ELSEVIER

Theoretical Computer Science 233 (2000) 1–18

Theoretical Computer Science

www.elsevier.com/locate/tcs

the operational and the logical view

Angelo Monti [a], Adriano Peron [b,*]

[a] *Dipartimento Scienze dell'Informazione, Università di Roma (La Sapienza), 00198, Via Salaria 113, Italy*

[b] *Dipartimento di Matematica e Informatica, Università di Udine, 33100, Via delle Scienze 206, Italy*

Received May 1995; revised September 1997
Communicated by A. Salomaa

Abstract

The class of ω-languages recognized by systolic (binary) tree automata is introduced. This class extends the class of Büchi ω-languages though maintaining the closure under union, intersection and complement and the decidability of emptiness. The class of systolic tree ω-languages is characterized in terms of a (suitable) concatenation of (finitary) systolic tree languages. A generalization of Büchi Theorem is provided which establishes a correspondence between systolic tree ω-languages and a suitable extension of the sequential calculus $S1S$.

Keywords: Systolic automata; Automata on ω-words; Decidable monadic second order logic

1. Introduction

The subject of automata accepting infinite sequences was established in the 1960s by Büchi, McNaughton and Rabin (for a survey, see [9]). Their work opened connections between automata theory and fields of logic and set-theoretic topology. The early papers were motivated by decision problems in mathematical logic (e.g. see [1]). One motivation for considering automata on infinite sequences (Büchi automata) was the analysis of the sequential calculus ($S1S$), a system of monadic second-order logic for the formalization of properties of sequences. Büchi showed that any condition on sequences that it is written in this calculus can be reformulated as a statement about acceptance of sequences by a Büchi automaton (Büchi Theorem). The resulting theory

* Corresponding author.
E-mail address: peron@dimi.uniud.it. (A. Peron)

0304-3975/00/$ - see front matter
PII: S0304-3975(97)00257-0

is fundamental for those areas in computer science where non-terminating computations are studied (e.g. modal logics of programs and specification and verification of concurrent programs).

Systolic tree automata were introduced in the 1980s by Culik II, Salomaa and Wood (see [2]) as a tool for studying computational power and properties of systolic systems. From a formal language viewpoint, the main interest of systolic tree automata is that the class of (finitary) languages they recognize strictly includes the class of regular languages still preserving decidability and closure properties of regular languages. Systolic automata as recognizers of finitary languages have been largely investigated (e.g. see [5]), but they have never been considered as acceptors of infinite sequences.

We show that, as the class of systolic tree languages is a proper extension of the class of regular languages, so is the class of systolic tree ω-languages a proper extension of the class of Büchi ω-languages. We prove also that systolic tree ω-languages enjoy the same nice properties of Büchi ω-languages, namely, the emptiness problem for systolic tree automata on infinite sequences is decidable and the class of recognized ω-languages is closed under boolean operations. Moreover, as Büchi ω-languages can be characterized in terms of union and concatenation of regular sets, so systolic tree ω-languages can be characterized in terms of union and a restricted concatenation of systolic tree (finitary) languages.

The correspondence of systolic tree ω-languages with a system of monadic second order logic is then established. We provide a characterization of systolic tree ω-languages as languages defined by a suitable (decidable) extension of $S1S$. We extend $S1S$ by a unary function f which allows to impose on natural numbers a structure appropriate for simulating a binary tree systolic computation. The function f is related to the characteristic function χ of the predicate "is a power of 2", but it is more expressive than χ. The extension of $S1S$ by χ has been proved decidable by Elgot and Rabin in the 1960s (see [3]). We believe that the extension of $S1S$ by f is not equivalent to the one by χ and that the method proposed by Elgot and Rabin cannot be exploited for obtaining our extension. However, the main concern of this work is not the investigation of decidable extensions of $S1S$ and the importance of our extension rely on the fact that it gives the logical view of systolic tree ω-languages.

The theory, which has been developed here in the particular case of systolic binary tree automata, can be rephrased with obvious changes for systolic automata over trees of any degree. The present paper is an extension of [6]. The remainder of the paper is organized as follows. In Section 2 we recall the definitions of Büchi automata and *SBTA*. Then, in Section 3 we define an acceptance condition for *SBTA* suitable for ω-words. In Section 4, we prove the closure of systolic tree ω-languages under union, intersection and complementation and we provide a characterization of that class of languages. In Section 5, we give complexity of emptiness problem. In Section 6, we establish the relationship between systolic tree ω-languages and regular ω-languages. Finally in Section 7, we provide an extension $S1S^{+}$ of $S1S$ that gives the logical counterpart of systolic tree ω-languages.

