

## TD2 Licence INF208 Théorie des langages, 2006/07

### Exo 1 (Glushkov)

Un langage  $L \subseteq \Sigma^*$  est dit *local* s'il existe des ensembles de symboles  $P, S \subseteq \Sigma$  ainsi qu'un ensemble de mots de 2 lettres  $T \subseteq \Sigma^2$  tels que  $L \setminus \{\epsilon\} = (P\Sigma^* \cap \Sigma^*S) \setminus \Sigma^*T\Sigma^*$ .

1. Montrer que tout langage local est accepté par un NFA ayant  $n + 1$  états, où  $n = |\Sigma|$ .
2. Montrer que le langage  $(ab)^*$  est local.
3. Donner un exemple de langage régulier qui n'est pas local.
4. Montrer que pour tout langage régulier  $L \subseteq \Sigma^*$  il existe un alphabet  $\Gamma$ , un langage local  $K \subseteq \Gamma^*$  et un morphisme alphabétique  $h : \Gamma \rightarrow \Sigma$  tels que  $L = h(K)$ .
5. Soit  $E$  une expression rationnelle sur  $\Sigma$  dans laquelle chaque lettre apparaît au plus une fois. Montrer (en utilisant la construction de Glushkov) que le langage associé  $\mathcal{L}(E)$  est local.

### Exo 2 (algorithme de minimisation de Brzozowski)

Soit  $\mathcal{A} = (Q, \Sigma, \delta, I, F)$  un NFA sans transitions- $\epsilon$ . Le langage à gauche  $L_g^{\mathcal{A}}(q)$  de l'état  $q \in Q$  est le langage accepté par l'automate  $\mathcal{A}_g(q) = (Q, \Sigma, \delta, I, q)$ . Le langage à droite  $L_d^{\mathcal{A}}(q)$  de l'état  $q \in Q$  est le langage accepté par l'automate  $\mathcal{A}_d(q) = (Q, \Sigma, \delta, q, F)$ .

Soit  $d(\mathcal{A})$  l'automate des parties associé à  $\mathcal{A}$  (en ne gardant que les états accessibles). Soit  $r(\mathcal{A})$  l'automate miroir associé à  $\mathcal{A}$  (obtenu en renversant les flèches et en échangeant  $I$  et  $F$ ).

1. Montrer que  $r(\mathcal{A})$  reconnaît les miroirs de mots dans  $\mathcal{L}(\mathcal{A})$  :  $\mathcal{L}(r(\mathcal{A})) = \mathcal{L}(\mathcal{A})^R$ .
2. Montrer que le langage à gauche  $L_g^{r(\mathcal{A})}(q)$  de l'état  $q$  dans  $r(\mathcal{A})$  est le miroir du langage à droite  $L_d^{\mathcal{A}}(q)$  de  $q$  dans  $\mathcal{A}$ . (Pareil pour gauche et droite inversés.)
3. Montrer que chaque langage à droite  $L_d^{d(\mathcal{A})}(R)$  associé à un état  $R$  de  $d(\mathcal{A})$  est l'union des langages à droite  $L_d^{\mathcal{A}}(r)$ , où  $r \in R$ .
4. Montrer que  $\mathcal{A}$  est déterministe ssi les langages à gauche associés à ses états sont deux à deux disjoints et  $|I| = 1$ .
5. Montrer qu'un DFA complet, dont tous les états sont accessibles, est minimal ssi les langages à droite associés à ses états sont deux à deux différents.
6. Montrer que  $d(r(d(r(\mathcal{A}))))$  est l'automate minimal de  $\mathcal{A}$ .

