Automates d’arbre

Homework - correction

**Theorem.** If \( L \subseteq T(\Sigma) \) is recognized by a NFHA then \( \text{fcns}(L) \) is recognized by a NFTA.

Let \( A = \langle \Sigma, Q, \Delta \rangle \) be a NFHA recognizing \( L \). We can assume that \( A \) is normalized i.e. for every \( a \in \Sigma \) and \( q \in Q \), there is exactly one transition of the form \( a(L_{a,q}) \rightarrow q \) in \( \Delta \). Let \( B = \langle P_{a,q}, Q, p_{a,q}, F_{a,q}, \Delta \rangle \) be a FWA (on the alphabet \( Q \)) recognizing \( L_{a,q} \). Note \( P = \bigcup_{a \in \Sigma, q \in Q} P_{a,q} \)

\[
F = \bigcup_{a \in \Sigma, q \in Q} F_{a,q}.
\]
Define the NFTA \( A' = \langle Q', \mathcal{F}_{\text{fcns}}', Q', \Delta' \rangle \) where :

- \( Q' = \{ q_f, q_\# \} \cup P \)
- \( Q'_f = \{ q_f \} \)
- \( \mathcal{F}_{\text{fcns}}' = \{ \{ a(2) \mid a \in \Sigma \} \cup \{ # (0) \} \} \)
- \( \Delta' \) contains :
  1. \( # \rightarrow p \) for all \( p \in F \)
  2. \( # \rightarrow q_\# \)
  3. \( b(p_1, p_2) \rightarrow p_3 \) if there are \( a \in \Sigma \) and \( q, q' \in Q \) such that \( p_1 = p_{b,q}^0 \) and \( (p_3, q', p_2) \in \Delta_{a,q} \)
  4. \( a(p, q_\#) \rightarrow q_f \) if there is \( q \in Q_f \) such that \( p = p_{a,q}^0 \)

We will prove the following :

**Invariant.** For every edge \( t_1...t_n \ (n \geq 0) \) and every \( p \in P_{a,q} \) for some \( a \in \Sigma \) and \( q \in Q \), \( \text{fcns}(t_1...t_n) \rightarrow_{A'} p \) if and only if there are \( q_1, ..., q_n \in Q \) such that for all \( i \in \{ 1, ..., n \} \), \( t_i \rightarrow_{A} q_i \) and \( q_1...q_n \) has a run in \( B_{a,q} \) starting from \( p \) and ending in a final state.

From this invariant, let us prove that \( L(A') = \text{fcns}(L) \).

First, let \( t \in L(A') \). Notice that the last rule of an accepting run for \( t \) can only be one of the form 4) \( a(p_{a,q}^0, q_\#) \rightarrow q_f \) for some \( q \in Q_f \) and then \( t = a(t', #) \) for some \( t' \rightarrow_{A'} p_{a,q}^0 \). As \( \text{fcns} \) is a bijection, there is an edge \( t_1...t_n \) such that \( t' = \text{fcns}(t_1...t_n) \) and so \( t = \text{fcns}(a(t_1, ..., t_n)) \).

By the invariant, there are \( q_1, ..., q_n \in Q \) such that for all \( i \), \( t_i \rightarrow_{A} q_i \) and \( q_1...q_n \) has an accepting run in \( B_{a,q} \).

Conversely, let \( t = a(t_1, ..., t_n) \in L \). So necessarily, there is \( q \in Q_f \), \( q_1...q_n \in L_{a,q} \). By the invariant, \( f \text{cns}(t_1...t_n) \rightarrow_{A'} p_{a,q}^0 \) and by applying the rule 4) and the rule 2), \( f \text{cns}(t) = a(f \text{cns}(t_1...t_n), #) \rightarrow_{A'} a(p_{a,q}^0, #) \rightarrow_{A} a(p_{a,q}, q_\#) \rightarrow_{A'} q_f \) i.e. \( f \text{cns}(t) \in L(A') \).

**Proof of the invariant :**

Let us prove it by induction on the size of the hedge \( t_1...t_n \). Let \( a \in \Sigma, q \in Q \) and \( p \in P_{a,q} \).

