ est une solution de PCP modifié, u_{i_1} est un préfixe de $v_{i_1} \cdots v_{i_m}$ et donc v_{i_2} a pour première lettre q_0.

Par définition des v_i, si la première lettre de v_{i_2} est dans Q, v_{i_2} est de longueur au moins 2. Donc $q_0\$$ est un préfixe de v_{i_2}. Ceci n'est possible que si $i_2 = 3.*$ ou $i_2 = 5.*$. Dans les deux cas $v_{i_2} = q_0\$$ et $u_{i_2} = \$q_1$ si $\delta(q_0, \$) = q_1, \$, \rightarrow$ et $u_{i_2} = q_1\$$ si $\delta(q_0, \$) = q_1, \$, \downarrow$.

Dans le premier cas, $u_{i_1} \cdot u_{i_2} =\lhd q_0\$ \star \$q_1$ et $v_{i_1} \cdot v_{i_2} =\lhd q_0\$$.

Dans le deuxième cas, $u_{i_1} \cdot u_{i_2} =\lhd q_0\$ \star q_1\$$ et $v_{i_1} \cdot v_{i_2} =\lhd q_0\$$.

Dans les deux cas, $u_{i_1} \cdot u_{i_2} =\lhd \gamma_0 \star t$, $v_{i_1} \cdot v_{i_2} =\lhd t'$ avec $t' \vdash_M t$.

Si $k \geq 2$ **et l'invariant est satisfait pour** k , comme $v_{i_1} \cdots v_{i_k} \cdot v_{i_{k+1}}$ est un préfixe de $u_{i_1} \cdots u_{i_m}$, si α est la première lettre de $v_{i_{k+1}}$, $t'\alpha$ est un préfixe de $\gamma_p\star$.

**to be completed **

ou bien s_p est une configuration finale de la machine, ou bien
— t' est un préfixe de s_p et t est un préfixe de s_{p+1}
— Si t' ne contient pas de symbole d'état et ne se termine pas par $\star$, alors $t' = t$, sinon $t' \vdash_M t$.

Pour $k = 1$, $u_{i_1} = u_1 =\lhd\, q_0\$w\star$ et $v_{i_1} = v_1 =\lhd$. On a bien, pour $p = 1$, $u_{i_1} =\lhd\, s_1\star$, $t = t' = \epsilon$.

Si la propriété est vraie pour k, considérons la paire suivante $u_{i_{k+1}}, v_{i_{k+1}}$. Si s_p était déjà une configuration finale, il n'y a rien à faire. Sinon, $s_p = t' \cdot t'_p$ et $s_{p+1} = t \cdot t_p$. De plus, comme on a une solution de PCP modifié, $v_{i_{k+1}}$ est un préfixe de $t'_p \star tu_{i_{k+1}}$. Considérons successivement tous les $v_{i_{k+1}}$ possibles :
— 1. est impossible car $\lhd$ n'apparait pas dans $t'_p \star tu_{i_{k+1}}$
— 2. Dans ce cas, $t'_p = v_{i_{k+1}} t''_p$ et on a la propriété voulue
— 3, 4,5 : $t'_p = v_{i_{k+1}} t''_p$, t' ne contient pas de symbole d'état, et donc $t = t'$ et $t' \cdot v_{i_{k+1}} \vdash_M t \cdot u_{i_{k+1}}$
— 6,7,8 : comme t'_p ne contient pas $\star$, $t'_p\star = v_{i_{k+1}}$. Donc $s_p\star = t' \cdot v_{i_{k+1}}$. De plus, $t = t'$ et $t \cdot u_{i_{k+1}} = s_{p+1}\star$. Il suffit alors de choisir des mots vides pour les nouveaux t, t'
— 9 : t' doit être vide puisque $v_{i_{k+1}}$ est un préfixe de $t'_p \star tu_{i_{k+1}}$.De plus, on doit avoir $t'_p = \$q_e$, ce qui n'est pas possible. Ce cas n'a pas lieu
— 10. t'_p est vide et il suffit de choisir des mots vides pour les nouveaux t, t'
— 11. Impossible pour des raisons identiques à 9.
— 12,13 : l'état final a été atteint.

Maintenant, reste à montrer que PCP lui-même est indécidable. Pour cela on réduit PCP modifié à PCP comme suit, en supposant (sans perte de généralité) qu'il n'y a pas de paire (ϵ, ϵ).

