Tree Automata and Applications: TD1
Recognizable Tree Languages and Finite Tree Automata

Emile Contal*
emile.contal@ens-cachan.fr

September 20, 2013

Exercise 1. Let $\mathcal{F} = \{f(,), g(), a\}$. Define a top-down NFTA, a NFTA and a DFTA for the set $G(t)$ of ground instances of term $t = f\big(f(a,x), g(y)\big)$ which is defined by $G(t) = \Big\{f\big(f(a,u), g(v)\big) \mid u, v \in T(\mathcal{F})\Big\}$.

Is it possible to define a top-down DFTA for this language?

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

Is it possible to define a top-down DFTA for this language? A finite union of top-down DFTA ?

Exercise 3. Let $\mathcal{F} = \{g(), a\}$. Is the set of ground terms whose height is even recognizable?

Let $\mathcal{F} = \{f(,), g(), a\}$. Is the set of ground terms whose height is even recognizable?

Exercise 4. Let $\mathcal{F} = \{f(,), a, b\}$. Define a recognizable language L for which there exists no top-down DFTA recognizing L. Don't forget to prove that L is recognizable.

Exercise 5. Let $\mathcal{F} = \{f(,), a\}$. Prove that the set $L = \{f(t,t) \mid t \in T(\mathcal{F})\}$ is not recognizable.

Let $\mathcal{F}$ be any ranked alphabet which contains at least one constant symbol a and one binary symbol $f(,)$. Prove that the set $L = \{f(t,t) \mid t \in T(\mathcal{F})\}$ is not recognizable.

Exercise 6. Let $\mathcal{F}_1 = \{or(,), and(,), not(), 0, 1, x\}$. A ground term over $\mathcal{F}$ can be viewed as a boolean formula over variable x. Define a DFTA which recognizes the set of satisfiable boolean formulae over x.

Let $\mathcal{F}_n = \{or(,), and(,), not(), 0, 1, x_1, \ldots, x_n\}$. A ground term over $\mathcal{F}_n$ can be viewed as a boolean formula over variables $x_1, \ldots, x_n$. Define a DFTA which recognizes the set of satisfiable boolean formulae over $x_1, \ldots, x_n$.

*Taken from the TATA book.

Exercise 7. *Motivation: Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two FTAs. We remind that the size of the construction for the FTA recognizing $\mathcal{L}(\mathcal{A}_1) \cap \mathcal{L}(\mathcal{A}_2)$ is the product of the respective sizes of $\mathcal{A}_1$ and $\mathcal{A}_2$. Therefore, the intersection non-emptiness problem for tree automata is* EXPTIME*-complete.*

A *flat automaton* is a tree automaton which has the following property: there is an ordering $>$ on the states and a particular state $q_\top$ such that the transition rules have one of the following forms:

1. $f(q_\top, \ldots, q_\top) \to q_\top$
2. $f(q_1, \ldots, q_n) \to q$ with $q > q_i$ for every i
3. $f(q_\top, \ldots, q_\top, q, q_\top, \ldots, q_\top) \to q$.

Moreover, we assume that all terms are accepted in the state $q_\top$. (The automaton is called *flat* because there are no nested loop). Prove that the intersection of two flat automata is a finite union of automata whose size is linear in the sum of the two original automata.

Deduce from the above result that the intersection non-emptiness problem for flat automata is in NP.