**case** \( n = 0 \) : \( \# \rightarrow_{A'} p \) if \( p \in F_{a,q} \) iff \( \epsilon \) has a run from \( p \) to a final state in \( B_{a,q} \).

**case** \( n \geq 1 \) : in this case, \( t_1 = b(t'_1, ..., t'_m) \) and \( f \text{cns}(t_1...t_n) = b(f \text{cns}(t'_1...t'_m), f \text{cns}(t_2...t_n)) \).

Observing that the only possible last rule is one of the form 3), \( f \text{cns}(t_1...t_n) \rightarrow_{A'} p \) iff there are \( q' \in Q \) and \( p' \in P_{a,q} \) such that \( f \text{cns}(t'_1...t'_m) \rightarrow_{A'} p_{b,q}^0 \), \( (p, q', p') \in \Delta_{a,q} \) and \( f \text{cns}(t_2...t_n) \rightarrow_{A'} p' \). By induction hypothesis, this holds iff there are \( q' \in Q \), \( p' \in P_{a,q} \) and \( q'_1, ..., q'_m \in Q \) such that for all \( i \), \( t'_i \rightarrow_{A} q'_i \) and \( q'_1...q'_m \) has an accepting run in \( B_{b,q'} \), \( (p, q', p') \in \Delta_{a,q} \) and \( f \text{cns}(t_2...t_n) \rightarrow_{A'} p' \). By induction hypothesis again, this
holds iff there are \( q' \in Q, p' \in P_{a,q} \) and \( q_2, ..., q_n \in Q \) such that \( t_1 \rightarrow_A^* q', (p, q', p') \in \Delta_{a,q} \), for all \( i \geq 2 \) \( t_i \rightarrow_A^* q_i \) and \( q_2...q_n \) has a run from \( p' \) to a final state in \( B_{a,q} \) iff there are \( q_1, ..., q_n \in Q \) such that for all \( i \in \{1, ..., n\} \), \( t_i \rightarrow_A^* q_i \) and \( q_1...q_n \) has a run in \( B_{a,q} \) starting from \( p \) and to a final state in \( B_{a,q} \).

\[Q.E.D.\]

**Theorem.** If \( K \subseteq T(\mathcal{F}^\Sigma_{\text{fcns}}) \) is recognized by a NFTA then \( \text{fcns}^{-1}(K) \cap T(\Sigma) \) is recognized by a NFTA.

Let \( \mathcal{A} = (Q, F^\Sigma_{\text{fcns}}, Q_f, \Delta) \) be a NFTA recognizing \( K \). Define the NFHA \( \mathcal{A}' = (Q', \Sigma, Q'_{f}, \Delta') \) where:

- \( Q' = \Delta \cap Q^2 \times \Sigma \times Q = \{a(p, p') \rightarrow p'' \in \Delta\} \)
- \( Q'_{f} = \{a(p, p') \rightarrow p'' \mid p'' \in Q_f \land \# \rightarrow p' \in \Delta\} \)
- \( \Delta' \) contains the transitions \( a(L_r) \rightarrow r \) for \( r \in \Delta \) of the form \( a(p, p') \rightarrow p'' \) where \( L_r \) is recognized by the FWA \( B_r = (Q, Q', p, \bar{F}, \bar{\Delta}) \) on the alphabet \( Q' \) with:
  * \( \bar{F} = \{ q \mid \# \rightarrow q \in \Delta \} \)
  * \( \bar{\Delta} = \{ (q'', r, q') \in Q \times Q' \times Q \mid r \in \Delta \text{ of the form } b(q, q') \rightarrow q'' \} \)

We will prove the following:

**Invariant.** For every hedge \( t_1...t_n \) and every \( p \in Q \),

\[ \text{fcns}(t_1...t_n) \rightarrow_A^* p \text{ if and only if there are } r_1, ..., r_n \in Q' \text{ such that for all } i \in \{1, ..., n\}, t_i \rightarrow_A^* r_i \text{ and } r_1...r_n \text{ has a run from } p \text{ to a state in } \bar{F} \text{ using transitions in } \bar{\Delta}. \]

From this invariant, let us prove that \( L(\mathcal{A}') = T(\Sigma) \cap \text{fcns}^{-1}(K) \).

