TD 8 – Indécidabilité, le retour

1 Décidabilité et grammaires algébriques

Soit Σ un alphabet. Étant donné $m \in \mathbb{N}_{>0}$ et une suite $u = (u_1, u_2, \ldots, u_m)$ de m mots sur Σ, on introduit un nouvel alphabet $A = \{a_1, a_2, \ldots, a_m\}$ disjoint de Σ et on définit

$$L_u = \{u_{i_1} u_{i_2} \cdots u_{i_n} a_{i_n} \cdots a_{i_2} a_{i_1} \mid n \in \mathbb{N}_{>0} \wedge \forall j \in [\![1, n]\!], i_j \in [\![1, m]\!]\} \ ,$$
$$L'_u = \{w a_{i_n} \cdots a_{i_2} a_{i_1} \mid n \in \mathbb{N} \wedge \forall j \in [\![1, n]\!], i_j \in [\![1, m]\!] \wedge w \in \Sigma^* \wedge w \neq u_{i_1} u_{i_2} \cdots u_{i_n}\} \cup \{\varepsilon\} \ .$$

Exercice 1. Montrer que les langages ainsi définis sont algébriques (« context-free »).

Solution 1. La grammaire $\mathcal{G} = (\{S, T\}, \Sigma \cup A, R, S)$ où R contient les règles

$$\begin{aligned} S &\to u_{i_1} T a_{i_1} \mid u_{i_2} T a_{i_2} \mid \cdots \mid u_{i_m} T a_{i_m} \\ T &\to u_{i_1} T a_{i_1} \mid u_{i_2} T a_{i_2} \mid \cdots \mid u_{i_m} T a_{i_m} \mid \varepsilon \end{aligned}$$

génère L_u.

La grammaire $\mathcal{G}' = (\{S', S'', T', U', V'\}, \Sigma \cup A, R, S')$ où R contient les règles

$$\begin{array}{ll} S' \to S'' \mid \varepsilon & \\ S'' \to u_i S'' a_i & \forall i \in [\![1, m]\!] \\ S'' \to bT' & \forall b \in \Sigma \\ S'' \to vU'a_i & \forall i \in [\![1, m]\!], v \in \Sigma^*, |v| < |u_i| \\ S'' \to vV'a_i & \forall i \in [\![1, m]\!], v \in \Sigma^{|u_i|} \setminus \{u_i\} \\ T' \to bT' & \forall b \in \Sigma \\ T' \to \varepsilon & \\ U' \to U'a_i & \forall i \in [\![1, m]\!] \\ U' \to \varepsilon & \\ V' \to V'a_i & \forall i \in [\![1, m]\!] \\ V' \to bV' & \forall b \in \Sigma \\ V' \to \varepsilon & \end{array}$$

génère quant à elle L'_u.

Exercice 2. Montrer, en utilisant l'indécidabilité de PCP, que les langages suivants sont indécidables.

1. $\mathrm{DISJ}_{\mathrm{CFG}} = \{\langle \mathcal{G}_1, \mathcal{G}_2 \rangle \mid \mathcal{G}_1$ et $\mathcal{G}_2$ sont des grammaires algébriques et $\mathcal{L}(\mathcal{G}_1) \cap \mathcal{L}(\mathcal{G}_2) = \emptyset\}$.
2. $\mathrm{EQ}_{\mathrm{CFG}} = \{\langle \mathcal{G}_1, \mathcal{G}_2 \rangle \mid \mathcal{G}_1$ et $\mathcal{G}_2$ sont des grammaires algébriques et $\mathcal{L}(\mathcal{G}_1) = \mathcal{L}(\mathcal{G}_2)\}$.
3. $\mathrm{ALL}_{\mathrm{CFG}} = \{\langle \mathcal{G} \rangle \mid \mathcal{G}$ est une grammaire algébrique sur un alphabet Σ et $\mathcal{L}(\mathcal{G}) = \Sigma^*\}$.
4. $\mathrm{REG}_{\mathrm{CFG}} = \{\langle \mathcal{G} \rangle \mid \mathcal{G}$ est une grammaire algébrique et $\mathcal{L}(\mathcal{G})$ est rationnel$\}$.
5. $\mathrm{IDEM}_{\mathrm{CFG}} = \{\langle \mathcal{G} \rangle \mid \mathcal{G}$ est une grammaire algébrique et $\mathcal{L}(\mathcal{G}) = \mathcal{L}(\mathcal{G})\mathcal{L}(\mathcal{G})\}$.