2. Preliminaries

Throughout this paper Σ denotes an alphabet and Σ^* (resp.: Σ^ω) denotes the set of (finite) words (resp.: ω-*words*) on Σ. Finite words are indicated by $u, v, w, \ldots$ and sets of finite words by $U, V, W, \ldots$. Letters $\alpha, \beta, \ldots$ are used for ω-words and $L, L', \ldots$ for sets of ω-words. For an ω-word α, $\alpha(i)$, with $i \in \mathbb{N}$, denotes the ith element of α; $\alpha(m,n)$ denotes the segment $\alpha(m)\ldots\alpha(n)$ of α. The symbol . denotes concatenation on strings and V^ω is the set of ω-words having the form $v_0 . v_1 . v_2 \ldots$ with $v_i \in V$ for $i \in \mathbb{N}$.

Definition 1. A *Büchi automaton* is a tuple $\mathscr{B} = \langle \Sigma, \mathscr{Q}, q_0, \Delta, \mathscr{F} \rangle$, where
- Σ is the finite *input alphabet*;
- $\mathscr{Q}$ is the finite set of *states*;
- q_0 is the *initial state*;
- $\Delta \subseteq \mathscr{Q} \times \Sigma \times \mathscr{Q}$ is the *transition relation*;
- $\mathscr{F} \subseteq Q$ is the set of *final states.*

Automaton $\mathscr{B}$ is *deterministic* iff $\langle q, a, q' \rangle, \langle q, a, q'' \rangle \in \mathscr{Q}$ implies $q' = q''$.
A *run* of $\mathscr{B}$ on an ω-word $\alpha \in \Sigma^\omega$ is a ω-word $\sigma \in \mathscr{Q}^\omega$ such that $\sigma(0) = q_0$ and $\langle \sigma(i), \alpha(i), \sigma(i+1) \rangle \in \Delta$, for $i \geqslant 0$. A run σ is *successful* iff some state of $\mathscr{F}$ occurs infinitely often in σ.

Automaton $\mathscr{B}$ *accepts* α iff there is a successful run σ on α. The ω-language accepted by $\mathscr{B}$, denoted as $\mathscr{L}_\omega(\mathscr{B})$, is the set $\{\alpha \in \Sigma^\omega \colon \alpha \text{ is accepted by } \mathscr{B}\}$.

An ω-language L is *regular*, iff there is a Büchi automaton $\mathscr{B}$ such that $\mathscr{L}_\omega(\mathscr{B}) = L$. The class of regular ω-languages is denoted by $\mathscr{L}_\omega(BA)$. The class of ω-languages recognized by deterministic Büchi automata is denoted by $\mathscr{L}_\omega(DBA)$.

We recall some results about regular ω-languages (see [9]):
1. $\mathscr{L}_\omega(BA)$ is (effectively) closed under union, intersection and complementation;
2. $\mathscr{L}_\omega(DBA) \subset \mathscr{L}_\omega(BA)$;
3. $L \in \mathscr{L}_\omega(BA)$ iff L is the finite union of sets of ω-words having the form $U.V^\omega$, where $U, V \subseteq \Sigma^*$ are regular sets;
4. the emptiness problem for Büchi automata is decidable.

Systolic languages are sets of (finite) words accepted by systolic automata (see [2]). In the following we give an informal description of *Systolic Binary Tree Automata* (shortly *SBTA*). A systolic automaton consists of an infinite number of nodes which can be interpreted as memoryless processors. Nodes are linked among them and the resulting structure is an (infinite) leafless perfectly balanced binary tree. In order to process a word w, the first level m of the tree is chosen which has at least $|w|$ nodes.[1] Now, the automaton is fed in such a way that adjacent processors at level m are fed with adjacent symbols of w, and that the rightmost processor is fed with the last symbol of w. If the number of processors is greater than the word length, then exceeding processors (i.e. each ith processor, for $1 \leqslant i < 2^m - |w|$) are fed with a special

[1] By $|w|$ we denote the length of w.