First, let \( t \in L(\mathcal{A}') \), \( t = a(t_1, ..., t_n) \). It is clear that \( t \in T(\Sigma) \). Then, there are \( r \in Q' \) of the form \( a(p, p') \rightarrow p'' \) and \( r_1, ..., r_n \in Q \) such that for all \( i \in \{1, ..., n\} \), \( t_i \rightarrow_A^* r_i \) and \( r_1...r_n \in L_r \) i.e. \( p'' \in Q_f, \# \rightarrow p' \in \Delta \) and \( r_1...r_n \) has a run from \( p \) to a state in \( \bar{F} \). So, by the invariant, \( \text{fcns}(t_1...t_n) \rightarrow_A^* p \). Then, \( \text{fcns}(a(t_1, ..., t_n)) = \text{fcns}(t_1...t_n, \#) \rightarrow_A^* \text{fcns}(a(p, \#) \rightarrow_A a(p, p') \rightarrow_A p'' \in Q_f \) i.e. \( \text{fcns}(t) \in K \).

Conversely, let \( t \in T(\Sigma) \cap \text{fcns}^{-1}(K) \), \( t = a(t_1, ..., t_n) \). Then, \( \text{fcns}(t) = a(\text{fcns}(t_1...t_n), \#) \in K \).

That means there are \( q \in Q_f, q', q'' \in Q \) such that \( \text{fcns}(t_1...t_n) \rightarrow_A q', \# \rightarrow q'' \in \Delta \) and \( r = a(q, q') \rightarrow q'' \in \Delta \). In particular, \( r \in Q' \) by the invariant, there are \( r_1, ..., r_n \in Q' \) such that for all \( i \in \{1, ..., n\} \), \( t_i \rightarrow_A^* r_i \) and \( r_1...r_n \in L_r \). So, \( a(t_1, ..., t_n) \rightarrow_A^* a(r_1, ..., r_n) \rightarrow_A^* r \in Q_f \) i.e. \( t \in L(\mathcal{A}') \).

**Proof of the invariant :**

Let us prove it by induction on the size of the hedge \( t_1...t_n \). Let \( p \in Q \).

- **Case** \( n = 0 \): \( \# \rightarrow_A^* p \) if \( p \in \bar{F} \) iff \( \epsilon \) has a run from \( p \) to a state in \( \bar{F} \) using transitions in \( \bar{\Delta} \).
- **Case** \( n \geq 1 \): in this case, \( t_1 = b(t'_1, ..., t'_{m}) \) and \( \text{fcns}(t_1...t_n) = b(\text{fcns}(t_1...t_{m}), \text{fcns}(t_2...t_{n})) \).

\( \text{fcns}(t_1...t_n) \rightarrow_A^* p \) iff there are \( p', p'' \in Q \) such that \( \text{fcns}(t'_1...t'_{m}) \rightarrow_A^* p', \text{fcns}(t_2...t_n) \rightarrow_A^* p'' \) and \( r_1 = b(p', p'') \rightarrow p \in \Delta \).

By the induction hypothesis, this holds iff there are \( p', p'' \in Q \) and \( r'_1, ..., r'_{m} \in Q' \) such that for all \( i \in \{1, ..., m\} \), \( t'_i \rightarrow_A^* r'_i \) and \( r'_1...r'_m \) has an accepting run in \( B_{a,q} \), \( \text{fcns}(t_2...t_{n}) \rightarrow_A^* p'' \) if there are \( p', p'' \in Q \) such that \( t_1 \rightarrow_A^* r_1 \) and \( (p, r_1, p'') \in \Delta \). Again by the induction hypothesis, this holds iff there are \( p', p'' \in Q \) and \( r_2, ..., r_n \in Q' \) such that \( t_1 \rightarrow_A^* r_1, (p, r_1, p'') \in \Delta \), for all \( i \in \{2, ..., n\} \), \( t_i \rightarrow_A^* r_i \) and \( r_2...r_n \) has a run from \( p'' \) to a state in \( \bar{F} \) iff there are \( r_1, ..., r_n \in Q' \) such that for all \( i \in \{1, ..., n\} \), \( t_i \rightarrow_A^* r_i \) and \( r_1...r_n \) has a run from \( p \) to a state in \( \bar{F} \) using transitions in \( \bar{\Delta} \).

\[Q.E.D.\]