Learning of Regular ω-Tree Languages

M. Jayasrirani[1], M.H. Begam[1], and D.G. Thomas[2]

[1] Arignar Anna Government Arts College, Walajapet
[2] Madras Christian College, Chennai - 600 059
dgthomasmcc@yahoo.com

Abstract. We introduce two subclasses of regular ω-tree languages called local ω-tree languages and Buchi local ω-tree languages. Automata characterization for these ω-tree languages is given. For these subclasses and ω-regular tree languages learning algorithms are given.

1 Introduction

The theory of tree automata and tree languages emerged in the middle of 1960s. Saoudi et al. [2] have considered infinite trees (ω-trees), recognizable ω-tree languages and regular ω-tree languages. Infinite trees are useful to decide second order theories. In this paper local ω-tree languages and Buchi local ω-tree languages are defined and automata characterization for ω-regular tree languages in terms of local ω-tree languages and Buchi local ω-tree languages is given. There is no learning algorithm so far in the literature for the local ω-tree languages, Buchi local ω-tree languages and regular ω-tree languages. We give learning algorithms for these classes of ω-tree languages. Our approach is similar to the one given in [3].

2 Definitions and Results

Definitions concerning trees, root of a tree, frontier of a tree, forks of a tree, infinite trees, automata on infinite trees and ultimately periodic infinite trees can be found in [1,2].
T_Σ stands for the set of all finite trees over Σ.
T_Σ^ω stands for the set of all infinite trees over Σ.
$root(t)$ stands for root of a tree t.
$fork(t)$ stands for fork of a tree t.
$fork(\Sigma)$ stands for the set of all forks of Σ-trees.
$Frfork(t)$ stands for the set of all forks of a tree t that end with frontiers of t.

Definition 1. *A ω-tree language $L \subseteq T_\Sigma^\omega$ is called a local ω-tree language if there exists a pair $S = \{R, E\}$ (called a local system) where $R \subseteq \Sigma$ and $E \subseteq fork(\Sigma)$ such that*

$$L = \{t \in T_\Sigma^\omega : root(t) \in R, fork(t) \subseteq E\}$$

The elements in $fork(\Sigma)$ occur infinitely many times. In this case we write $L = L^\omega(R, E)$. The set of all local ω-tree languages is denoted by $\mathcal{L}^\omega$.
$L = \{a(b^\omega, c^\omega), a(c^\omega, b^\omega)\}$ is a local ω-tree language.

A. Clark, F. Coste, and L. Miclet (Eds.): ICGI 2008, LNAI 5278, pp. 295–297, 2008.

Definition 2. *A Buchi local system over Σ is an ordered triple $S = \{R, E, E'\}$ where $R \subseteq \Sigma$, $E \subseteq fork(\Sigma)$ and $E' \subseteq E$. We denote $L^\omega(R, E, E')$ a Buchi local ω-tree language defined as*

$$L'(R, E, E') = \{t \in T_\Sigma : root(t) \in R, fork(t) \subseteq E,\ inf\, fork(t) \cap E' \neq \phi\}$$

where $inf\, fork(t)$ is the set of elements in $fork(t)$ which occur infinitely many times in t. An ω-tree language $L \subseteq T^\omega_\Sigma$ is called a Buchi local ω-tree language if there exists a Buchi local system such that $L = L^\omega(R, E, E')$. The set of all Buchi local ω-tree languages is denoted by $\mathcal{L}^\omega_{BE}$. $L = \{a(b^\omega, c^\omega), a(c^\omega, b^\omega)\}$ is a Buchi local ω-tree language.

Theorem 1. *Every regular ω-tree language (recognizable ω-tree language) is an alphabetic homomorphic image of a Buchi local (local) ω-tree language.*

We can give construction procedures for deterministic Buchi k-ary ω-tree automaton M such that $L = L^\omega(M)$ where L is a local (Buchi local) ω-tree language.

