Automates d'arbre

TD n°2 : Decision problems & tree homomorphisms

September 19, 2017

Exercise 1 : Back from TD1 : an abstract language.

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

Solution:

1) Do the case where $\mathcal{E}$ is a singleton $\{t\}$, t linear (the general case can be deduced by finite union). $Red(\{t\}$ is recognized by the following NFTA : $Q = \{q_\perp\} \cup Pos(t)$, $F = \{\epsilon\}$ and $\Delta =$
   - ⋆ $f(q_1, ..., q_n) \longrightarrow q_\perp$ for all $f \in \mathcal{F}$, q_1, ..., $q_n \in Q$
   - ⋆ $q_\perp \longrightarrow p$ for all $p \in Pos(t)$ such that $t(p)$ is a variable
   - ⋆ $f(p.1, ..., p.n) \longrightarrow p$ if $t(p) = f$
   - ⋆ $f(q_1, ..., q_n) \longrightarrow \epsilon$ for all $f \in \mathcal{F}$ and q_1, ..., $q_n \in Q$ such that there exists $i \in \{1, ..., n\}$ such that $q_i = \epsilon$
2) Let $St(\mathcal{E})$ be the set of all subterms of a term in $\mathcal{E}$. Then the following DFTA works : $Q = \{q_t \mid t \in St(\mathcal{E})\} \cup \{q_\perp, q_\top\}$, $F = \{q_\top\}$ and $\Delta =$
   - ⋆ $f(q_{t_1}, ..., q_{t_n}) \longrightarrow q_{f(t_1,...,t_n)}$ if $f(t_1, ...t_n) \in St(\mathcal{E}) \setminus \mathcal{E}$
   - ⋆ $f(q_{t_1}, ..., q_{t_n}) \longrightarrow q_\top$ if $f(t_1, ..., t_n) \in \mathcal{E}$
   - ⋆ $f(q_{t_1}, ..., q_{t_n}) \longrightarrow q_\perp$ else
   - ⋆ $f(q_1, ..., q_n) \longrightarrow q_\top$ if there is at least one $q_i = q_\top$
   - ⋆ $f(q_1, ..., q_n) \longrightarrow q_\perp$ else

Exercise 2 : Back from TD1 : Commutative closure

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 $L_2 = \{t \in T(\mathcal{F}) \mid |t|_a = |t|_b\}$ 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.

Solution:

1) $Q = \{q_a, q_b, q_f, q_\top\}$, $F = \{q_\top\}$ and $\Delta =$
   - ⋆ $a \longrightarrow q_a$
   - ⋆ $b \longrightarrow q_b$
   - ⋆ $f(q_a, q_b) \longrightarrow q_\top$
   - ⋆ $f(q_\top, q_b) \longrightarrow q_f$

- $\star$ $f(q_a, q_f) \longrightarrow q_\top$

2) By the pumping lemma.
3) Given a NFTA for L, construct a NFTA for $C(L)$ by adding for every rule of the form $f(q, q') \longrightarrow q''$, the rule $f(q', q) \longrightarrow q''$.
4) $AC(L_1) = L_2$.

Exercise 3 : Decisions problems

We consider the **(GII)** problem (ground instance intersection) :

Instance : 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.
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 complement problem :
   Instance : 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 ?
   is decidable.

Solution:

1) in P : use a construction similar to exercise 1, intersect with $\mathcal{A}$ and test the non-emptiness.
   P-hard : testing the emptiness of $\mathcal{A}$ is equivalent to testing **(GII)** on $\mathcal{A}$ and a variable.
2) in NP : guess for each variable an accessible state of $\mathcal{A}$ and verify that you can complete this to an accepting run by running the automata.
   NP-hard : We reduce **(SAT)** : let $\mathcal{F} = \{\neg(1), \vee(2), \wedge(2), \bot(0), \top(0)\}$ and $\mathcal{A}_{SAT}$ the DFTA with $Q = \{q_\top, q_\bot\}$, $F = \{q_\top\}$ and $\Delta =$
   - $\star$ $\bot \longrightarrow q_\bot$
   - $\star$ $\top \longrightarrow q_\top$
   - $\star$ $\neg(q_\alpha) \longrightarrow q_{\neg\alpha}$
   - $\star$ $\vee(q_\alpha, q_\beta) \longrightarrow q_{\alpha\vee\beta}$
   - $\star$ $\wedge(q_\alpha, q_\beta) \longrightarrow q_{\alpha\wedge\beta}$

   The language of $\mathcal{A}_{SAT}$ is the set of closed valid formulae.
   Let ϕ a CNF formula, $\phi = \bigwedge_{i=1}^{n} c_i$ where c_i are clauses. Define t_{c_i} by induction on the size of c_i :
   - if $c_i = x_j$, $t_{c_i} = x_j$
   - if $c_i = \neg x_j$, $t_{c_i} = \neg(x_j)$
   - if $c_i = x_j \vee c_i'$, $t_{c_i} = \vee(x_j, t_{c_i'})$
   - if $c_i = \neg x_j \vee c_i'$, $t_{c_i} = \vee(\neg(x_j), t_{c_i'})$

   Then $t_\phi = \wedge(t_{c_1}, \wedge(t_{c_2}, ..., \wedge(t_{c_{n-1}}, t_{c_n})...))$. ϕ is satisfiable iff a closed instance of t_ϕ is recognized by $\mathcal{A}_{SAT}$.
3) in EXP : for each coloring of t by states (exponentially many) :
   - check that the coloring of every occurrence of a variable is an accessible state (in P)
   - check that the coloring corresponds to an accepting run (in P)
   - for every variable, let $\{q_1, ..., q_n\}$ be the set of the colorings of all occurrence of x. Check that $L(\mathcal{A}_{q_1}) \cap ... \cap L(\mathcal{A}_{q_n})$ is non empty where A_q is the NFTA obtained from $\mathcal{A}$ by changing the set of final states to $\{q\}$ (in P)

