Langages formels

Devoir 1

à rendre le lundi 19 mars

Soit $\mathcal{F}$ un ensemble fini de symboles avec arités. Soit $\mathcal{X}$ un ensemble de variables. Un système de réécriture est un ensemble fini $\mathcal{S} \subseteq \mathsf{T}(\mathcal{F}, \mathcal{X})^2$ de couples de termes. Étant donné un système de réécriture $\mathcal{S}$, on définit la relation de réduction $\to$ sur $\mathsf{T}(\mathcal{F})$ par $t \to u$ si et seulement s'il existe un contexte $c \in \mathsf{T}_{\square}(\mathcal{F})$, une substitution $\sigma : \mathcal{X} \to \mathsf{T}(\mathcal{F})$ et un couple $(g, d) \in \mathcal{S}$ tels que $t = c \cdot \sigma(g)$ et $u = c \cdot \sigma(d)$. On note $\to^*$ la clôture réflexive transitive de $\to$. Un terme $t \in \mathsf{T}(\mathcal{F})$ est dit irréductible s'il n'existe aucun u tel que $t \to u$. On note Irr($\mathcal{S}$) l'ensemble des termes (clos) irréductibles pour $\mathcal{S}$. Étant donné un langage d'arbres $L \subseteq \mathsf{T}(\mathcal{F})$, on définit

$$\mathcal{S}^*(L) := \{ t \mid t \in \mathsf{T}(\mathcal{F}),\ \exists u \in L.\ u \to^* t \} \qquad \text{réduits de } L$$
$$\mathcal{S}(L) := \mathcal{S}^*(L) \cap \mathrm{Irr}(\mathcal{S}) \qquad \text{formes } \mathcal{S}\text{-normales de } L$$

1. Montrer que si $\mathcal{S}$ est linéaire à gauche, c'est-à-dire si pour tout $(g, d) \in \mathcal{S}$, g est linéaire, alors Irr($\mathcal{S}$) est reconnaissable.

2. Montrer que le nombre d'états d'un automate déterministe reconnaissant Irr($\mathcal{S}$) peut être parfois minoré par 2^{n-1}, n étant la somme des tailles des membres gauches de $\mathcal{S}$.

3. En général, si $\mathcal{S}$ n'est pas linéaire à gauche, Irr($\mathcal{S}$) est-il reconnaissable ? Supposons que l'on sache décider, étant donnés un langage reconnaissable $L \in \mathsf{T}(\mathcal{F})$ et un morphisme $h : \mathsf{T}(\mathcal{F}) \to \mathsf{T}(\mathcal{F})$, si $h(L)$ est reconnaissable. Montrer qu'on peut alors décider si Irr($\mathcal{S}$) est reconnaissable.

On se place dans le cas $\mathcal{F} = \{f(2), g(1), h(1), a(0)\}$.

4. On considère le langage $L = \{ f(t, u) \mid t, u \in \mathsf{T}(\{g(1), h(1), a(0)\}) \}$, et le système de réécriture $\mathcal{S}$ constitué des règles suivantes :

$$f(g(x), h(y)) \to f(x, y) \qquad f(h(x), g(y)) \to f(x, y)$$
$$g(h(x)) \to x \qquad h(g(x)) \to x$$
$$f(a, x) \to x \qquad f(x, a) \to x$$

Les langages L, $\mathcal{S}^*(L)$ et $\mathcal{S}(L)$ sont-ils reconnaissables ?

5. On considère le langage $L = \{ g(h^n(a)) \mid n \geqslant 0 \}$ et le système de réécriture $\mathcal{S}$ consitiué de l'unique règle $g(x) \to f(x, x)$. Les langages L, $\mathcal{S}^*(L)$ et $\mathcal{S}(L)$ sont-ils reconnaissables ?

On se place à nouveau dans le cas général. On appelle préservation de la reconnaissabilité la propriété selon laquelle, pour tout L reconnaissable, $\mathcal{S}^*(L)$ est reconnaissable.

6. On considère un système réécriture $\mathcal{S}$ qui préserve la reconnaissabilité. Que dire des deux problèmes de décision suivants ?
- Étant donnés $u, v \in \mathsf{T}(\mathcal{F})$, est-ce que $v \in \mathcal{S}^*(\{u\})$?
- Étant donnés $L, L' \subseteq (\mathcal{F})$ reconnaissables, est-ce que $L' \subseteq \mathcal{S}^*(L)$?

7. On suppose que $\mathcal{S}$ est linéaire et monadique, c'est-à-dire que, pour toute règle $(g, d) \in \mathcal{S}$, g et d sont linéaires, g est de hauteur au moins 1 et d est soit une variable soit un terme $f(x_1, \ldots, x_k)$ avec $f \in \mathcal{F}_k$ ($k \geqslant 0$) et les x_i tous distincts. Montrer que $\mathcal{S}$ préserve la reconnaissabilité.

8. En déduire que si $\mathcal{S}$ est linéaire et monadique, alors pour tout L reconnaissable $\mathcal{S}(L)$ est reconnaissable.

9. Montrer que la préservation de la reconnaissabilité n'est plus garantie si l'on supprime la condition de linéarité à droite.