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 \( \text{Red}(\mathcal{E}) = \{ C[t\sigma] \mid C \in \mathcal{C}(\mathcal{F}), t \in \mathcal{E}, \sigma \text{ ground substitution} \} \) is recognizable.

2) Prove that if \( \mathcal{E} \) contains only ground terms, then one 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 \( \mathcal{E} \).

Solution:

1) Do the case where \( \mathcal{E} \) is a singleton \{\( t \)\}, \( t \) linear (the general case can be deduced by finite union). \( \text{Red}(\{t\}) \) is recognized by the following NFTA: \( Q = \{ q_\perp \} \cup \text{Pos}(t) \), \( F = \{ \epsilon \} \) and \( \Delta = \):  
   - \( f(q_1,\ldots,q_n) \rightarrow q_\perp \) for all \( f \in \mathcal{F} \), \( q_1,\ldots,q_n \in Q \)
   - \( q_\perp \rightarrow p \) for all \( p \in \text{Pos}(t) \) such that \( t(p) \) is a variable
   - \( f(p_1,\ldots,p,n) \rightarrow p \) if \( t(p) = f \)
   - \( f(q_1,\ldots,q_n) \rightarrow \epsilon \) for all \( f \in \mathcal{F} \) and \( q_1,\ldots,q_n \in Q \) such that there exists \( i \in \{1,\ldots,n\} \) such that \( q_i = \epsilon \)

2) Let \( S(t) \) be the set of all subterms of a term in \( \mathcal{E} \). Then the following DFTA \( A \) works: \( Q = \{ q_t \mid t \in S(t) \} \cup \{ q_\perp,q_\top \} \), \( F = \{ q_\top \} \) and \( \Delta =, \forall f \in \mathcal{F} \):  
   - \( f(q_{t_1},\ldots,q_{t_n}) \rightarrow q_{f(t_{t_1},\ldots,t_n)} \) if \( f(t_{t_1},\ldots,t_n) \in S(t) \setminus \mathcal{E} \)
   - \( f(q_{t_1},\ldots,q_{t_n}) \rightarrow q_\top \) if \( f(t_{t_1},\ldots,t_n) \in \mathcal{E} \)
   - \( f(q_{t_1},\ldots,q_{t_n}) \rightarrow q_\perp \) else
   - \( f(q_{t_1},\ldots,q_{t_n}) \rightarrow q_\top \) if there is at least one \( q_i = q_\top \)
   - \( f(q_{t_1},\ldots,q_{t_n}) \rightarrow q_\perp \) else

We will, for once, and as you should at least for the first few questions of an exam, formally prove that this automaton recognizes the expected language.

We first prove by induction on the size of the terms, that \( \forall t \in S(t) \setminus \mathcal{E} \), \( L(q_a) = t \).

- If \( t = a/0 \in S(t) \setminus \mathcal{E} \), then, the only rule which can produce \( q_a \) is \( a \rightarrow q_a \), and we do have \( L(q_a) = a \).
- If \( t = f(t_{t_1},\ldots,t_{t_n}) \in S(t) \setminus \mathcal{E} \), the interesting rule is then \( f(q_{t_{t_1}},\ldots,q_{t_{t_n}}) \rightarrow q_{f(t_{t_1},\ldots,t_{t_n})} \). Thus, \( L(q_{f(t_{t_1},\ldots,t_{t_n})}) = \{ f(x_{t_{t_1}},\ldots,x_{t_{t_n}}) \mid \forall i \leq n, x_i \in L(q_{t_{t_i}}) \} \). By the induction hypothesis, we have \( \forall 1 \leq i \leq n, L(q_{t_i}) = t_i \). Thus, \( L(q_{f(t_{t_1},\ldots,t_{t_n})}) = f(t_{t_1},\ldots,t_{t_n}) \).

Now, we may prove that by induction on the size \( n \) of the terms that \( L(q_\perp) \supset T^{<n}(\mathcal{F},\mathcal{X}) \setminus (\text{Red}(\mathcal{E}) \cup S(t)) \land L(q_\top) \supset \text{Red}^{<n}(\mathcal{E}) \) (we denote with \( \text{Red}^{<n}(\mathcal{E}) \) all the terms of \( \mathcal{L} \) of size at most \( n \)).

- If \( t = a/0 \notin (\text{Red}(\mathcal{E}) \cup S(t)) \), then we have a transition \( a \rightarrow q_\perp \).
- If \( t = a/0 \in \mathcal{E} \) then we have a transition \( a \rightarrow q_\top \).