   EXP-hard : We reduce intersection non-emptiness : let $(A_k = (Q_k, \mathcal{F}, I_k, \Delta_k))_{k\in\{1,...,n\}}$ a finite sequence of top-down NFTA (we can transform a bottom-up NFTA to a top-down one in polynomial time). We suppose that all the Q_k are disjoint. Define :
   - $\mathcal{F}' = \mathcal{F} \cup \{h(n)\}$

— $t = h(x, ..., x)$
— $\tilde{\mathcal{A}} = (\bigsqcup Q_k \sqcup \{q_0\}, \mathcal{F}', \{q_0\}, \Delta' \sqcup \bigsqcup \Delta_k)$ where

$$\Delta' = \{q_0(h(x_1, ..., x_n)) \longrightarrow h(q_1(x_1), ..., q_n(x_n)) \mid \ for \ q_k \in I_k\}$$

Then $L(\mathcal{A}_1) \cap ... \cap L(\mathcal{A}_n) \neq \varnothing$ iff t has a closed instance in $L(\tilde{A})$.

4) Use question 3 and exercise 4 of TD1.

Exercise 4 : Stability vs Recognizability

We can see the set of runs of an NFTA $\mathcal{A} = (Q, \mathcal{F}, Q_f, \Delta)$ as a tree language on $\mathcal{F} \times Q = \{(f, q)(n) \mid f(n) \in \mathcal{F},\ q \in Q\}$ as the smallest set $Run(\mathcal{A})$ included in $T(\mathcal{F} \times Q)$ such that :

- if $a \to q \in \Delta$, then $(a, q) \in Run(\mathcal{A})$
- if $f(q_1, ..., q_n) \to q \in \Delta$ and $t_1, ..., t_n \in Run(\mathcal{A})$ with $t_i(\epsilon) = (\_, q_i)$ then $(f, q)(t_1, ..., t_n) \in Run(\mathcal{A})$.

Then the set of accepting runs can be seen as $Acc(\mathcal{A}) = \{t \in Run(\mathcal{A}) \mid t(\epsilon) = (\_, q),\ q \in Q_f\}$.

1) Prove that $Acc(\mathcal{A})$ is in the smallest class **Stab** of sets which contains all the $T(\mathcal{F})$ for any finite ranked set $\mathcal{F}$ and which is stable by image of linear morphisms and inverse image of morphisms. For example, you should be able to prove that $Acc(\mathcal{A}) = \beta^{-1}(\gamma(\delta^{-1}(T(\mathcal{F}'))))$ where γ is linear.
2) Deduce that **Stab** = **Rec**.

Solution:

1) We define β, γ and δ this way :
— $\delta : T(\mathcal{F}_\mathcal{A}) \longrightarrow T(\mathcal{F}_\perp)$ where $\mathcal{F}_\perp = \mathcal{F} \cup \{q(1) \mid q \in F\} \cup \{\perp(0)\}$ and $\mathcal{F}_\mathcal{A} = \{(f, q_1, ..., q_n, q)(n) \mid f(n) \in \mathcal{F}, q, q_i \in q\} \cup \{q(1) \mid q \in Q\}$ such that
 ⋆ $q(x) \mapsto q(x)$ if $q \in F$
 ⋆ $q(x) \mapsto \perp$ if $q \notin F$
 ⋆ $(f, q_1, ..., q_n, q)(x_1, ...x_n) \mapsto f(x_1, ..., x_n)$ if $f(q_1, ..., q_n) \longrightarrow q \in \Delta$
 ⋆ $(f, q_1, ..., q_n, q)(x_1, ...x_n) \mapsto \perp$ else
— $\gamma : T(\mathcal{F}_\mathcal{A}) \longrightarrow T(\mathcal{F}_Q)$ linear where $T(\mathcal{F}_Q) = \mathcal{F} \cup \{q(1) \mid q \in Q\}$ such that :
 ⋆⋆ $q(x) \mapsto q(x)$
 ⋆⋆ $(f, q_1, ..., q_n, q)(x_1, ..., x_n) \mapsto q(f(q_1(x_1), ..., q_n(x_n)))$
— $\beta : T(\mathcal{F} \times Q) \longrightarrow T(\mathcal{F}_Q)$ such that :
 ⋆⋆⋆ $(f, q)(x_1, ..., x_n) \mapsto q(q(f(x_1, ..., x_n)))$
Then $Acc(\mathcal{A}) = \beta^{-1}(\gamma(\delta^{-1}(T(\mathcal{F}_\perp \setminus \perp))))$.
2) **Stab** $\subseteq$ **Rec** : **Rec** is stable under inverse image, linear image and contains all the $T(\mathcal{F})$.
Rec $\subseteq$ **Stab** : Let $L \in$ **Rec** and $\mathcal{A}$ a NFTA recognizing L. Define $\alpha : T(\mathcal{F} \times Q) \longrightarrow T(\mathcal{F})$ linear such that :
⋆⋆⋆⋆ $(f, q)(x_1, ..., x_n) \mapsto f(x_1, ..., x_n)$
Then $L = \alpha(Acc(\mathcal{A}))$ and by 1), $L \in$ **Stab**.