Syntactic ∞-Tree Languages

Symeon Bozapalidis

Department of Mathematics,
Aristotle University of Thessaloniki
54124, Thessaloniki, Greece.

Abstract. The class of languages obtained by solving non-deterministic rational program schemes is shown to be closed under linear tree homomorphisms and inverse alphabetic tree homomorphisms. Moreover, for a given such language $\mathcal{F}$ and an infinite regular tree R we can decide whether or not $\mathcal{F}$ is finite and $R \in \mathcal{F}$. These results are also valid for Buchi tree languages.

1 INTRODUCTION

In 1977 Arnold and Nivat [AN1] showed that the set of trees computed by a non-deterministic recursive program scheme (NRvePS) is just a component of the greatest solution of the monotonic operator canonically associated with this scheme. The various semantics of such schemes have been investigated in a series of papers (cf. [Na], [ANN], [Po], [AN2], [AN3]). Here we try a linguistic study of languages of infinite trees determined by non-deterministic rational program schemes which are the first order variant of NRvePS's above.

Let Γ be a ranked alphabet and $X_n = \{x_1, \ldots, x_n\}$ a set of variables. As usual, $T_\Gamma(X_n)$, $T_\Gamma^\infty(X_n)$ stand for the sets of finite, infinite trees respectively over Γ and X_n.

A non-deterministic rational program scheme (NRPS) is a system $P = \langle \Gamma, X_n, \Sigma, x_1 \rangle$ where Γ, X_n are as above, x_1 is the starting variable and Σ is a system of equations of the form

$$x_i = L_i, \quad 1 \le i \le n$$

with $L_1, \ldots, L_n$ finite subsets of $T_\Gamma(X_n)$.

The syntactic ∞-tree languages defined by P are the x_1-component of the greatest OI-solution of the above system. Such languages can also be obtained as ∞-behaviours of top down tree automata using the D3-acceptance mode in the terminology of [NS].

Notice that the NRPS's and the rational logic programs of Kowalski (cf. [Ko]) are syntactically equivalent objects (cf. [NS]).

Two kinds of results are established:

- Decision problems.
  1. For a given syntactic language $\mathcal{F}$ we can decide whether $\mathcal{F}$ is finite or not. If it is finite, then $\mathcal{F}$ consists exclusively of regular trees.

2. Given a syntactic language $\mathcal{F} \subseteq T_{\Gamma}^{\infty}$ and a regular infinite tree R, we van decide whether or not $R \in \mathcal{F}$.

– Closure properties.

3. The class $STL(\Gamma)$ of syntactic tree languages over the alphabet Γ, is the smallest class containing the finite sets of finite trees and closed under OI-substitution and OI-STAR.
4. $STL(\Gamma)$ is closed under linear tree homomorphisms and inverse alphabetic tree homomorphisms.
5. The branches of a syntactic tree language, form a syntactic language, as well.

The techniques displayed to achieve the previous results are also valid for the case of Büchi tree automata. Hence the statements 1) - 5) above hold for Büchi tree languages also.

An excellent survey on infinite trees can be found in [Th].

2 PRELIMINARIES

Let $\Gamma = \bigcup_{k \geq 0} \Gamma_k$ be a ranked alphabet and $X = \{x_1, x_2, \ldots\}$ a set of variables. We put $X_n = \{x_1, \ldots, x_n\}$, $X_0 = \varnothing$. Also let Ω be a 0-ranked symbol and set $\mathbb{N}_0 = \{1, 2, ...\}$.

A *tree* is a partial function $T : \mathbb{N}_0^* \to \Gamma \cup \{\Omega\} \cup X_n$ whose domain is prefix closed and has the additional properties:

- for $a \in \mathbb{N}_0^*$, if $T(a) \in \Gamma_0 \cup \{\Omega\} \cup X_n$ then $T(ai)$ is not defined for all $i \in \mathbb{N}_0$.

- for $a \in \mathbb{N}_0^*$, if $T(a) \in \Gamma_k$ $(k \geq 1)$ then $T(ai)$ is defined iff $1 \leq i \leq k$.

Denote by $T_{\Gamma,\Omega}^{\infty}(X_n)$ the so obtained set and by $T_{\Gamma,\Omega}(X_n)$ the set of all trees with finite domain. For $T, T' \in T_{\Gamma,\Omega}^{\infty}(X_n)$ we write $T \leq_{\Omega} T'$ whenever $Dom(T) \subseteq Dom(T')$ and for any $a \in Dom(T)$ such that $T(a) \neq \Omega$ we have $T(a) = T'(a)$. $(T_{\Gamma,\Omega}^{\infty}(X_n), \leq_{\Omega})$ is an ω-complete set.