We do have our property for \( n = 0 \).

- If \( t = f(t_{t_1},\ldots,t_{t_n}) \in \mathcal{E} \), we have obtained previously that \( L(q_t) = t \).
- If \( t = f(t_{t_1},\ldots,t_{t_n}) \in \mathcal{E} \), the only interesting rule is \( f(q_{t_{t_1}},\ldots,q_{t_{t_n}}) \rightarrow q_\top \). As \( f(t_{t_1},\ldots,t_{t_n}) \in \mathcal{E} \), \( t_{t_1},\ldots,t_{t_n} \in \mathcal{E} \), we obtained that \( L(t_i) = q_{t_i} \), and we do have \( t \in L(q_\top) \).
Exercise 2: Decisions problems

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

Instance: $t$ a term in $T(F,X)$ and $A$ a NFTA

Question: Is there at least one ground instance of $t$ accepted by $A$?

1) Suppose that $t$ is linear. Prove that (GII) is P-complete.
2) Suppose that $A$ is deterministic. Prove that (GII) is NP-complete.
3) Prove that (GII) is EXPTIME-complete.

Solution:

1) in P: use a construction similar to exercise 1, intersect with $A$ and test the non-emptiness.

P-hard: testing the emptiness of $A$ is equivalent to testing (GII) on $A$ and a variable.

2) in NP: guess for each variable an accessible state of $A$ and verify that you can complete this to an accepting run by running the automata.

NP-hard: We reduce (SAT) : let $F = \{¬(1), ∨(2), ∧(2), ⊥(0), ⊤(0)\}$ and $A_{SAT}$ the DFTA with $Q = \{ q_T, q_⊥ \}$, $F = \{ q_T \}$ and $Δ =$

* $⊥ \rightarrow q_⊥$
* $⊤ \rightarrow q_T$
* $¬(q_α) \rightarrow q_¬α$
* $∨(q_α, q_β) \rightarrow q_{α∨β}$
* $∧(q_α, q_β) \rightarrow q_{α∧β}$

The language of $A_{SAT}$ is the set of closed valid formulæ.

Let $φ$ a CNF formula, $φ = \bigwedge_{i=1}^n c_i$ 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 $c_i = ¬x_j$, $t_{c_i} = ¬(x_j)$
- if $c_i = x_j ∨ c'_i$, $t_{c_i} = ∨(x_j, t'_{c_i})$
- if $c_i = ¬x_j ∨ c'_i$, $t_{c_i} = ∨(¬(x_j), t'_{c_i})$

Then $t_φ = ∨(t_{c_1}, ∨(t_{c_2}, ..., ∨(t_{c_n−1}, t_{c_n})...))$. $φ$ is satisfiable if and only if a closed instance of $t_φ$ is recognized by $A_{SAT}$.

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 \( L(A_{q_1}) \cap ... \cap L(A_{q_n}) \) is non empty where \( A_q \) is the NFTA obtained from \( A \) by changing the set of final states to \( \{q\} \) (in \( P \)).

EXP-hard : We reduce intersection non-emptiness : let \((A_k = (Q_k, F_k, I_k, \Delta_k))_{k \in \{1, ..., n\}}\) 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 :

1. \( F' = F \cup \{h(n)\} \)
2. \( t = h(x, ..., x) \)
3. \( \tilde{A} = (\bigsqcup Q_k \sqcup \{q_0\}, F', \{q_0\}, \Delta' \sqcup \bigsqcup \Delta_k) \) where

\[
\Delta' = \{q_0(h(x_1, ..., x_n)) \rightarrow h(q_1(x_1), ..., q_n(x_n)) \mid \text{for } q_k \in I_k\}
\]

Then \( L(A_1) \cap ... \cap L(A_n) \neq \emptyset \) iff \( t \) has a closed instance in \( L(\tilde{A}) \).

4) Use question 3 and exercise 4 of TD1.

***

Homework for next week : Direct images of an homomorphism

Let \( F = \{f/2, g/1, a\} \) and \( F' = \{f'/2, g/1, a\} \). Let us consider the tree homomorphism \( h \) determined by \( h_F \) defined by : \( h_F(f) = f'(x_1, x_2), h_F(g) = f'(x_1, x_1), \) and \( h_F(a) = a \).

1. Is \( h(T(F)) \) recognizable ?
2. Let \( L_1 = \{g^i(a) \mid i \geq 0\} \), then \( L_1 \) is a recognizable tree language, is \( h(L_1) \) recognizable ?