Solution 2. Soit Σ un alphabet. Considérons $(u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m)$ avec $m \in \mathbb{N}_{>0}$ une suite de dominos sur Σ.

On pose $u = (u_1, u_2, \ldots, u_m)$ et $v = (v_1, v_2, \ldots, v_m)$ et on considère les langages L_u, L'_u, L_v, L'_v tels que définis plus haut sur l'alphabet $\Sigma \cup A$ où $A = \{a_1, a_2, \ldots, a_m\}$ est disjoint de Σ.

Observons que $(\Sigma \cup A)^* \setminus L_u = L'_u \cup ((\Sigma \cup A)^* \setminus \Sigma^* A^*)$ et $(\Sigma \cup A)^* \setminus L_v = L'_v \cup ((\Sigma \cup A)^* \setminus \Sigma^* A^*)$.

Dans la suite, on donne l'idée de la réduction que l'on applique étant donné le mot $\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle$, réduction qui est à chaque fois calculable. Pour le cas des mots qui ne correspondent pas à un encodage valide d'une instance pour PCP, on associe simplement un mot fixé qui convient.

1. Soit $\mathcal{G}_u$ et $\mathcal{G}_v$ des grammaires algébriques engendrant, respectivement, L_u et L_v. On observe que
$$\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle \in \text{PCP} \Leftrightarrow L_u \cap L_v \neq \emptyset \Leftrightarrow \langle \mathcal{G}_u, \mathcal{G}_v \rangle \notin \text{DISJ}_{\text{CFG}} \ .$$
2. Soit $\mathcal{G}'_u$ et $\mathcal{G}_\cup$ des grammaires algébriques engendrant, respectivement, L'_u et $L'_u \cup L_v$. On observe que
$$\begin{aligned}\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle \in \text{PCP} &\Leftrightarrow L_u \cap L_v \neq \emptyset \\ &\Leftrightarrow L_v \nsubseteq L'_u \\ &\Leftrightarrow L'_u \neq L'_u \cup L_v \\ &\Leftrightarrow \langle \mathcal{G}'_u, \mathcal{G}_\cup \rangle \notin \text{EQ}_{\text{CFG}} \ .\end{aligned}$$
3. Soit $\mathcal{G}''_\cup$ une grammaire algébrique engendrant $(\Sigma \cup A)^* \setminus L_u \cup (\Sigma \cup A)^* \setminus L_v$. On observe que
$$\begin{aligned}\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle \in \text{PCP} &\Leftrightarrow L_u \cap L_v \neq \emptyset \\ &\Leftrightarrow (\Sigma \cup A)^* \setminus L_u \cup (\Sigma \cup A)^* \setminus L_v \neq (\Sigma \cup A)^* \\ &\Leftrightarrow \langle \mathcal{G}''_\cup \rangle \notin \text{ALL}_{\text{CFG}} \ .\end{aligned}$$
4. On suppose d'abord qu'il n'existe pas $i \in [\![1, m]\!]$ tel que $(u_i, v_i) = (\varepsilon, \varepsilon)$.
Soit $\mathcal{G}'_\cup$ une grammaire algébrique engendrant $L'_u \cup L'_v$.
Si $\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle \notin \text{PCP}$, alors $L_u \cap L_v = \emptyset$ et il est direct de voir que $L'_u \cup L'_v = \Sigma^* A^*$, qui est un langage rationnel.