The basic operation on trees is substitution. Let $T \in T_{\Gamma,\Omega}^{\infty}(X_n)$ and put

$$V_i(T) = \{a \mid a \in \mathbb{N}_0^*,\, T(a) = x_i\},\ 1 \leq i \leq n$$

be the set of all vertices of T labelled by x_i $(1 \leq i \leq n)$ and consider families $\overrightarrow{S}_i = (S_a^i)_{a \in V_i(T)}$ of trees in $T_{\Gamma,\Omega}^{\infty}(X_n)$ $(1 \leq i \leq n)$. Then

$$T\left[\overrightarrow{S}_1/x_1, \ldots, \overrightarrow{S}_n/x_n\right] \text{ or just } T\left[\overrightarrow{S}_1, \ldots, \overrightarrow{S}_n\right]$$

is the tree whose domain is

$$(Dom\,(T) - V\,(T)) \cup \left(\bigcup_{\substack{a \in V_i(T) \\ 1 \leq i \leq n}} \alpha Dom(S_a^i)\right)$$

with $V(T) = \bigcup_{1 \le i \le n} V_i(T)$. Moreover for all $\beta \in \mathbb{N}_0^*$ we have

$$T\left[\overrightarrow{S}_1, \dots, \overrightarrow{S}_n\right](\beta) = T(\beta), \text{ if} \beta \in Dom(T) - V(T)$$
$$= S_a^i(\gamma), \text{ if } \beta = \alpha\gamma,\ \alpha \in V_i(T),\ \gamma \in Dom(S_\alpha).$$

Associativity holds for the above vectorial tree substitution.

Now, given languages of infinite trees $\mathcal{L}, \mathcal{A}_1, \dots, \mathcal{A}_n \subseteq T^{\infty}_{\Gamma,\Omega}(X_n)$ we define

$$\mathcal{L}\left[\mathcal{A}_1, \dots, \mathcal{A}_n\right]_{OI} = \left\{T\left[\overrightarrow{S}_i, \dots, \overrightarrow{S}_n\right] \mid T \in \mathcal{L},\ \overrightarrow{S}_i \in \mathcal{A}_i^{V_i(T)},\ 1 \le i \le n\right\}.$$

A straightforward argument shows that

Proposition 1. *OI-substitution of infinite tree languages is associative.*

3 SYNTACTIC LANGUAGES

Let Γ be a finite ranked alphabet and $X_n = \{x_1, \dots, x_n\}$ a set of variables and consider the system

$$(\Sigma) \quad x_i = L_i \ \ , \ \ L_i \subseteq T_\Gamma(X_n) \ \ , \ \ 1 \le i \le n.$$

We say that $(\mathcal{A}_1, \dots, \mathcal{A}_n) \in \mathcal{P}(T^{\infty}_{\Gamma})^n$ is an *OI-solution* of (Σ) whenever

$$\mathcal{A}_i = L_i\left[\mathcal{A}_1, \dots, \mathcal{A}_n\right]_{OI} \ \ , \ \ 1 \le i \le n.$$

We say that the sequence $\mathbf{t} = (t_k)$ of trees in $T_\Gamma(X_n)$ is an *x_i-th OI-expansion* of (Σ) whenever

$$t_0 \in L_i,\ t_{k+1} \in t_k\left[L_1, \dots, L_n\right]_{OI}.$$

$EXP_{OI}(\Sigma, x_i)$ stands for the so defined set.

We put

$$\mathbf{t}_\Omega = (t_{k,\Omega})$$

where $t_{k,\Omega} = t_k\left[\Omega/x_1, \dots, \Omega/x_n\right]$. $\mathbf{t}_\Omega$ is an increasing sequence in $T^{\infty}_{\Gamma,\Omega}$ and thus $\sup \mathbf{t}_\Omega$ exists and is denoted by $\widehat{\mathbf{t}}$, i.e.

$$\widehat{\mathbf{t}} = \sup \mathbf{t}_\Omega.$$

Theorem 1. *The n-tuple*

$$(s) \quad (\mathcal{F}(\Sigma, x_1), \dots, \mathcal{F}(\Sigma, x_n)) \in \mathcal{P}(T^{\infty}_{\Gamma})^n,$$

where $\mathcal{F}(\Sigma, x_i) = \left\{\widehat{\mathbf{t}} \mid \mathbf{t} \in EXP_{OI}(\Sigma, x_i)\right\}$ *is the greatest, by inclusion, OI-solution of* (Σ).

