

Automates d'arbres pour XML

Pierre-Alain Reynier

21 novembre 2011

1 Automates d'arbres d'arité fixe

Les automates d'arbres considérés ici sont sur des arbres d'arité fixe.

Exercice 1 : Expressivité des automates déterministes descendants. Nous allons démontrer dans cet exercice que les automates d'arbres descendants déterministes sont strictement moins expressifs que les automates d'arbres ascendants.

Nous considérons le langage d'arbres formé des deux arbres $f(a, b)$ et $f(b, a)$. L'alphabet est donc $\Sigma = \{f, a, b\}$, f est un symbole d'arité 2, a et b sont des symboles d'arité 0 (feuilles), et nous considérons le langage $L = \{f(a, b), f(b, a)\}$.

1. Montrer que L est accepté par un automate ascendant (déterministe) \mathcal{A} .
2. Supposer que L est accepté par un automate descendant déterministe \mathcal{B} , et montrer qu'alors l'arbre $f(a, a)$ est également accepté par \mathcal{B} .
3. Conclure.

Exercice 2 : Donner un automate d'arbre (ascendant puis descendant) acceptant les arbres binaires tels que les noeuds internes sont étiquetés par le symbole b , et les feuilles par le symbole a . On discutera de la possibilité d'obtenir des automates déterministes.

Exercice 3 : Donner un automate d'arbre (ascendant puis descendant) acceptant les arbres binaires tels que la racine est étiquetée par le symbole a , le sous-arbre gauche est non vide et tous ses noeuds sont étiquetés par le symbole b et le sous-arbre droit est non vide et tous ses noeuds sont étiquetés par le symbole c . On discutera de la possibilité d'obtenir des automates déterministes.

Exercice 4 : Donner un automate d'arbre (ascendant puis descendant) acceptant les arbres binaires tels que le nombre de noeuds internes est pair. On discutera de la possibilité d'obtenir des automates déterministes.

2 Automates de haies

Exercice 5 : On considère l'automate de haies descendant suivant : $\mathcal{A} = (\Sigma, Q, \delta, I, F)$ avec :

- $\Sigma = \{\text{films, film, titre, annee, acteur}\}$.
- $Q = \{q_0, q_1, q_2, q_3, q_4, q_f\}$.
- $\delta(\text{films}, q_0) = q_1^*$, $\delta(\text{film}, q_1) = q_2q_3q_4^+$, $\delta(\text{titre}, q_2) = q_f$,
 $\delta(\text{annee}, q_3) = q_f$ et $\delta(\text{acteur}, q_4) = q_f$,
- $I = \{q_0\}$, et
- $F = \{q_f\}$.

1. Donnez un arbre appartenant au langage de cet automate.
2. Modifiez cet automate pour qu'il autorise qu'il n'y ait pas d'année associée à un film.
3. Donnez une DTD associée à ce nouvel automate.

Exercice 6 : Transformez la DTD suivante en un automate de haies (descendant).

```
<!ELEMENT stock      (item*) >
<!ATTLIST stock
          nom CDATA #REQUIRED >
<!ELEMENT item      (marque,modele,annee) >
<!ATTLIST item
          prix NMTOKEN #REQUIRED >
<!ELEMENT marque    (#PCDATA) >
<!ELEMENT modele    (#PCDATA) >
<!ELEMENT annee     (#PCDATA) >
```

Rappel : on ignore les types dans la représentation des documents XML par des arbres.