Informatique théorique III
Série 5

Prof. Nestmann, 2004

1. Langage non régulier et propriétés de stabilité

Soit $\Sigma = \{a, b\}$. On admet dans cet exercice que $L \triangleq \{a^n b^n \mid n \in \mathbb{N}\}$ n'est pas régulier.

Montrer que les langages suivants ne sont pas réguliers :

1. $L_1 \triangleq \{w \in \Sigma^* \mid |w|_a = |w|_b\}$
2. $L_2 \triangleq \{a^i b^j \mid i, j \in \mathbb{N}, i \neq j\}$
3. $L_3 \triangleq \{b^{2n} a^n \mid n \in \mathbb{N}\}$

On pourra raisonner par l'absurde et exprimer L en fonction du langage à considérer et d'opérations qui conservent la propriété de régularité d'un langage.

2. Propriétés de stabilité

Soit Σ un alphabet.

Soit $u \in \Sigma^*$. On dit que $v \in \Sigma^*$ est un *préfixe propre* de u si v est un préfixe de u (i.e. s'il existe $w \in \Sigma^*$ tel que $u = v \cdot w$) et $v \neq u$.

On a donc l'équivalence suivante :

$$v \text{ est un préfixe propre de } u \Leftrightarrow \exists w \in \Sigma^* \setminus \{\epsilon\} : u = v \cdot w$$

Soit L un langage régulier.

Montrer que $\min(L) = \{u \in L \mid u \text{ ne possède pas de préfixe propre dans } L\}$ est un langage régulier en exprimant $\min(L)$ en fonction de L et d'autres langages réguliers et en utilisant les propriétés de stabilité des langages réguliers.

3. Propriétés de stabilité (artisanal)

Soit Σ un alphabet et L un langage régulier sur Σ. Montrer que les langages suivants sont réguliers :

1. $\mathrm{mirror}(L) = \{u \in \Sigma^* \mid \overline{u} \in L\}$ (où $\overline{u}$ a été défini dans la série 1)
2. $\frac{1}{2}L = \{u \in \Sigma^* \mid \exists v \in \Sigma^* : |u| = |v| \land u \cdot v \in L\}$

Indication : Prendre un AFD M reconnaissant L et construire un AFD, AFN ou AFN$_{\mathbf{e}}$ (défini en fonction de M) reconnaissant $f(L)$.

4. Langages finis et langages réguliers

Un langage est *fini* s'il contient un nombre fini de mots. Il est *infini* dans le cas contraire.

Démontrez le théorème suivant.

Théorème 4.1 *Tous les langages finis sont réguliers.* □

5. Langages non réguliers

Montrez à l'aide du lemme de gonflement la non-régularité des langages suivants.

1. $L = \{w \in \{a, b\}^* \mid |w|_a = |w|_b\}$

2. Un *palindrome* est un mot sur un alphabet Σ que l'on peut lire de manière identique dans un sens ou dans l'autre[1]. En d'autre termes on a :

$$\mathrm{Pal}(w) \Leftrightarrow \forall i \in [1, |w|] : (w)_i = (w)_{|w|+1-i}$$

$L = \{w \in \{a, b\}^* \mid \mathrm{Pal}(w)\}$

3. Soit $\Sigma = \{(,), a\}$, un mot w sur Σ dont les parenthèses sont *équilibrées* satisfait la propriété suivante :

$$\mathrm{Equil}(w) \Leftrightarrow (|w|_(= |w|_)) \wedge (\forall u, v \in \Sigma^* : w = uv \Rightarrow |u|_(\geq |u|_))$$

Par exemple les mots $a(a()(((a)))a)$ et $a()()$ sont équilibrés, mais les mots $(a((()))$ et $))(($ ne le sont pas.

$L = \{w \in \Sigma \mid \mathrm{Equil}(w)\}$ □

6. Langages réguliers infinis

L'algorithme suivant permet de décider la finitude d'un langage régulier L sur un alphabet Σ. L'algorithme *accepte* un langage L si et seulement si il est fini.

1. *Trouver un nombre n satisfaisant la condition du lemme de gonflement $\mathbf{G}(L)$ (le plus petit par exemple).*
2. *Si pour tout $w \in \Sigma^*$ tel que $n \leq |w| \leq 2n - 1$ on a $w \notin L$ alors L est accepté.*
   Sinon L est rejeté.

Démontrer la validité de l'algorithme, c'est à dire que pour tout langage régulier L, L est accepté ssi L est fini. Pour cela :

1. Démontrer que pour tout langage régulier L, le point 1 peut être satisfait.
2. Démontrer la proposition suivante qui est implicitement contenue dans le point 2 de l'algorithme.

   Soit L un langage régulier et n un nombre satisfaisant la condition du lemme de gonflement $\mathbf{G}(L)$. Alors :

$$\exists w \in L : n \leq |w| \leq 2n - 1 \Rightarrow L \text{ infini} \quad \text{(a)}$$
$$\neg\exists w \in L : n \leq |w| \leq 2n - 1 \Rightarrow L \text{ fini} \quad \text{(b)}$$

   N.B. Un langage L est fini ssi il existe une borne finie sur la longueur de ses mots. □

[1] « Ésope reste élu par cette crapule et se repose » est un palindrome modulo les accents et les espaces. Un grand palindrome écrit en français (environ 1500 mots) a été publié par Georges Perec sous le titre 9691 dans Oulipo, *La littérature potentielle*, Gallimard 1973 (seconde édition Folio, 1988).