A *non deterministic rational program scheme* (NRPS) is a 4-tuple $P = \langle \Gamma, X_n, \Sigma, x_1 \rangle$, where Γ, X_n are as above, (Σ) is a system whose all right-hand side members are finite subsets of $T_\Gamma(X_n)$ and x_1 is the starting variable. The language computed by P is the first component of the greatest OI-solution of (Σ).

We say that $F \subseteq T_\Gamma^\infty$ is a *syntactic* language whenever it is computed by a NRPS P. $STL\,(\Gamma)$ denotes the class of syntactic languages over Γ.

Example 1. The whole T_Γ^∞ is a syntactic language since it is the greatest OI-solution of the equation

$$x = \{f\,(x, \ldots, x) \mid f \in \Gamma_k, k \geq 0\}\,.$$

The presence of variables in the right-hand side members of a system does not affect the formalization of its maximal solution. Indeed, for a given system

$$(\Sigma) \quad x_i = L_i \quad , \quad 1 \leq i \leq n$$

with $L_1, \ldots, L_n$ finite subsets of $T_\Gamma(X_n)$, let us write $i \to j$ whenever $x_j \in L_i$. Then

Lemma 1. *The systems* (Σ) *above and*

$$(\Sigma') \quad x_i = (L_i - X_n) \cup \left(\bigcup_{i \xrightarrow{k} j} (L_j - X_n) \right)$$

have the same greatest OI-solution.

The NRPS's $P = \langle \Gamma, X_n, \Sigma, x_1 \rangle$ and $P' = \langle \Gamma, X_m, \Sigma', x_1 \rangle$ are said to be *equivalent* whenever they compute the same language: $\mathcal{F}(\Sigma, x_1) = \mathcal{F}(\Sigma', x_1)$.

Proposition 2. *Given a NRPS* $P = \langle \Gamma, X_n, \Sigma, x_1 \rangle$ *we can effectively construct an equivalent* $P' = \langle \Gamma, X_m, \Sigma', x_1 \rangle$ *satisfying the condition*

$$L'_j \subseteq T_\Gamma(X_m) - X_m \quad , \quad j = 1, \ldots, m.$$

The graph associated with a system

$$(\Sigma) \quad x_i = L_i \quad , \quad L_i \subseteq T_\Gamma(X_n) \quad , \quad 1 \leq i \leq n$$

denoted by $Gr(\Sigma)$ has $X_n \cup \{\#\}$ as set of vertices while we draw an edge $x_i \to x_j$ (resp. $x_i \to \#$) whenever there is a tree $t \in L_i$ in which the variable x_i occurs (resp. $t \in L_i \cap T_\Gamma$).

It is not hard to see that the language computed by $P = \langle \Gamma, X_n, \Sigma, x_1 \rangle$ is non-empty if and only if

$$(c) \begin{cases} \text{for every path} \\ x_1 \to x_{i_1} \to \ldots \to x_{i_p} \to x_i \\ \text{in } Gr(\Sigma) \text{ with } i_1, \ldots, i_p, \; i \text{ distinct elements of} \\ \{2, 3, \ldots, n\} \text{ it holds } L_i \neq \emptyset. \end{cases}$$

From now on, without any loss of generality, we may deal with NRPS $P = \langle \Gamma, X_n, \Sigma, x_1 \rangle$ whose corresponding system

$$(\Sigma) \quad x_i = L_i \;\; , \;\; 1 \leq i \leq n$$

is such that

i) $L_1, \ldots, L_n$ are finite subsets of $T_\Gamma(X_n) - X_n$ and
ii) the condition (c) above is satisfied.

Call $P = \langle \Gamma, X_n, \Sigma, x_1 \rangle$ *deterministic* whenever all right-hand side members of (Σ) are singletons

$$(\Sigma) \quad x_i = t_i \;\; , \;\; t_i \in T_\Gamma(X_n) - X_n \;\; , \;\; 1 \leq i \leq n \; .$$

For each $i = 1, \ldots, n$ there is just one x_i-th OI-expansion, namely

$$\mathbf{s}^i = (s^i_k) \;\; , \;\; s^i_0 = t_i \;\; , \;\; s^i_{k+1} = t_i[s^1_k, \ldots, s^n_k] \; .$$

Therefore $\left(\{\widehat{\mathbf{s}}^1\}, \ldots, \{\widehat{\mathbf{s}}^n\}\right)$ is the greatest by inclusion OI-solution of (Σ) where $\left(\widehat{\mathbf{s}}^1, \ldots, \widehat{\mathbf{s}}^n\right)$ is the least with respect to $\leq_\Omega$-solution of (Σ).