3 Learning Buchi Local ω-Tree Languages

Definition 3. *Let $L \in \mathcal{L}^\omega_{BE}$ be such that $L = L^\omega(S)$ for some Buchi local system $S = (R, E, E')$ over an alphabet Σ. S is said to be minimal for L, if for any other Buchi local system $S_1 = (R_1, E_1, E'_1)$ over Σ with $L = L^\omega(S_1)$ we have $R \subseteq R_1$, $E \subseteq E_1$ and $E' \subseteq E'_1$.*

Definition 4. *Let K be a finite sample of ultimately periodic infinite trees. Let $R_K = root(K) = \{root(t) : t \in K\}$, $E_K = fork(K) = \cup_{t \in K} fork(t)$*

$$E'_K = \cup_{a(b^\omega, c^\omega)} F_r fork(t)$$

$S_K = (R_K, E_K, E'_K)$ is called a Buchi local system associated with K and $L = L^\omega(S_K)$ is called Buchi local ω-tree language associated with K.

Theorem 2. *If K, K' are finite samples of ultimately periodic ω-trees of T^ω_Σ then*

1. $K \subseteq L^\omega(S_K)$
2. $K \subseteq K'$ *implies* $L^\omega(S_K) \subseteq L^\omega(S_{K'})$
3. $L \in \mathcal{L}^\omega_{BE}$ *with* $K \subseteq L$ *implies* $L^\omega(S_K) \subseteq L$

Definition 5. *Let L be a local (Buchi local) ω-tree language. A finite subset F of T^ω_Σ is called a characteristic sample for L if L is the smallest local (Buchi local) ω-tree language containing F.*

Theorem 3. *If F is the characteristic sample for a local (Buchi local) ω-tree language and $F \subseteq K \subseteq L$ then $L = L^\omega(S_K)$.*

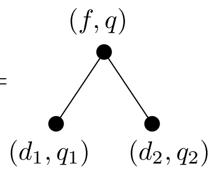
Theorem 4. *There effectively exists a characteristic sample for any local (Buchi local) ω-tree language.*

Theorem 5. *Given an unknown local (Buchi local) ω-tree language we give an algorithm that learns in the limit from positive data, a local system (Buchi local system) S_F such that $L^\omega(S_F) = L$.*

4 Learning Regular ω-Tree Languages

In this section we give a learning algorithm for regular ω-tree languages from positive data and restricted superset queries.

If L is a regular ω-tree language over Σ and if L is recognized by a Buchi k-ary ω-tree automaton $M = < Q, \Sigma, \delta, q_0, F >$ then by theorem 1 there is a Buchi local ω-tree language U over Ω and a strictly alphabetic morphism $h : \Omega \rightarrow \Sigma$ such that $h(U) = L$. Consider an ultimately periodic infinite tree $a(b^\omega, c^\omega)$ in U. Let $t = a(b(b, b), c(c, c))$ where $a, b, c \in \Sigma$. Let $g(t) = (f, q) < (d_1, q_1), (d_2, q_2) >$ where $d_i =$ root of (t_i) $(i = 1, 2, t_i$ are the subtrees of $t)$ be a tree over Ω. i.e.,

$g(t) =$ 

The tree $g(t)$ is said to be a good tree for t if d_1, d_2 are the children of f. Let $G(t)$ be the set of all good trees in $h^{-1}(t)$ for t. If H_U is a characteristic sample for U, then there exists a finite set of positive data S_L of L such that $H_U \subseteq h^{-1}(S_L)$.

We provide a learning algorithm for regular ω-tree languages from positive data and restricted superset queries.

Theorem 6. *Given an unknown regular (recognizable) ω-tree language L, we can give algorithm that effectively learns from positive data and restricted superset queries, a Buchi local (local) system S such that $L = h(L^\omega(S))$.*

References

1. Gecseg, F., Steinby, M.: Tree languages. In: Handbook of Formal Languages, vol. 3, pp. 1–68. Springer, Heidelberg (1997)
2. Saoudi, A.: Rational and recognizable infinite tree sets. In: Tree Automata and Languages, pp. 225–234. Elsevier Science, Amsterdam (1992)
3. Saoudi, A., Yokomori, T.: Learning local and recognizable ω-languages and Monadic Logic Programs. In: Proc. EuroColt, 1993, pp. 157–169. Oxford Univ. Press, Oxford (1994)