Problèmes indécidables sur les grammaires algébriques

Manon Ruffini

Théorème 1

Soient G, G' des grammaires algébriques d'axiomes respectifs S et S'. Alors, les problèmes suivants sont indécidables :

- *Est-ce que $L_G(S) \cap L_{G'}(S') = \emptyset$?*
- *Est-ce que $L_G(S) = L_{G'}(S')$*
- *Est-ce que $L_G(S) = \mathcal{A}^*$?*
- *Est-ce que G est ambigüe ?*

Proposition 1 (*Rappel*)

Le problème de correspondance de Post (PCP) est indécidable.

Entrée : Une suite $(u_1, v_1), \ldots, (u_m, v_m)$ de paires de mots sur un alphabet Σ

Sortie : Oui, ssi il existe une suite d'indices $i_1, \ldots, i_n$ de $[|1, n|]$ telle que $u_{i_1} \ldots u_{i_n} = v_{i_1} \ldots v_{i_n}$

Étape 1

Soit $(u_1, v_1) \ldots (u_m, v_m)$ une instance de PCP. Soit $A = \{a_1, \ldots, a_m\}$ un alphabet tel que $A \cap \Sigma = \emptyset$. On définit les langages suivants :

$$L_u = \{u_{i_1} \ldots u_{i_n} a_{i_n} \ldots a_{i_1} | n \geqslant 0 \text{ et } 1 \leqslant i_k \leqslant m\}$$

$$L'_u = \{w a_{i_n} \ldots a_{i_1} | n \geqslant 0, w \in \Sigma^* \text{ et } w \neq u_{i_1} \ldots u_{i_n}\}$$

Ces langages sont algébriques et $(A + \Sigma)^* \setminus L_u = L'_u \cup ((A + \Sigma)^* \setminus \Sigma^* A^*)$

Le langage L_u est engendré par la grammaire :

$$S \to \sum_{i=1}^{m} u_i S a_i + \varepsilon$$

Le langage L'_u est engendré par la grammaire suivante :

$$\begin{aligned}
S &\to \sum_{i=1}^{m} u_i S a_i + \sum_{\substack{1 \leqslant i \leqslant m \\ |u| = |u_i|, u \neq u_i}} u R a_i + \sum_{\substack{1 \leqslant i \leqslant m \\ |u| < |u_i|}} u T a_i + \sum_{\substack{1 \leqslant i \leqslant m \\ b \in \Sigma}} u_i b V a_i \\
R &\to \sum_{i=1}^{m} R a_i + \sum_{b \in Sigma} b R + \varepsilon \\
T &\to \sum_{i=1}^{m} T a_i + \varepsilon \\
V &\to \sum_{b \in \Sigma} b V + \varepsilon
\end{aligned}$$

L'égalité découle du fait que $L_u \subset \Sigma^* A^*$ et $\Sigma^* A^* \setminus L_u = L'_u$.

Étape 2

Soit $(u_1, v_1), \ldots, (u_m, v_m)$ une instance de PCP. Les langages L_u et L_v définis comme précé-

demment sont algébriques et calculables par machine de Turing. De plus, l'instance de PCP est positive ssi $L_u \cap L_v)\emptyset$.
On a réduit PCP à (1). Comme PCP est indécidable, (1) est indécidable.

Étape 3

$$L_u \cap L_v = \emptyset \text{ ssi } ((A+\Sigma)^* \setminus L_u) \cup ((A+\Sigma)^* \setminus L_v) = (A+\Sigma)^*$$

Or $(A+\Sigma)^* \setminus L_u = L'_u \cup ((A+\Sigma)^* \setminus \Sigma^* A^*)$ *est algébrique car* L'_u *et* $(A+\Sigma)^* \setminus \Sigma^* A^*$ *le sont. Donc,* $((A+\Sigma)^* \setminus L_u) \cup ((A+\Sigma)^* \setminus L_v)$ *est algébrique et on a réduit PCP à (3). Ainsi, (3) est indécidable.*

Étape 4

Comme (3) est un sous-problème de (2) et que (3) est indécidable, (2) est indécidable.

Étape 5

$(L_u + L_v) \setminus \{\varepsilon\} = L_G(S)$ *où la grammaire* G *est donnée par :*

$$\begin{aligned}
&S \to S_1 + S_2 \\
&S_1 \to \sum_{i=1}^{m} u_i S_1 a_i + \sum_{i=1}^{m} u_i a_i \\
&S_2 \to \sum_{i=1}^{m} v_i S_2 a_i + \sum_{i=1}^{m} v_i a_i
\end{aligned}$$

La grammaire G *est ambigüe ssi* $L_u \cap L_v \neq \emptyset$. *On a réduit (1) à (4). Donc (4) est indécidable.*

Références

[1] Olivier Carton, *Langages formels*. Vuibert, 2014.