We conclude

Proposition 3. *If $T \in T^\omega_\Gamma$ is a regular infinite tree, then the singleton $\{T\}$ is a syntactic language.*

Proposition 4. *Every non-empty syntactic language of T^ω_Γ contains at least one infinite regular tree.*

Theorem 2. *A finite subset $\mathcal{F} = \{T_1, \ldots, T_k\}$ of T^∞_Γ is syntactic if and only if all the trees $T_1, \ldots, T_k$ are regular.*

Proposition 5. *The finiteness problem for syntactic languages is decidable.*

4 KLEENE THEOREM

Let $\mathcal{A} \subseteq T_\Gamma^\infty(X_n)$. The greatest ∞-OI-solution of the equation $x_\kappa = \mathcal{A}$ is denoted by $\mathcal{A}^{*,OI,\kappa}$ and is the greatest (by inclusion) part of $T_\Gamma^\infty(X_n - \{x_\kappa\})$ such that

$$\mathcal{A}^{*,OI,\kappa} = \mathcal{A}\left[x_1, \ldots, x_{\kappa-1}, \mathcal{A}^{*,OI,\kappa}, x_{\kappa+1}, \ldots, x_n\right]_{OI}.$$

Next theorem confirms that the greatest solution of a system can be obtained by solving the system step by step.

Theorem 3. *Consider the system*

$$(\Sigma_n) \qquad \begin{cases} x_1 = \mathcal{A}_1 \\ \quad\vdots \\ x_n = \mathcal{A}_n \end{cases} \qquad , \quad \mathcal{A}_i \subseteq\ T_\Gamma^\infty(X_n) \quad , \quad 1 \leq i \leq n.$$

If $\mathcal{A}_n^{*,OI,n}$ is the greatest ∞-OI-solution of the last equation and $(\mathcal{F}_1, \ldots, \mathcal{F}_{n-1})$ is the greatest ∞-OI-solution of the system

$$(\Sigma_{n-1}) \begin{cases} x_1 = \mathcal{A}_1\left[x_1, \ldots, x_{n-1}, \mathcal{A}_n^{*,OI,n}\right]_{OI} \\ \quad\vdots \\ x_{n-1} = \mathcal{A}_{n-1}\left[x_1, \ldots, x_{n-1}, \mathcal{A}_n^{*,OI,n}\right]_{OI} \end{cases}$$

then

$$(s) \qquad \left(\mathcal{F}_1, \ldots, \mathcal{F}_{n-1}, \mathcal{A}_n^{*,OI,n}\left[\mathcal{F}_1, \ldots, \mathcal{F}_{n-1}\right]_{OI}\right)$$

is the greatest OI-solution of (Σ_n).

Next closure properties come by applying the above elimination procedure.

Theorem 4. *Suppose $\mathcal{A}, \mathcal{A}_1, \ldots, \mathcal{A}_n \in STL(\Gamma)$. Then $\mathcal{A}\left[\mathcal{A}_1, \ldots, \mathcal{A}_n\right]_{OI} \in STL(\Gamma)$.*

The main result of this section is next Kleene-like theorem.

Theorem 5. *$STL(\Gamma)$ is the least class of $\mathcal{P}(T_\Gamma^\infty)$ containing finite languages of T_Γ and closed under OI-substitution and OI-star-operation.*

Corollary 1. *Consider the system*

$$(\Sigma) \qquad x_i = \mathcal{L}_i \ , \quad \mathcal{L}_i \subseteq T_\Gamma^\infty(X_n) \quad , \quad 1 \leq i \leq n$$

and its OI-greatest solution $(\mathcal{F}_1, \ldots, \mathcal{F}_n)$. If $\mathcal{L}_1, \ldots, \mathcal{L}_n$ are syntactic languages, then so are $\mathcal{F}_1, \ldots, \mathcal{F}_n$.

5 TREE HOMOMORPHISMS AND SYNTACTIC LANGUAGES

First recall that for two given ranked alphabets Γ and Δ a homomorphism from Γ to Δ is simply a sequence of functions

$$h_\kappa : \Gamma_\kappa \to T_\Delta(\xi_1, \ldots, \xi_\kappa) \quad , \quad \kappa = 0, 1, ..$$

where $\Xi = \{\xi_1, \xi_2, \ldots\}$ is a set of auxiliary variables, $\Xi_\kappa = \{\xi_1, \ldots, \xi_\kappa\}, \kappa \geq 0$.

