

# Automates d'arbre

TD n°4 : Hedge Automata, alternating word automata and membership

**Exercise 1 :**

Let  $\Sigma = \{a, b, c\}$ . Define a NFHA  $\mathcal{A}$  whose language is the trees which have two nodes labelled by  $b$  whose greatest common ancestor is labelled by  $c$ .

If  $\mathcal{X}$  is a set of propositional variables, let  $\mathbb{B}(\mathcal{X})$  be the set of positive propositional formulae on  $\mathcal{X}$ , i.e., formulae generated by the grammar  $\phi ::= \perp \mid \top \mid x \in X \mid \phi \vee \phi \mid \phi \wedge \phi$ .

A AWA (Alternating Word Automata) is a tuple  $\mathcal{A} = (Q, \Sigma, Q_0, Q_f, \delta)$  where  $\Sigma$  is a finite set (alphabet),  $Q$  is a finite set (of states),  $Q_0 \subseteq Q$  (initial states),  $Q_f \subseteq Q$  (final states) and  $\delta$  is a function from  $Q \times \Sigma$  to  $\mathbb{B}(Q)$  (transition function). A run of  $\mathcal{A} = (Q, \Sigma, Q_0, Q_f, \delta)$  on a word  $w$  is a tree  $t$  labelled by  $Q \times \mathbb{N}$  such that :

- if  $w = \varepsilon$ , then  $t = (q_0, 0)$  with  $q_0 \in Q_0$ .
- if  $w = a.w'$ , then  $t = (q_0, k)(t_1, \dots, t_n)$  where  $k$  is the length of  $w$ ,  $q_0 \in Q_0$  and such that for all  $i$ ,  $t_i$  is a run of  $w'$  on  $(Q, \Sigma, \{q_i\}, Q_f, \delta)$  for some  $q_i$  satisfying  $\{q_1, \dots, q_n\} \models \delta(q_0, a)$ .

We say that a run is accepting if every leaf of the form  $(q, 0)$  satisfies that  $q \in Q_f$ . Notice that a run may have leaves of the form  $(q, i)$  with  $i > 0$  if  $\emptyset \models \delta(q_0, a)$ . Those leaves are considered as 'success' leaves in this semantic. The language of a AWA is the set of words which have an accepting run.

**Exercise 2 :**

- 1) Assume given a simple AWA  $\mathcal{A}$ , i.e., an AWA for which the transition function is with values in formulae of the form  $q \wedge q'$ ,  $q \vee q'$ ,  $\top$  or  $\perp$  with  $q, q'$  states. Given a word  $w$ , construct a top-down NFTA  $\mathcal{A}_w$ , in time  $O(|w|.|\mathcal{A}|)$ , such that  $w \in L(\mathcal{A})$  iff  $L(\mathcal{A}_w) \neq \emptyset$ .  
*hint : If  $\mathcal{A} = (Q, \Sigma, Q_0, Q_f, \delta)$ , you will construct a top-down NFTA of the form  $(Q', \Sigma', Q'_0, \Delta')$  with  $\Sigma' = \{\wedge(2), \vee(1), \top(0), \text{succ}(0)\}$  and  $Q' = \{0, 1, \dots, |w|\} \times Q$ . The general idea is to have a correspondence between accepting runs of  $w$  in  $\mathcal{A}$  and accepting runs in  $\mathcal{A}_w$ .*
- 2) Generalize this construction (with the same complexity) to general AWA.  
*hint : The alphabet will be the same, but  $Q'$  will be something like*

$$\{0, \dots, |w|\} \times (Q \sqcup \bigsqcup_{(a,q) \in \Sigma \times Q} Pos(\delta(q, a)))$$

where  $Pos(\phi)$  is the set of positions of the formula  $\phi$  seen as a tree. You may use  $\varepsilon$ -transitions.

- 3) Deduce that membership in AWA is solvable in time  $O(|w|.|\mathcal{A}|)$ .
- 4) Prove that this problem is P-hard by reducing emptiness in NFTA.

**Exercise 3 :**

- 1) Prove that the membership problem for a NFHA whose horizontal languages are given by (non necessarily deterministic) word automata is in PTIME. Prove that in the case where the NFHA is deterministic and every horizontal languages are given by deterministic automata, then your algorithm is in linear time.

- 2) Prove that membership for DFHA where horizontal languages are given by AWA is in PTIME.
- 3) Let  $\Phi$  be a propositional formula in CNF with  $n$  variables  $x_1, \dots, x_n$ . Construct, in polynomial time, an AWA  $\mathcal{A}_\Phi$  whose language is  $\{w \in \{0, 1\}^n \mid w \models \Phi \text{ i.e. } \Phi_{[x_i \leftarrow w_i]} = \top\}$ .
- 4) Deduce that membership for NFHA where horizontal languages are given by AWA is NP-complete.