Tree Automata and Applications
Exercise session 3

Luc Lapointe
luc.lapointe@ens-paris-saclay.fr
home.lmf.cnrs.fr/LucLapointe/

Exercise 1 - Recognizing an abstract language

1. Let $\mathcal{E}$ be a finite set of linear terms on $T(\mathcal{F}, \mathcal{X})$. Prove that
$$\mathrm{Red}(\mathcal{E}) \stackrel{\text{def}}{=} \{C[t\sigma] \mid C \in \mathcal{C}(\mathcal{F}), t \in \mathcal{E}, \sigma \text{ ground substitution}\}$$
is recognizable.
2. Prove that if $\mathcal{E}$ contains only ground terms, then one can construct a DFTA recognizing $\mathrm{Red}(\mathcal{E})$ whose number of states is at most $n + 2$, where n is the number of nodes of $\mathcal{E}$.

Exercise 2 - Decision problems

Let $\mathcal{F}$ a signature. We consider the Ground Instance Intersection **(GII)** problem:
Input : t a term in $T(\mathcal{F}, \mathcal{X})$ and $\mathcal{A}$ a NFTA.
Question : Is there at least one ground instance of t accepted by $\mathcal{A}$?

1. Suppose that t is linear. Prove that **(GII)** is P-complete.
Hint: The emptiness problem for tree automata is P-complete.
2. Suppose that $\mathcal{A}$ is deterministic. Prove that **(GII)** is NP-complete.
3. Prove that **(GII)** is EXPTIME-complete.
Hint: For the hardness, reduce the intersection non-emptiness problem (admitted to be EXPTIME-complete).
4. Deduce that the following problem is decidable:
Input : t a term in $T(\mathcal{F}; \mathcal{X})$ and linear terms $t_1, ..., t_n$.
Question : Is there a ground instance of t which is not an instance of any t_i?

Exercise 3 - Residuals

Exercise from January 2023 exam.

For $\mathcal{F} = \{f(2), a(0)\}$ and $n > 0$, let L_n be the language of trees that have at least one branch of length exactly n, i.e.

$$L_n \overset{\text{def}}{=} \{t \in T(\mathcal{F}) \mid \exists p \in \text{Pos}(t) : |p| = n - 1 \wedge t(p) = a\}$$

e.g. $f(a, f(a, f(a, a))) \in L_3$ because it contains one branch of length 3 (as well as one of length 2 and two of length 4).

1. Give a bottom-up NFTA for L_n with $n + 1$ states.
2. Show that the minimal DFTA for L_n has at least 2^{n-1} states.
   Hint: For $I \subseteq \{2, ..., n\}$, build a tree t_I that has a branch of length i iff $i \in I$.

Let $L \subseteq T(\mathcal{F})$ be a language of trees and $C \in \mathcal{C}(\mathcal{F})$ a context. The *residual* of L by C is defined as

$$C^{-1}L \overset{\text{def}}{=} \{t \in T(\mathcal{F}) \mid C[t] \in L\}.$$

We define $R(L) \overset{\text{def}}{=} \{C^{-1}L \mid C \in \mathcal{C}(\mathcal{F})\}$ as the set of residuals of L.

3. Show that if L is recognizable, then $|R(L)|$ is finite.

Homework - Direct images of a homomorphism

Let $\mathcal{F} = \{f(2), g(1), a(0)\}$ and $\mathcal{F}' = \{f'(2), g(1), a(0)\}$. Let us consider the tree homomorphism h determined by $h_{\mathcal{F}}$ defined by

$$\begin{aligned} h_{\mathcal{F}}(f) &= f'(x_1, x_2) \\ h_{\mathcal{F}}(g) &= f'(x_1, x_1) \\ h_{\mathcal{F}}(a) &= a. \end{aligned}$$

1. Is $h(T(\mathcal{F}))$ recognizable?
2. Let $L = \{g^i(a) \mid i \geq 0\}$, then L is a recognizable tree language. Is $h(L)$ recognizable?