The above sequence $(h_\kappa)_{\kappa \geq 0}$ gives rise to a single function

$$h : T_\Gamma(X_n) \to T_\Delta(X_n) \quad , \quad X_n = \{x_1, \ldots, x_n\}$$

defined by the inductive formula

- $h(x_i) = x_i \quad , \quad 1 \leq i \leq n$
- $h(c) = h_0(c) \quad , \quad c \in \Gamma_0$
- $h(f(t_1, \ldots, t_\kappa)) = h_\kappa(f)[h(t_1)/\xi_1, \ldots, h(t_\kappa)/\xi_\kappa], f \in \Gamma_\kappa, t_i \in T_\Gamma(X_n), 1 \leq i \leq \kappa$.

A homomorphism h from Γ to Δ is said to be *linear* whenever for all $\kappa \geq 1$ and $f \in \Gamma_\kappa$ the tree $h_\kappa(f)$ is Ξ_κ-linear (i.e. each variable ξ_i occurs in $h_\kappa(f)$ at most once).

Every tree homomorphism $h : T_\Gamma(X_n) \to T_\Delta(X_n)$ preserves tree substitution, that is

$$h(t[s_1, \ldots, s_n]) = h(t)[h(s_1), \ldots, (s_n)]$$

for all $t, s_1, \ldots, s_n \in T_\Gamma(X_n)$, where the above substitutions take place at the variables $x_1, \ldots, x_n$.

We can extend $h : T_\Gamma(X_n) \to T_\Delta(X_n)$ to $h : T^\infty_{\Gamma,\Omega}(X_n) \to T^\infty_{\Delta,\Omega}(X_n)$ by setting

$$h(T) = \sup_{t \leq T}(t), \ T \in T^\infty_{\Gamma,\Omega}(X_n)$$

where the above ordering is the syntactic tree ordering described in Section 2.

Theorem 6. *If $h : T^\infty_{\Gamma,\Omega}(X_n) \to T^\infty_{\Delta,\Omega}(X_n)$ is a linear tree homomorphism and $(\mathcal{F}_1, \ldots, \mathcal{F}_n)$ is the greatest OI-solution of the system*

$$(\Sigma) \qquad x_i = L_i \quad , \quad L_i \subseteq T_\Gamma(X_n) \quad , \quad 1 \leq i \leq n$$

then $(h(\mathcal{F}_1), \ldots, h(\mathcal{F}_n))$ is the greatest OI-solution of the system

$$(h\Sigma) \qquad x_i = h(L_i) \quad , \quad 1 \leq i \leq n \ .$$

Consequently,

$$\mathcal{F} \in STL(\Gamma) \text{ implies } h(\mathcal{F}) \in STL(\Delta).$$

In the next section we shall display an example of a non-linear homomorphism not preserving syntactic languages.

Actually, syntactic languages are closed under the branching operator. Recall that the branching alphabet $b(\Gamma)$ associated with a ranked alphabet Γ is the monadic alphabet

$$b(\Gamma)_0 = \Gamma_0, b(\Gamma)_1 = \{[f,i] \mid f \in \Gamma_\kappa, \kappa \geq 1 \text{ and } i = 1, \ldots, \kappa\}.$$

The mapping

$$br : T_\Gamma(X_n) \to \mathcal{P}\left(T_{b(\Gamma)}(X_n)\right)$$

is defined by

- $br(a) = \{a\}, a \in \Gamma_0 \cup X_n$
- $br\left(f(t_1, \ldots, t_\kappa)\right) = [f,1]br(t_1) \cup \cdots \cup [f,\kappa]br(t_\kappa)$.

We state

Theorem 7. *If $(\mathcal{F}_1, \ldots, \mathcal{F}_n)$ is the greatest OI-solution of*

$$(\Sigma) \qquad x_i = L_i, \quad L_i \subseteq T_\Gamma(X_n), 1 \leq i \leq n$$

then $(br(\mathcal{F}_1), \ldots, br(\mathcal{F}_n))$ is the greatest solution of the system

$$br(\Sigma) \qquad x_i = br(L_i) \ \ , \ \ 1 \leq i \leq n.$$

Consequently, if $T \in T^\omega_\Gamma$ is an infinite regular tree, then $br(T)$ is a syntactic language of $T^\infty_{b(\Gamma)}$.

Corollary 2. *We can decide whether or not a regular tree $T \in T^\infty_\Gamma$ is finite or not (provided Γ has no 1-ranked symbols, $\Gamma_1 = \emptyset$).*

Theorem 8. *Inverse linear alphabetic homomorphisms preserve ∞-OI-regular languages.*

6 OI-RECOGNIZABILITY

In this section we shall relate syntactic languages with infinite behaviours of tree automata.

