Grenoble INP - Ensimag 1ère année — Janvier 2019

Théorie des langages 1

Durée : 2h.
Documents : tous les documents sont autorisés.

Nb : le barême est donné à titre indicatif; la rigueur des preuves et la clarté de la rédaction seront prises en compte dans la notation.

Exercice 1 - Transformation d'automates (5 points)

On considère l'automate $A = (Q, \{a, b\}, \delta, \{q_0\}, \{q_1, q_3\})$ dont la relation de transition δ est définie dans le tableau suivant :

δ	a	b
q_0	q_5	q_1
q_1	q_2	q_3
q_2	q_4	-
q_3	-	q_3
q_4	-	-
q_5	q_6	q_3
q_6	q_0	q_4

▷ **Question 1** *(3 points). En utilisant les méthodes vues en cours, construire le système d'équations associé à cet automate et en déduire une expression régulière représentant le langage reconnu par A.*
Nb : *on éliminera les variables dans l'ordre décroissant (x_6 puis x_5 puis x_4...)*

▷ **Question 2** *(2 points). Construire un automate minimal équivalent à A en détaillant le calcul des classes d'équivalence.*

Exercice 2 - Images réciproques (8 points)

Soient V et W des vocabulaires, et $h : V \to W^*$ une fonction (totale). Cette fonction est étendue à une fonction sur V^* par induction en posant [1] :

— $h(\varepsilon) = \varepsilon$,

— pour tout $a \in V$ et pour tout $w \in V^*$, $h(a.w) = h(a).h(w)$.

Pour $w \in W^*$, on note $h^{-1}(w)$ l'image réciproque de w. Formellement :

$$h^{-1}(w) = \{v \in V^* \mid h(v) = w\}.$$

1. On rappelle que l'opérateur de concaténation peut être omis : on peut aussi bien écrire $w.w'$ que ww'.

Images réciproques

On pose $V = \{a, b, c\}$, $W = \{0, 1\}$ et on considère la fonction h_1 définie par :

$$h_1 : \begin{array}{ccc} a & \to & 01 \\ b & \to & 10 \\ c & \to & 0110 \end{array}$$

Ainsi, $h_1(bac) = 10010110$, $h_1(cca) = 0110011001$ et $h_1^{-1}(0110) = \{c,\ ab\}$.

▷ **Question 1** *(2 points). Calculer, sans donner de justification, les images réciproques par h_1 des mots suivants :*

1. $w_1 = \varepsilon$
2. $w_2 = 0101$
3. $w_3 = 01011010$
4. $w_4 = 0101101$

Les images réciproques sont étendues aux langages de la façon suivante :

$$\forall L \subseteq W^*,\ h^{-1}(L) = \bigcup_{w \in L} h^{-1}(w).$$

▷ **Question 2** *(2 points). Calculer, sans donner de justification, les images réciproques par h_1 des langages suivants :*

1. $L_1 = 0^*1^* + 1^*0^*$
2. $L_2 = (0110)^*$
3. $L_3 = \{(01)^n(10)^n \mid n \geq 1\}$
4. $L_4 = (0 + \varepsilon)(10)^*(1 + \varepsilon)$

Soit $A = (Q, W, \delta, \{q_0\}, F)$ un automate fini **déterministe** (et donc complet) et considérons une fonction quelconque $h : V \to W^*$. On définit l'automate $h^{-1}(A) \stackrel{\text{def}}{=} (Q, V, \mu, \{q_0\}, F)$, où la fonction de transition μ est définie pour tout $q \in Q$ et pour tout $a \in V$ par [2] $\mu(q, a) \stackrel{\text{def}}{=} \delta^*(q, h(a))$. L'objet des questions suivantes est de prouver que si A reconnaît un langage L, alors $h^{-1}(A)$ reconnaît le langage $h^{-1}(L)$.