Autrement, si $\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle \in \text{PCP}$, alors il existe $n \in \mathbb{N}_{>0}$ et $i_1, i_2, \ldots, i_n \in [\![1, m]\!]$ tels que $u_{i_1} u_{i_2} \cdots u_{i_n} a_{i_n} \cdots a_{i_2} a_{i_1} = v_{i_1} v_{i_2} \cdots v_{i_n} a_{i_n} \cdots a_{i_2} a_{i_1} \in L_u \cap L_v$. Notons $w = u_{i_1} u_{i_2} \cdots u_{i_n} = v_{i_1} v_{i_2} \cdots v_{i_n}$ et $x = a_{i_n} \cdots a_{i_2} a_{i_1}$. Par hypothèse, on a nécessairement $w \neq \varepsilon$, d'où il s'ensuit que pour tous $j, k \in \mathbb{N}_{>0}, j \neq k$, $(w^j)^{-1}(L'_u \cup L'_v) \neq (w^k)^{-1}(L'_u \cup L'_v)$ puisque $w^j x^j \in L_u \cap L_v \subseteq (\Sigma \cup A)^* \setminus (L'_u \cup L'_v)$ alors que $w^k x^j \in L'_u \cup L'_v$. On en déduit donc que $L'_u \cup L'_v$ a une infinité de quotients à gauche, d'où $L'_u \cup L'_v$ n'est pas rationnel.
On en conclut que
$$\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle \in \text{PCP} \Leftrightarrow L'_u \cup L'_v \notin \mathcal{R}\mathsf{eg} \Leftrightarrow \langle \mathcal{G}'_\cup \rangle \notin \text{REG}_{\text{CFG}} \ .$$
Pour le cas où il existe $i \in [\![1, m]\!]$ tel que $(u_i, v_i) = (\varepsilon, \varepsilon)$, on associe simplement un mot fixé qui convient.
5. On suppose d'abord qu'il n'existe pas $i \in [\![1, m]\!]$ tel que $(u_i, v_i) = (\varepsilon, \varepsilon)$.
Soit $\mathcal{G}''_\cup$ une grammaire algébrique engendrant $(\Sigma \cup A)^* \setminus L_u \cup (\Sigma \cup A)^* \setminus L_v = L''_\cup$.
Si $\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle \notin \text{PCP}$, alors $(\Sigma \cup A)^* \setminus L_u \cup (\Sigma \cup A)^* \setminus L_v = (\Sigma \cup A)^*$ qui vérifie bien évidemment $(\Sigma \cup A)^* = (\Sigma \cup A)^* (\Sigma \cup A)^*$.
Autrement, si $\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle \in \text{PCP}$, alors il existe $n \in \mathbb{N}_{>0}$ et $i_1, i_2, \ldots, i_n \in [\![1, m]\!]$ tels que $u_{i_1} u_{i_2} \cdots u_{i_n} a_{i_n} \cdots a_{i_2} a_{i_1} = v_{i_1} v_{i_2} \cdots v_{i_n} a_{i_n} \cdots a_{i_2} a_{i_1} \in L_u \cap L_v$. Notons $w = u_{i_1} u_{i_2} \cdots u_{i_n} = v_{i_1} v_{i_2} \cdots v_{i_n}$ et $x = a_{i_n} \cdots a_{i_2} a_{i_1}$. Par hypothèse, on a nécessairement $w \neq \varepsilon$, d'où il s'ensuit que $w, x \in (\Sigma \cup A)^* \setminus L_u \cup (\Sigma \cup A)^* \setminus L_v$. Nous avons donc que $wx \in L''_\cup L''_\cup$, alors que $wx \notin L''_\cup$, d'où $L''_\cup \neq L''_\cup L''_\cup$.

On en conclut que

$$\langle (u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m) \rangle \in \mathrm{PCP} \Leftrightarrow L''_{\cup} \neq L''_{\cup} L''_{\cup} \Leftrightarrow \langle \mathcal{G}''_{\cup} \rangle \notin \mathrm{IDEM}_{\mathrm{CFG}} \ .$$

Pour le cas où il existe $i \in [\![1, m]\!]$ tel que $(u_i, v_i) = (\varepsilon, \varepsilon)$, on associe simplement un mot fixé qui convient.

Exercice 3. Montrer que

1. $\{\langle \mathcal{G}, \mathcal{A} \rangle \mid \mathcal{G}$ est une grammaire algébrique, $\mathcal{A}$ est un AFD et $\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{G})\}$ est indécidable ;
2. $\{\langle \mathcal{G}, \mathcal{A} \rangle \mid \mathcal{G}$ est une grammaire algébrique, $\mathcal{A}$ est un AFD et $\mathcal{L}(\mathcal{G}) \subseteq \mathcal{L}(\mathcal{A})\}$ est décidable.

Solution 3.

