

Décidabilité de problèmes de grammaires algébriques

Motivation : Problèmes très simples à formaliser dans les langages formels montrant une différence assez fondamentale entre les langages rationnels et algébriques.

On pose Σ un alphabet.

- Entrée : G grammaire
Sortie : $\varepsilon \in L_G(S)$?
est décidable
- Entrée : G grammaire, $w \in \Sigma^+$
Sortie : $w \in L_G(S)$?
est décidable (CYK).
- Entrée : G, G' grammaires
Sortie : $L_G(S) \cap L_{G'}(S) = \emptyset$?
est indécidable.

On montre l'indécidabilité par réduction au problème de Correspondance de Post (PCP).

PCP:

Entrées: $(u_1, v_1), \dots, (u_n, v_n)$ sur Σ

Sortie: $\exists i_1, \dots, i_k \in ([1; n])$, $u_{i_1} \dots u_{i_k} = v_{i_1} \dots v_{i_k}$

On admet ici que PCP est indécidable.

Preuve:

Soit $(u_1, v_1), \dots, (u_n, v_n)$ une instance de PCP sur Σ .

Soit $A = \{a_1, \dots, a_n\}$ un alphabet de n lettres avec $A \cap \Sigma = \emptyset$.

On définit sur $A + \Sigma$:

$$L_u = \{u_{i_1} \dots u_{i_k} a_{i_k} \dots a_{i_1}, n \in \mathbb{N} \text{ et } i_1, \dots, i_k \in ([1; n])\}$$

$$L'_u = \left\{ w a_{i_k} \dots a_{i_1}, n \in \mathbb{N}, i_1, \dots, i_k \in ([1; n]), w \in \Sigma^*, w \neq u_{i_1} \dots u_{i_k} \right\}$$

Lemme: L_u et L'_u sont algébriques.

On a: l'instance de PCP admet une solution si et seulement si $L_u \cap L_v \neq \emptyset$.

On a une réduction de PCP à ce pb donc ce pb est indécidable.

Preuve du lemme:

L_u est engendré par $S \rightarrow \sum_{i=1}^n u_i S a_i + \epsilon$.

$$L_u' = \Sigma^* A^* \setminus L_u$$

↳ Ne donne pas une preuve mais une intuition

L_u' est engendré par:

$$S \rightarrow \sum_{i=1}^n u_i S a_i + \sum_{\substack{i=1 \\ |u|=|u_i| \\ u \neq u_i}}^n u R a_i + \sum_{\substack{i=1 \\ |u| < |u_i|}}^m u T a_i + \sum_{\substack{i=1 \\ b \in \Sigma}}^n u_i b V a_i$$

$$R \rightarrow \sum_{i=1}^n R a_i + \sum_{b \in \Sigma} b R + \epsilon$$

$$T \rightarrow \sum_{i=1}^n T a_i + \epsilon$$

$$V \rightarrow \sum_{b \in \Sigma} b V + \epsilon$$

Idee: $S \rightarrow$ mot entier

↳ on continue de générer jusqu'à une "erreur"

$R \rightarrow$ on a fait une erreur non rattrapable
donc on peut générer comme on veut

$T \rightarrow$ il manque des lettres à gauche donc on ne
génère plus à droite pour se "compléter"

$V \rightarrow$ il y a trop de lettres à gauche donc on peut en rajouter.

Entrée : G grammaire

Sortie : $L_G(S) = \Sigma^*$?

Preuve : On remarque : $(A + \Sigma)^* \setminus L_u = L_u' \cup ((A + \Sigma)^* \setminus \Sigma^* A^*)$

car $L_u \subseteq \Sigma^* A^*$ et $\Sigma^* A^* \setminus L_u = L_u'$.
donc L_u et L_v sont algébriques.

On réduit le problème précédent à celui-ci :

$L_u \cap L_v$ est vide $\Leftrightarrow (A + \Sigma)^* \setminus L_u \cup (A + \Sigma)^* \setminus L_v = (A + \Sigma)^*$.

Comme le pb précédent est décidable, alors celui-ci aussi.