A top-down tree automaton over the ranked alphabet Γ is a 4-tuple

$$\mathcal{M} = (\Gamma, Q, I, \delta)$$

consisting of a finite ranked alphabet Γ of input symbols, a finite set Q of states, a set $I \subseteq Q$ of initial states and a finite set

$$\delta \subseteq \bigcup_{\kappa \geq 0} Q \times \Gamma_\kappa \times Q^\kappa$$

of transitions.

The behaviour of $\mathcal{M}$ is

$$|\mathcal{M}| = \bigcup_{q \in I} F_q$$

where $(F_q)_{q\in Q}$ is the least OI-solution of the system

$$\Sigma(\mathcal{M}) \qquad x_q = \{f\,(x_{q_1},\ldots,x_{q_\kappa}) \mid (q,f,q_1,\ldots,q_\kappa)\in\delta\}\,.$$

The behaviours of such automata coincide with the class of recognizable tree languages (cf. [GS]).

The ∞-*OI-behaviour* of a top-down automaton $\mathcal{M}$ is

$$|\mathcal{M}|^{\infty,OI} = \bigcup_{q\in I}\mathcal{F}\,(\Sigma(\mathcal{M}),x_q)$$

where $\left(\mathcal{F}\,(\Sigma(\mathcal{M}),x_q)_{q\in Q}\right)$ is the greatest OI-solution of $\Sigma(\mathcal{M})$. $\mathcal{F}\subseteq T_\Gamma^\infty$ is said to be ∞-*OI-recognizable* if it is the ∞-OI-behaviour of a top-down tree automaton $\mathcal{M}$; ∞-*OI-Rec*(Γ) stands for the so defined class.

Proposition 6. *The classes $STL(\Gamma)$ and $\infty - OI - Rec(\Gamma)$ coincide.*

Actually the above proposition states that a language is syntactic if and only if it is D3-recognizable in the terminology of [Sa], [NS].

Hence,

Corollary 3. *(cf. [Sa], [NS]) The syntactic subsets of T_Γ^∞ are closed under intersection.*

In order to render more apparent the use of tree runs in the formation of ∞-OI-behaviour of a top-down tree automaton $\mathcal{M}=(\Gamma,Q,I,\delta)$ we introduce the ranked alphabet $\mathbf{Q}$,$\mathbf{Q}_n=Q$ for $n\geq 0$, as well as the product alphabet

$$\Gamma\times\mathbf{Q},(\Gamma\times\mathbf{Q})_n=\Gamma_n\times\mathbf{Q}_n, n\geq 0.$$

Denote by $(loc(\mathcal{M})_{<\gamma,q>})_{<\gamma,q>\in\Gamma\times Q}$ the greatest ∞-OI-solution of the system

$$x_{<\gamma,q>} = \left\{<\gamma,q>\left(x_{<\gamma_1,q_{i_1}>},\ldots,x_{<\gamma_\kappa,q_{i\kappa}>}\right) \mid (q,\gamma,q_{i_1},\ldots,q_{i_\kappa})\in\delta\right\},$$

$\langle\gamma,q\rangle\in\Gamma\times\mathbf{Q}$, and set

$$loc(\mathcal{M}) = \bigcup_{<\gamma,q>\in\Gamma\times I} loc(M)_{<\gamma,q>}.$$

$loc(\mathcal{M})$ is the local language defined by $\mathcal{M}$.

The composition

$$T_\Gamma^\infty \overset{pr_\Gamma}{\longleftarrow} T_{\Gamma_x}^\infty \overset{-\cap loc(\mathcal{M})}{\longrightarrow} T_{\Gamma\times\mathbf{Q}}^\infty \overset{pr_{\mathbf{Q}}}{\longrightarrow} T_{\mathbf{Q}}^\infty$$

is by definition the relation $\overset{run_{\mathcal{M}}}{\longrightarrow}$; for $T\in T_\Gamma^\infty$,

$$run_{\mathcal{M}}(T) = pr_{\mathbf{Q}}\left(pr_\Gamma^{-1}(T)\cap loc(\mathcal{M})\right).$$

It is easily seen that

$$|\mathcal{M}|^{\infty,OI} = \{T \mid run_{\mathcal{M}}(T)\neq\emptyset\}\,.$$

Proposition 7. *For a given top-down automaton $\mathcal{M}$ the language $run_{\mathcal{M}}\left(|\mathcal{M}|^{\infty,OI}\right) \subseteq T_{\mathbf{Q}}^{\infty}$ is syntactic.*

Also, for any infinite regular tree $T \in |\mathcal{M}|^{\omega,OI}$, $run_{\mathcal{M}}(T)$ is a syntactic language of $T_{\mathbf{Q}}^{\infty}$.