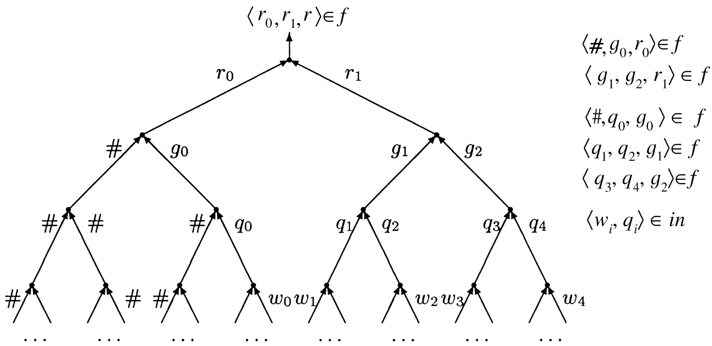

Fig. 1. A computation of a *SBTA* fed with $w_0 . w_1 . w_2 . w_3 . w_4$.

symbol #. In Fig. 1 an example is given for an input word of length five. Now, all the processors at level m synchronously output, according to the *input relation*, a symbol belonging to the *state alphabet* Q. Each processor at level $m-1$ receives the couple of states output by its pair of sons and it synchronously (with respect to processors at the same level) outputs a symbol belonging to Q according to the *transition relation*. Therefore, information flows bottom-up, in parallel and synchronously, level by level. The word is accepted whenever the root of the tree outputs a symbol belonging to the given set of *final states*.

Definition 2. A *systolic automaton* is a tuple $\mathscr{A} = \langle \Sigma, Q, in, f, F \rangle$, where
- Σ is the finite *input alphabet*;[2]
- Q is the finite set of *states*;
- $in \subseteq (\Sigma \cup \{\#\}) \times (Q \cup \{\#\})$ is the *input relation* such that
  $\langle x, \# \rangle \in in$ iff $x = \#$;
- $f \subseteq (Q \cup \{\#\}) \times (Q \cup \{\#\}) \times (Q \cup \{\#\})$ is the *transition relation* such that
  $\langle p, q, \# \rangle \in f$ iff $p = q = \#$;
- $F \subseteq Q \cup \{\#\}$ is the set of *final states*.

An automaton $\mathscr{A}$ is *deterministic* if $\langle a, q' \rangle, \langle a, q'' \rangle \in in$ implies $q' = q''$ and $\langle q_1, q_2, q' \rangle$, $\langle q_1, q_2, q'' \rangle \in f$ implies $q' = q''$.

The relation $O_{\mathscr{A}} \subseteq (\Sigma \cup \{\#\})^* \times Q$ is recursively defined as follows:
- if $|w| = 1$, then $\langle w, q \rangle \in O_{\mathscr{A}}$ iff $\langle w, q \rangle \in in$;
- if $2^{m-1} < |w| \leqslant 2^m$, with $m > 0$, then $\langle w, q \rangle \in O_{\mathscr{A}}$ iff $\langle q_1, q_2, q \rangle \in f$ where q_1, q_2 are such that $\langle w_1, q_1 \rangle, \langle w_2, q_2 \rangle \in O_{\mathscr{A}}$ with $|w_1| = |w_2| = 2^{m-1}$ and $w_1 . w_2 = \#^{2^m - |w|} . w$.

The language recognized by $\mathscr{A}$, denoted as $\mathscr{L}(\mathscr{A})$, is the set

$$\{w \in \Sigma^* : \langle w, q \rangle \in O_{\mathscr{A}}, q \in F\}.$$

[2] Assume that $\# \notin \Sigma$.

The class of languages recognized by *SBTA* (resp.: by deterministic *SBTA*) is denoted by $\mathscr{L}(SBTA)$ (resp.: $\mathscr{L}(DSBTA)$).

Remark. The given definition of systolic tree automaton differs from the standard definition of [2] on the way processors are fed with an input word. In the standard definition, adjacent processors at the input level are fed with adjacent symbols of the input word, but the leftmost processor is fed with the first symbol of the word. The class of systolic binary tree languages recognized by automata defined as in [2] and the class $\mathscr{L}(SBTA)$ are uncomparable. However, the class $\mathscr{L}(SBTA)$ and the standard class of systolic binary tree languages enjoy analogous properties. In particular, proofs of properties of $\mathscr{L}(SBTA)$ can be obtained by slightly modifying those given in literature for the corresponding properties of the standard class. For this reason, we shall argue a property of $\mathscr{L}(SBTA)$ by referring to the proof of the corresponding property in the standard case. We modify the standard notion of systolic automaton only for technical reasons, i.e. for achieving a more elegant characterization of the class of systolic ω-languages in terms of finitary systolic languages.

We recall some properties of $\mathscr{L}(SBTA)$ (see [5]):

1. $\mathscr{L}(SBTA)$ is (effectively) closed under union, intersection and complementation;
2. $\mathscr{L}(DSBTA) = \mathscr{L}(SBTA)$;
3. the emptiness problem for systolic tree automata is decidable;
4. the class of regular languages is a proper subclass of $\mathscr{L}(SBTA)$.

