Automates d’arbre

TD n°1 : Recognizable Tree Languages and Finite Tree Automata

Exercise 1 :
Let \( \mathcal{F} = \{ f(2), g(1), a(0) \} \). Give a DFTA and a top-down DFTA for the set \( G(t) \) of ground instances of the term \( t = f(f(a,x), g(y)) \) which is defined by :

\[
G(t) = \{ f(f(a,u), g(v)) \mid u, v \in T(\mathcal{F}) \}
\]

Exercise 2 :
Let \( \mathcal{F} = \{ f(2), g(1), a(0) \} \). Give a DFTA and a top-down NFTA for the set \( M(t) \) of terms which have a ground instance of the term \( t = f(a,g(x)) \) as a subterm, ie. \( M(t) = \{ C[f(a,g(u))] \mid C \in \mathcal{C}(\mathcal{F}), u \in T(\mathcal{F}) \} \).

Prove that \( M(t) \) is not recognizable by a top-down DFTA or even by a finite union of top-down DFTA.

*hint* : you can use without proof the following fact (prove it if you have time) : let \( M(t) \) be a tree language. The path language \( \pi(t) \) is defined by :

- if \( t \) is a constant, \( \pi(t) = \{ t \} \)
- if \( t = f(t_1, ..., t_n) \), \( \pi(t) = \bigcup_{i=1}^{n} \{ fiw \mid w \in \pi(t_i) \} \)

Let \( L \) be a tree language. The path language of \( L \) is \( \pi(L) = \bigcup_{t \in L} \pi(t) \). The path closure of \( L \) is defined by \( \text{pathclosure}(L) = \{ t \mid \pi(t) \subseteq \pi(L) \} \).

\( L \) is recognizable by a top-down DFTA iff \( L \) is recognizable and path closed, ie. \( L = \text{pathclosure}(L) \).

Exercise 3 :
Are the following tree languages recognizable (by a bottom-up FTA) ?

- \( \mathcal{F} = \{ g(1), a(0) \} \) and \( L \) the set of ground terms of even height.
- \( \mathcal{F} = \{ f(2), g(1), a(0) \} \) and \( L \) the set of ground terms of even height.
- \( \mathcal{F} = \{ f(2), a(0) \} \) and \( L = \{ f(t,t) \mid t \in T(\mathcal{F}) \} \).
- \( \mathcal{F} \) any ranked alphabet which contains a constant symbol \( a \) and a binary symbol \( f \) and \( L = \{ f(t,t) \mid t \in T(\mathcal{F}) \} \).

Exercise 4 :
1) Let \( \mathcal{E} \) be a finite set of linear terms on \( T(\mathcal{F}, \chi) \). Prove that \( \text{Red}(\mathcal{E}) = \{ C[t\sigma] \mid C \in \mathcal{C}(\mathcal{F}), t \in \mathcal{E}, \sigma \text{ ground substitution} \} \) is recognizable.
2) Let \( R \) be a left linear finite rewrite system. Prove that the set of irreducible ground terms is recognizable.
3) Give a non left linear and non redundant finite rewrite system whose set of irreducible ground terms is recognizable.
4) Give a recognizable set which is not the set of irreducible ground terms of a rewrite system.
5) Prove that the number of states of a DFTA recognizing \( \text{Red}(\mathcal{E}) \) can be at least exponential in the number of nodes of \( \mathcal{E} \).

*hint* : use for example \( \mathcal{E} = \{ h(f(x_1, f(x_2, ..., f(x_{p-1}, f(a, x_p))))...)) \} \) with \( p \) fixed.
6) Prove that if \( \mathcal{E} \) contains only ground terms, we can construct a DFTA recognizing \( \text{Red}(\mathcal{E}) \) whose number of states is at most \( n + 2 \), where \( n \) is the number of nodes of \( \mathcal{E} \).

*taken from Tree Automata Techniques and Applications*
Exercise 5:

Let \( \mathcal{F} = \{ f(2), a(0), b(0) \} \).

1) Let \( L_1 \) be the smallest set such that:
   - \( f(a, b) \in L_1 \)
   - \( t \in L_1 \Rightarrow f(a, f(t, b)) \in L_1 \)
   Prove that \( L_1 \) is recognizable.

2) Prove that \( L_2 = \{ t \in \mathcal{T}(\mathcal{F}) \mid |t|_a = |t|_b \} \) is not recognizable.

3) Let \( L \) be recognizable on \( \mathcal{F} \) and \( C(L) \) be the closure of \( L \) by the congruence generated by the equation \( f(x, y) = f(y, x) \). Prove that \( C(L) \) is recognizable.

4) Let \( L \) be recognizable on \( \mathcal{F} \) and \( AC(L) \) be the closure of \( L \) by the congruence generated by the equations \( f(x, y) = f(y, x) \) and \( f(x, f(y, z)) = f(f(x, y), z) \). Prove that \( AC(L) \) is not recognizable in general.

5) Let \( L \) be recognizable on \( \mathcal{F} \) and \( A(L) \) be the closure of \( L \) by the congruence generated by the equation \( f(x, f(y, z)) = f(f(x, y), z) \). Prove that \( A(L) \) is not recognizable in general.