Corollary 4. *Given a syntactic language $\mathcal{F} \subseteq T_{\Gamma}^{\infty}$ and a regular tree $T \in T_{\Gamma}^{\omega}$, we can decide whether $T \in \mathcal{F}$ or not.*

A useful pumping lemma can be obtained in this setup. We denote by P_{Γ}^{∞} the free monoid generated by the trees

$$f\left(T_1, \ldots, T_{i-1}, x, T_{i+1}, \ldots, T_p\right), f \in \Gamma_p, p \geq 1, T_j \in T_{\Gamma}^{\infty}, j \neq i.$$

Clearly P_{Γ}^{∞} acts on T_{Γ}^{∞} via substitution at x.

Lemma 2. *For every syntactic language $\mathcal{F} \subseteq T_{\Gamma}^{\infty}$, there is a number $N > 0$ so that each $T \in \mathcal{F} \cap T_{\Gamma}^{\omega}$ admits a decomposition*

$$T = S_1 \cdot S_2 \cdot T_1 \text{ such that } S_1, S_2 \in P_{\Gamma}^{\infty}, T_1 \in T_{\Gamma}^{\omega}, |S_2| > 0 \text{ and}$$

$$S_1\left(S_2\right)^{\kappa} T_1, S_1 \cdot S_2^{\omega} \in \mathcal{F}, \kappa = 0, 1, \ldots$$

We shall apply this lemma to show non-closure of syntactic languages under non-linear tree homomorphisms.

Example 2. Consider the ranked alphabets $\Gamma = \{f, g\}$ and $\Gamma_1 = \{f_1, g_1\}$ with $rank(f) = 2 = rank(g)$ and $rank(f_1) = 1 = rank(g_1)$ respectively.

Let $h : T_{\Gamma_1}^{\infty} \to T_{\Gamma}^{\infty}$ be the homomorphism defined by

$$h(f_1) = f(x, x), h(g_1) = g(x, x)$$

and take $\mathcal{F} = h\left(T_{\Gamma_1}^{\infty}\right)$ and $T = h(W)$ with

$$W = f_1 g_1 f_1^2 g_1^2 f_1^3 g_1^3 \cdots$$

If $\mathcal{F}$ was syntactic then by virtue of the pumping lemma above

$$T = S_1 \cdot S_2 \cdot T_1$$

with $|S_2| > 0$ and $S_1(S_2)^{\kappa} T_1 \in \mathcal{F}$ for $\kappa = 0, 1, \ldots$. But for κ large enough this is not true.

7 Büchi ∞-TREE LANGUAGES

A *Büchi tree automaton* is a system $\mathcal{M} = (\Gamma, Q, \delta, q_0, F)$ where (Γ, Q, δ, q_0) is a top-down tree automaton and $F \subseteq Q$ is the set of final states of $\mathcal{M}$.

The *behaviour* of $\mathcal{M}$, denotes by $|\mathcal{M}|^{Büchi}$, consists of all trees $T \in T_\Gamma^\infty$ such that there is a run $R \in run_\mathcal{M}(T)$ such that in every branch W of R there appear infinitely many final states.

$\mathcal{F} \subseteq T_\Gamma^\infty$ is a *Büchi language* whenever $\mathcal{F} = |\mathcal{M}|^{Büchi}$ for some automaton $\mathcal{M}$.

It is well known that

Theorem 9. *(cf. [Ta], [AN4]) $\mathcal{F} \subseteq T_\Gamma^\infty$ is a Büchi language iff it is the first component of the maximal OI-solution of a system*

$$(\Sigma) \qquad x_i = L_i,\ 1 \le i \le n$$

with $L_1, \ldots, L_n$ recognizable subsets of $T_\Gamma(X_n)$.

Given a system

$$(\Sigma_r) \qquad x_i = L_i \ \ ,\ \ 1 \le i \le n$$

with $L_1, \ldots, L_n$ recognizable subsets of $T_\Gamma(X_n) - X_n$, we can effectively construct a finite graph $Gr(\Sigma_r)$ by taking $X_n \cup \{\#\}$ as its set of vertices while we draw an edge $x_i \to x_j$ (resp. $x_i \to \#$) whenever x_j occurs in a tree $t \in L_i$ (resp. $L_i \cap T_\Gamma(X_n) \neq \emptyset$).

Since L_i is recognizable, it is decidable whether the set

$$L_i x_j^{-1} = \{\tau \mid \tau \in P_\Gamma(X_n), \tau x_j \in L_i\}$$

is empty or not.

