Langages Formels

TD n°2

Emilie Grienenberger
emilie.grienenberger@lsv.fr

31 mars 2021

Exercise 1 : Arden Lemma

Let A, B be two languages.

1. Prove that the language $L = A^*B$ is the smallest solution to the equation :
$$X = (A \cdot X) \cup B$$
2. Prove that if $\varepsilon \notin A$, then it is the only solution.

Exercise 2 : Regular identities

We study identities on regular expressions r, s, t. Here, $r = s$ means $\mathcal{L}(r) = \mathcal{L}(s)$.

1. Prove the following identities :
   (a) $(r + s)t = rt + st$
   (b) $(r^*)^* = r^*$
   (c) $(rs + r)^*r = r(sr + r)^*$
2. Prove or disprove the following identities :
   (a) $(r + s)^* = r^* + s^*$
   (b) $(r^*s^*)^* = (r + s)^*$
   (c) $s(rs + s)^*r = rr^*s(rr^*s)^*$

Exercise 3 : Antimirov's automaton

The goal of this exercise is to build an automaton from a regular expression. We will define a *partial derivative* operation $\partial_a(E)$ which corresponds to $a^{-1}\mathcal{L}(E)$ (via interpretation of expressions). Formally, for every expression E and letter a, we define the *set of expressions* $\partial_a(E)$ as follows :

$$\begin{aligned}
\partial_a(\underline{\emptyset}) &= \emptyset \\
\partial_a(\underline{b}) &= \begin{cases} \emptyset & \text{if } a \neq b \\ \{\underline{\emptyset}^*\} & \text{else} \end{cases} \\
\partial_a(E + E') &= \partial_a(E) \cup \partial_a(E') \\
\partial_a(E^*) &= \partial_a(E) \cdot \{E^*\} \\
\partial_a(E \cdot E') &= \begin{cases} \partial_a(E) \cdot \{E'\} & \text{if } \varepsilon \notin E \\ (\partial_a(E) \cdot \{E'\}) \cup \partial_a(E') & \text{else} \end{cases}
\end{aligned}$$

where concatenation is naturally extended over sets of expressions.
We define $\partial_w(E)$ for a word w inductively with $\partial_\varepsilon(E) = \{E\}$ and $\partial_{wa}(E) = \partial_a(\partial_w(E))$, where $\partial_w(S) = \bigcup_{E\in S} \partial_w(E)$ when S is a set of expressions.
Given a set of regular expressions S, we denote by $\mathcal{L}(S)$ the set $\bigcup_{E\in S} \mathcal{L}(E)$.

1. Give the partial derivatives of $(ab + b)^*ba$ by a and b.
2. Prove that for every $L, L' \subseteq \Sigma^*$ and $a \in \Sigma$,

$$\begin{aligned} a^{-1}(L \cup L') &= (a^{-1}L) \cup (a^{-1}L') \\ a^{-1}L^* &= (a^{-1}L) \cdot L^* \\ a^{-1}(L \cdot L') &= \begin{cases} a^{-1}L \cdot L' & \text{si } \varepsilon \notin L \\ (a^{-1}L \cdot L') \cup (a^{-1}L') & \text{sinon} \end{cases} \end{aligned}$$

3. Show that $\mathcal{L}(\partial_w(E)) = w^{-1}\mathcal{L}(E)$.
4. We define the set of non empty suffixes of a word :

$$\mathrm{Suf}(w) = \{\ v \in \Sigma^+\ :\ \exists u, w = uv\ \}$$

   Show that for every $w \in \Sigma^+$:

$$\begin{aligned} \partial_w(E + E') &= \partial_w(E) \cup \partial_w(E') \\ \partial_w(E \cdot E') &\subseteq (\partial_w(E) \cdot E') \cup \bigcup_{v\in\mathrm{Suf}(w)} \partial_v(E') \\ \partial_w(E^*) &\subseteq \bigcup_{v\in\mathrm{Suf}(w)} \partial_v(E) \cdot E^* \end{aligned}$$

5. Let $\|E\|$ be the number of occurences of letters of Σ in E. Show that the set of partial derivatives different to E has at most $\|E\| + 1$ elements.
6. Conclude, and apply the construction to the expression $(ab + b)^*ba$.

Exercise 4 : Closure by morphism

1. Let h be the morphism $h(a) = 01$ and $h(b) = 0$. Give $h(a(a + b)^*)$.
2. Apply the construction of closure by morphism to this example.
3. Let h' be the morphism $h(0) = ab$, $h(1) = \varepsilon$. Give $h^{-1}(\{\ abab,\ baba\ \})$.
4. Apply the construction of closure by inverse morphism to this example.
5. Let $L = (00\cup1)^*$, $h(a) = 01$ and $h(b) = 10$. What is $h^{-1}(1001)$? $h^{-1}(010110)$? $h^{-1}(L)$? What is $h(h^{-1}(L))$, and is it related to L ? Apply the construction by inverse morphism to this example.

Exercise 5 : Characterizing recognizability

We want to show a converse to the pumping lemma. We say that a language L satisfies P_h if for all $uv_1 \dots v_h w$ avec $|v_i| \geq 1$, there exists $0 \leq j < k \leq h$ such that

$$uv_1 \dots v_h w \in L \Leftrightarrow uv_1 \dots v_j v_{k+1} \dots v_h w \in L.$$

The theorem of Ehrenfeucht, Parikh & Rozenberg states that L is rational iff there exists h such that L satisfies P_h.

1. Show that if L satisfies P_h, then $w^{-1}L$ also does for every word $w \in \Sigma^*$.
2. Let $h \in \mathbb{N}$. We want to show that the number of languages satisfying P_h is finite. We use the following statement of Ramsey's theorem :

   > For every k there is N such that, for every set E of cardinal greater than N and every bipartition $\mathcal{P}$ of $\mathfrak{P}_2(E) = \{ \{e, e'\} \ : \ e, e' \in E, \ e \neq e' \}$, there exists a subset $F \subseteq E$ of cardinal k such that $\mathfrak{P}_2(F)$ is contained in one of the classes of $\mathcal{P}$.

   Let N be the natural number given by Ramsey's theorem for $k = h + 1$. Let L and L' be two languages satisfying P_h and coinciding on words of size smaller than N. Prove that they coincide on words or size $M \geq N$, by induction on M. You may consider, for a word $f = a_1 \dots a_N t$ of size M (with $a_i \in \Sigma$), the following partition of $\mathfrak{P}_2([0; N])$:

   $$X_f = \{ (j, k) \ : \ 0 \leq j < k \leq N, \ a_1 \dots a_j a_{k+1} \dots a_N t \in L \}$$

   $$Y_f = \mathfrak{P}_2([0; N]) \setminus X_f$$

   Conclude.
3. Conclude that if a language L satisfies P_h for some h, then L is regular.

Exercise 6 : A rational slice ? (open exercise)

Let L be a rational language over a finite alphabet Σ. Is $\mathsf{FH}(L) = \{ f \in \Sigma^* \ : \ \exists h \in \Sigma^*. \ |h| = |f|, \ fh \in L \}$ rational ?