Si $(u_1, \ldots, u_n)$, $(v_1, \ldots, v_n)$ est une instance de PCP modifié, on considère un alphabet augmenté des lettres $\bullet, \rhd, \lhd$ et l'instance de PCP : $(\lhd\, \overline{u_1}, \overline{u_1}, \ldots, \overline{u_n}, \bullet \rhd)$, $(\lhd \bullet\widetilde{v_1}, \widetilde{v_1}, \ldots, \widetilde{v_n}, \rhd)$ où $\overline{\epsilon} = \widetilde{\epsilon} = \epsilon$ et $\overline{a \cdot w} = \bullet a\overline{w}$ et $\widetilde{a \cdot w} = a \bullet \widetilde{w}$.

Si PCP modifié a une solution $u_{i_1} \cdots u_{i_m} = v_{i_1} \cdots v_{i_m}$, alors $\lhd\, \overline{u_{i_1}} \cdots \overline{u_{i_m}} \bullet \rhd = \lhd \bullet\widetilde{v_{i_1}} \cdots \widetilde{v_{i_m}} \rhd$ est une solution de PCP.

Réciproquement, si PCP admet une solution, notons $(u'_0, \ldots, u'_n, u'_{n+1})$ et $(v'_0, \ldots, v'_{n+1})$ l'instance du problème : il existe une suite d'indices telle que $u'_{i_1} \cdots u'_{i_m} = v'_{i_1} \cdots v'_{i_m}$. Notons que, pour tout i, $u'_i = \epsilon$ ou bien u'_i commence par $\bullet$, ou bien $i = 1$. Si $u'_{i_1} = \epsilon$, alors soit k le plus petit indice tel que $u'_{i_k} \neq \epsilon$. Par hypothèse et par construction, $v'_{i_1} \neq \epsilon$ et sa première lettre est $a \notin \{\bullet, \lhd\}$. À l'inverse, la première lettre de u'_{i_k} est dans $\{\bullet, \lhd\}$. Ce qui est absurde. Il en résulte que $u'_{i_1} \neq \epsilon$. Dans ce cas, la première lettre de u'_{i_1} est dans $\{\bullet, \lhd\}$ et donc aussi la première lettre du premier v_{i_k} non vide. Ce n'est possible que si $i_1 = 1$ et cette première lettre est $\lhd$.

Soit maintenant ϕ le morphisme défini sur $(\Sigma \cup \{\bullet, \lhd, \rhd\})^*$ par $\phi(\bullet) = \phi(\lhd) = \phi(\rhd) = \epsilon$ et $\phi(a) = a$ sinon. On montre, par récurrence sur k que $\phi(u'_{i_1} \cdots u'_{i_k}) = u_1 \cdot u_{i_2} \cdots u_{i_k}$ et $\phi(v_{i_1} \cdots v_{i_k}) = v_1 \cdot v_{i_2} \cdots v_{i_k}$, si $1 \leq k < m$. Il en résulte que, si PCP a une solution, alors PCP modifié aussi, en prenant l'image par ϕ.

Exercice 173 (6)
Si $\mathcal{E}$ est un ensemble fini de matrices carrées, le *semi-groupe engendré par* $\mathcal{E}$ est le plus petit ensemble $\mathcal{S}(\mathcal{E})$ qui contient $\mathcal{E}$ et clos par produit : si $M, N \in \mathcal{S}(\mathcal{E})$ alors leur produit MN est dans $\mathcal{S}(\mathcal{E})$.

On veut montrer que le problème suivant est indécidable (*problème de la mortalité de* $\mathcal{E}$) :

Donnée : un ensemble fini de matrices $\mathcal{E}$, à coefficients entiers.

Question : est ce que la matrice nulle est dans le semi-groupe engendré par $\mathcal{E}$?

.

1. Montrer que le problème suivant est indécidable :

   Donnée : un ensemble fini $\mathcal{E}$ de matrices 3×3 à coefficients entiers.

   Question : Existe-t-il une matrice dans $\mathcal{S}(\mathcal{E})$ de la forme $\begin{pmatrix} \alpha & 0 & 0 \\ \beta & 1 & \beta \\ 0 & 0 & \alpha \end{pmatrix}$?

   (**Ind :** On pourra utiliser PCP)

2. Montrer que le problème suivant est indécidable :

   Donnée : un ensembe fini de matrices 3×3,$\mathcal{E}$

   Question : existe-t-il une matrice dans $\mathcal{S}(\mathcal{E})$ dont le coin supérieur gauche est nul ? (i.e. de la forme $\begin{pmatrix} 0 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$)

3. En déduire que le problème de la mortalité d'un ensemble fini de matrices 3×3 à coefficients entiers est indécidable

Exercice 174 (6)
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 ?