We have the following important result:

Theorem 10. *We can decide whether a given Büchi tree language is finite or not.*

Theorem 11. *Büchi-tree languages are closed under linear tree homomorhisms and inverse alphabetic homomorphisms.*

Proposition 8. *If $\mathcal{F} \subseteq T_\Gamma^\infty$ is a Büchi language then so is $br(\mathcal{F}) \subseteq T_{b(\Gamma)}^\infty$.*

Let $\mathcal{M} = (\Gamma, Q, \delta, q_0, F)$ be a Büchi tree automaton and consider the top down automaton

$$\widehat{\mathcal{M}}_q = (\Gamma \cup X_F, Q, \widehat{\delta}, q), q \in Q$$

where $X_F = \{x_p \mid p \in F\}$ and $\widehat{\delta} = \delta \cup \{(x_p, p) \mid p \in F\}$.

Also, let $loq(\widehat{\mathcal{M}}_q)$ be the local set associated with $\widehat{\mathcal{M}}_q$ and

$$(\widehat{\Sigma}) \qquad \xi_q = loc(\widehat{\mathcal{M}}_q)\,[\xi_p/(x_p, p)]_{p \in F}\,.$$

We set

$$LOC(\mathcal{M}) = \mathcal{F}(\widehat{\Sigma}, \xi_{q_0}).$$

Then $pr_\Gamma\left(LOC(\mathcal{M})\right) = |\mathcal{M}|^{Büchi}$ and the function

$$RUN_{\mathcal{M}} : T_\Gamma^\infty \to \mathcal{P}\left(T_{\mathbf{Q}}^\infty\right),$$

$$RUN_{\mathcal{M}}(T) = pr_{\mathbf{Q}}\left(pr_\Gamma^{-1}(T) \cap LOC(\mathcal{M})\right), T \in T_\Gamma^\infty$$

preserves Büchi tree languages. Since the non emptiness problem for Büchi tree languages is decidable, we get

Theorem 12. *We can decide whether or not a regular tree $T \in T_\Gamma^\infty$ belongs to a Büchi language $\mathcal{F} \subseteq T_\Gamma^\infty$.*

References

[AD] Arnold, A. Dauchet, M., *Forêts Algébriques et Homomorphismes Inverses*, Information and Control, Vol.37., No 2, 1978, 182-196

[ANN] Arnold, A., Naudin, P., Nivat, M., *On Semantics of Non-Deterministic Recursive Program Schemes*, in Algebraic Methods in Semantics, Nivat and Reynolds Eds., Cambridge University Press.

[AN1] Arnold, A., Nivat, M., *Non deterministic recursive programs*, in Fundamentals of Computation Theory, 12-21, Lecture Notes in Computer Science, No 56, Springer-Verlag, Heidelberg, 1977.

[AN2] Arnold, A., Nivat, M., *Formal Computations of Non-Deterministic Recursive Program Schemes*, Math. System Theory, vol 13, 1980, 219-236

[AN3] Arnold, A., Nivat, M., *Metric Interpretations of Infinite Trees and Semantics of Non-Deterministic Recursive Program Schemes*, TCS, vol 11, 1980, 181-205.

[AN4] Arnold, A., Niwinski, D., *Fixed Point Characterization of Büchi Automata on Infinite Trees*, J. Inf. Process Cybern. EIK 26, 1990, 451-459

[ES] Engelfriet, J., Schmidt, E.-M., *IO and OI*, I,J. Comput. System Sci. 15, 1998

[GS] Giesceg, F., Steinby, M., *Tree Automata*, Akademiai Kiado, Budapest, 1984

[Ko] Kowalski, R., *Algorithm = Logic+Control,* Communications of the ACM, Vol.22., No 7, 1979, 424-436

[Na] Naudin, P., *Comparison et Equivalence de Semantiques pour les schémas de Programmes non déterministes*, Informatique théorique et Applications, vol 21, no 1, 1987, 59-91.

[NS] Nivat, M., Saoudi, A., *Automata on Infinite Objects and their Applications to Logic and Programming*, Information and Computation 83, 1989, 41-64

[Po] Poigné, A., *Effective Computations of Non-Determnistic Schemes*, Lecture Notes in Computer Science, vol. 137, 1982.

[Sa] Saoudi, A., *Generalized Automata on Infinite Trees and Muller-McNaughton's Theorem*, Theoret. Comput. Sci. 84, 1991,165-177

[Ta] Takahashi, M., *The Greatest Fixed-points and Rational Omega Tree Languages*, Theoret. Comput. Sci. 44, 1986, 259-274

[Th] Thomas, W., *Automata on Infinite Objects*, Handbook of Theoretical Computer Science, Vol.B, Elsevier, 1990, 133-191