1. On réduit $\mathrm{ALL}_{\mathrm{CFG}}$ au langage considéré, en observant que, pour toute grammaire algébrique $\mathcal{G}$ sur un alphabet Σ,
$$\mathcal{L}(\mathcal{G}) = \Sigma^* \Leftrightarrow \mathcal{L}(\mathcal{A}_{\mathrm{ALL}}) \subseteq \mathcal{L}(G)$$
où $\mathcal{A}_{\mathrm{ALL}}$ est un AFD reconnaissant Σ^*.
2. Étant donné une grammaire algébrique $\mathcal{G}$ et un AFD $\mathcal{A}$ sur un alphabet Σ, on observe que
$$\mathcal{L}(\mathcal{G}) \subseteq \mathcal{L}(\mathcal{A}) \Leftrightarrow \mathcal{L}(\mathcal{G}) \cap \Sigma^* \setminus \mathcal{L}(\mathcal{A}) = \emptyset \ .$$
De plus, puisque $\mathcal{L}(\mathcal{A})$ est rationnel, $\Sigma^* \setminus \mathcal{L}(\mathcal{A})$ l'est aussi et donc $\mathcal{L}(\mathcal{G}) \cap \Sigma^* \setminus \mathcal{L}(\mathcal{A})$ est algébrique. Puisque le langage $\{\langle \mathcal{G} \rangle \mid \mathcal{G}$ est une grammaire algébrique et $\mathcal{L}(\mathcal{G}) = \emptyset\}$ est décidable, on peut conclure.

2 Séparabilité

Exercice 4. Soit Σ un alphabet. Soit A et B deux langages disjoints sur Σ. On dit qu'un langage C sur Σ *sépare A et B* lorsque $A \subseteq C$ et $B \subseteq \bar{C}$.

1. Si deux langages A et $\bar{A}$ sur Σ ne sont séparés par aucun langage décidable, que peut-on en déduire ?
2. Montrer que deux langages A et B sur Σ sont séparés par un langage décidable s'ils sont co-reconnaissables et disjoints.
3. Montrer qu'il existe des langages A et B sur Σ qui sont reconnaissables, mais ne sont séparés par aucun langage décidable.

Solution 4.

1. On peut en déduire que ni A, ni $\bar{A}$ ne sont décidables.
2. Étant donné une MT $\mathcal{M}_{\bar{A}}$ reconnaissant $\bar{A}$ et une MT $\mathcal{M}_{\bar{B}}$ reconnaissant $\bar{B}$, on construit une MT $\mathcal{M}$ qui, étant donné un mot $w \in \{0,1\}^*$ en entrée, exécute $\mathcal{M}_{\bar{A}}(w)$ et $\mathcal{M}_{\bar{B}}(w)$ en parallèle, en acceptant si $\mathcal{M}_{\bar{A}}$ accepte en premier, et en rejetant si $\mathcal{M}_{\bar{B}}$ accepte en premier.
Quel que soit le mot $w \in \{0,1\}^*$ en entrée, $\mathcal{M}$ s'arrêtera puisqu'il y a toujours au moins une des deux MT $\mathcal{M}_{\bar{A}}$ et $\mathcal{M}_{\bar{B}}$ qui acceptera avec w en entrée, de par le fait que A et B sont disjoints.
Le langage $\mathcal{L}(\mathcal{M})$ est donc décidable par construction et vérifie que pour tout $w \in \{0,1\}^*$,
— si $w \in A$, alors $w \in A \cap \bar{B}$ et donc $w \notin \mathcal{L}(\mathcal{M})$;
— si $w \in B$, alors $w \in \bar{A} \cap B$ et donc $w \in \mathcal{L}(\mathcal{M})$.
Par conséquent, $\mathcal{L}(\mathcal{M})$ sépare A et B.
3. On pose
$$A = \{\langle \mathcal{M} \rangle \mid \mathcal{M} \text{ MT normalisée acceptant sur le mot vide}\}$$
et
$$B = \{\langle \mathcal{M} \rangle \mid \mathcal{M} \text{ MT normalisée rejetant sur le mot vide}\} \ ,$$
puis on observe que si A et B pouvaient être séparés par un langage décidable, il existerait une MT résolvant « Consistent guessing » (voir le polycopié du cours, dans la partie traitant de la décidabilité de théories logiques).