### Exo 3

Donner des formules MSO  $\phi_L$  pour chacun des langages réguliers  $L$  suivants :

1.  $L_1 = \{w \in \{a, b\}^* \mid |w|_a \text{ est divisible par } 3\}$
2.  $L_2 \subseteq \{a, b\}^*$  contient tous les mots  $w$  tels que dans chaque préfixe  $v$  de  $w$ , la différence  $|v|_a - |v|_b$  est 0, 1 ou 2.
3.  $L_3 \subseteq \{a, b\}^*$  contient tous les mots  $w$  qui n'ont pas de facteur  $aab$ .
4.  $L_4 \subseteq \{a, b\}^*$  contient tous les mots  $w$  qui n'ont pas de sous-mot  $aab$ .
5.  $L_5 \subseteq \{a, b\}^*$  contient tous les mots  $w$  qui ont un nombre pair de facteurs  $abba$ .

**Exo 4** Montrer que tout langage rationnel sur un alphabet  $\Sigma$  est définissable dans MSOL par induction structurelle sur une expression rationnelle définissant le langage.

Étant donné une expression rationnelle  $E$ , on pourra construire une formule  $\varphi_E(x, y)$  avec 2 variables libres  $x$  et  $y$  telle qu'un mot  $w$  satisfait  $\varphi_E(x, y)$  sous une interprétation  $\{x \mapsto i, y \mapsto j\}$  ssi  $i \leq j$  et le facteur du mot  $w$  compris entre les positions  $i$  et  $j$  est dans  $\mathcal{L}(E)$ .

**Exo 5** On considère la logique FOL(+) dans laquelle les formules sont construites à partir des formules atomiques  $a(x)$ , qui s'interprète comme "la position  $x$  porte la lettre  $a$ ", et  $x = y + z$  avec l'interprétation naturelle de l'addition.

1. Montrer que tout langage définissable dans FOL est aussi définissable dans FOL(+).
2. Montrer que  $\{a^n b^n \mid n > 0\}$  est définissable dans FOL(+).
3. Dédire que la relation ternaire  $R(x, y, z) \equiv x = y + z$  n'est pas définissable dans MSOL.

**Exo 6**

Soit  $\Sigma_0 = \{a\}, \Sigma_1 = \{g\}, \Sigma_2 = \{f\}$ .

1. Construire un automate d'arbre montant (descendant, respectivement) qui reconnaît l'ensemble des arbres de la forme  $f(f(a, t_1), g(t_2))$ , où  $t_1, t_2 \in T(\Sigma)$ ,  $\Sigma = \Sigma_0 \cup \Sigma_1 \cup \Sigma_2$ .
2. Construire un DTA (automate déterministe montant) qui reconnaît l'ensemble des arbres de  $T(\Sigma_0 \cup \Sigma_1)$  dont la hauteur est paire.

**Exo 7**

Soit  $\Sigma = \Sigma_0 \cup \Sigma_1 \cup \Sigma_2$ , où  $\Sigma_0 = \{0, 1\}, \Sigma_1 = \{\text{not}\}, \Sigma_2 = \{\text{and, or}\}$ .

Construire un DTA qui reconnaît les arbres de  $T(\Sigma)$  correspondant aux formules booléennes qui s'évaluent à *Vrai*.

**Exo 8**

On considère des arbres binaires (chaque nœud a 0 ou 2 fils) dont les nœuds sont étiquetés par  $\Sigma = \Sigma_0 \cup \Sigma_2$ , où  $\Sigma_0 = \Sigma_2 = \{a, b\}$ .

Construire des automates d'arbre (montants ou descendants) reconnaissant les langages suivants (si possible, proposez des automates déterministes) :

1. L'ensemble des arbres dont la feuille la plus à gauche est étiquetée par  $a$ .
2. L'ensemble des arbres contenant 2 nœuds  $u, v$  étiquetés par  $a$ , tels que  $u$  est ancêtre de  $v$ .
3. L'ensemble des arbres contenant 2 nœuds  $u, v$  étiquetés par  $a$ , tels que  $u$  et  $v$  ne sont pas ancêtre l'un de l'autre.
4. L'ensemble des arbres contenant un nombre pair de nœuds étiquetés par  $a$ .
5. L'ensemble des arbres dont la branche la plus à droite contient au moins un nœud étiqueté par  $a$ .