▷ **Question 3** *(2 points). Montrer par induction sur w que pour tout $w \in V^*$ et pour tout $q \in Q$, $\mu^*(q, w) = \delta^*(q, h(w))$.*
Remarque : On pourra admettre que $h(w.w') = h(w).h(w')$ et que pour tout $q \in Q$, $\delta^*(\delta^*(q, w), w') = \delta^*(q, ww')$.

▷ **Question 4** *(1 point). En déduire que $h^{-1}(A)$ reconnaît un mot w si et seulement si A reconnaît $h(w)$. Si L est régulier, que peut-on dire du langage $h^{-1}(L)$?*

▷ **Question 5** *(1 point). En admettant que $M \stackrel{\text{def}}{=} \{0^n1^n \mid n \geq 0\}$ n'est pas régulier et* ***sans se servir du lemme de l'étoile****, montrer que $N \stackrel{\text{def}}{=} \{a^nb^{2n} \mid n \geq 0\}$ n'est pas régulier.*

2. On notera que, par construction, $h^{-1}(A)$ est un automate déterministe.

Exercice 3 - Grammaires (7 points)

On rappelle que les règles d'une grammaire s'écrivent sous la forme $u \rightarrow v$ avec u une chaîne sur V^+ et v une chaîne sur V^* et $V = V_T \cup V_N$.

On s'intéresse ici à décrire le langage des règles d'une grammaire, et ceci pour les différentes classes de grammaires. On fixe le vocabulaire terminal à $\{a, b, c\}$ et le vocabulaire non-terminal à $\{X, Y, Z\}$. On choisit ici la notation Backus-Naur (::= au lieu de $\rightarrow$ et ; pour délimiter la fin de règle).

Voici une grammaire G décrivant ce langage :

$$\begin{array}{lcl}
LR & \rightarrow & R; \mid R; LR \\
R & \rightarrow & G ::= D \\
G & \rightarrow & T \mid N \mid TG \mid NG \\
D & \rightarrow & eps \mid G \\
T & \rightarrow & a \mid b \mid c \\
N & \rightarrow & X \mid Y \mid Z
\end{array}$$

On a ici $V_T = \{a, b, c, X, Y, Z, eps, ;, ::=\}$ et $V_N = \{LR, R, D, G, T, N\}$, LR étant l'axiome. Le symbole eps représente la chaîne vide.

▷ **Question 1** *(1 point)*
Soit la liste de règles suivante :

```
X ::= eps ;
X ::=  aXb ;
```

Donner l'arbre de dérivation par la grammaire ci-dessus de cette liste de règles.

▷ **Question 2** *(2 points)*
Le langage $L(G)$ est-il régulier ? Justifier votre réponse soit en donnant un automate ou une expression régulière soit en montrant l'aspect non régulier du langage.

▷ **Question 3** *(1 point)*
On rappelle que la classification de Chomsky est basée sur la forme des règles. Modifier la grammaire précédente pour que le langage engendré représente des grammaires n'admettant que des règles qui suivent le critère hors-contexte. On appellera LHC le langage associé à cette grammaire. Ce langage est-il régulier ?

▷ **Question 4** *(2 points)*
On rappelle que le critère pour les grammaires sous-contexte est le suivant : toutes les règles sont de la forme $u \rightarrow v$ (ici $u ::= v$) avec u une chaîne sur V^+, v une chaîne sur V^* et $|u| \leq |v|$, $|u|$ désignant la longueur de la chaine u.

Modifier la grammaire initiale G pour n'admettre que des règles qui suivent le critère sous-contexte. On appellera LSC ce langage.

▷ **Question 5** *(1 point) Prouvez par induction que la grammaire proposée est correcte, c'est-à-dire que les règles produites par la grammaire précédente vérifient bien la contrainte sous-contexte. On expliquera précisément sur quoi porte l'induction et les hypothèse d'induction.*

▷ **Question 6** *(bonus) Le langage LSC est-il régulier ? On argumentera la réponse.*