3. Systolic tree ω-languages

We introduce now a notion of stepwise systolic computation for ω-words. The run of a Büchi automaton on an ω-word α stores at the ith position the state q resulting from processing the prefix $\alpha(0, i-1)$ of α. The state resulting from processing the prefix $\alpha(0, i)$ can be obtained from q and $\alpha(i)$ according to the transition relation. So, the computation proceeds step by step by processing one symbol of α at a time. In the case of *SBTA*, the computation proceeds by processing at each time a segment of α whose length doubles step by step. In particular, an ω-word on the set of states Q stores at ith position the state q resulting from processing the prefix $\alpha(0, 2^i - 1)$ of α. The state resulting from processing the next prefix $\alpha(0, 2^{i+1} - 1)$ is obtained, according to the transition relation f, from q and from a state output by the systolic automaton fed with $\alpha(2^i, 2^{i+1} - 1)$. The "structure" of processing an ω-word α is given in Fig. 2. The left-hand side edge of the structure is constituted by nodes associated with states obtained by processing prefixes of α whose length is a power of two (that sequence of states is called *systolic run*).

Definition 3. For a systolic automaton $\mathscr{A} = \langle \Sigma, Q, in, f, F \rangle$, a *systolic run* of $\mathscr{A}$ on an ω-word $\alpha \in \Sigma^\omega$, is an ω-word $\sigma \in Q^\omega$ such that

– $\langle \alpha(0), \sigma(0) \rangle \in in$;
– $\langle \sigma(i-1), q, \sigma(i) \rangle \in f$, with $\langle \alpha(2^{i-1}, 2^i - 1), q \rangle \in O_{\mathscr{A}}$, for $i \geqslant 1$.

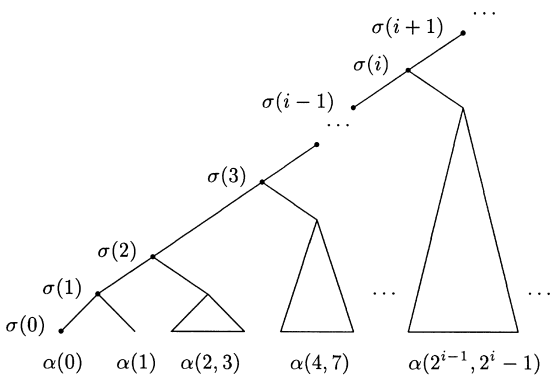

Fig. 2. A systolic run σ on α.

A systolic run σ is *successful* iff some state of F occurs infinitely often in σ.
Automaton $\mathscr{A}$ *accepts* α iff there is a successful systolic run on α.
The ω-language recognized by $\mathscr{A}$, denoted as $\mathscr{L}_\omega(\mathscr{A})$, is the set

$$\{\alpha \in \Sigma^\omega\colon \mathscr{A} \text{ accepts } \alpha\}.$$

The class of ω-languages recognized by systolic (resp.: by deterministic systolic) tree automata is denoted by $\mathscr{L}_\omega(SBTA)$ (resp.: $\mathscr{L}_\omega(DSBTA)$).

There is a strict relationship between the notion of run for Büchi automaton and that of systolic run. A systolic run of an automaton $\mathscr{A}$ on α can be alternatively viewed as a run of a Büchi automaton induced by the transition relation of $\mathscr{A}$ over an ω-word α' which results from preprocessing segments of α. An ω-word α' *is a preprocessing of* α *under* $\mathscr{A} = \langle \Sigma, Q, in, f, F\rangle$, if

1. $\langle \alpha(0), \alpha'(0)\rangle \in in$;
2. $\langle \alpha(2^i, 2^{i+1}-1), \alpha'(i)\rangle \in O_{\mathscr{A}}$, for $i > 0$.

The Büchi automaton *induced by* f is $\mathscr{B} = \langle Q, Q \cup \{q_0\}, q_0, \Delta, F\rangle$, where $q_0 \notin Q$ and $\Delta = \{\langle q_0, q, q\rangle : q \in Q\} \cup f$.

Now, it follows from the definition that there is a successful systolic run of $\mathscr{A}$ on $\alpha \in \Sigma^\omega$ iff there is a successful run of $\mathscr{B}$ on some preprocessing of α under $\mathscr{A}$.

The example below shows that, by introducing preprocessing, non-regular ω-languages can be recognized. Conversely, in Section 6, we shall show that, by introducing preprocessing, the ability of recognizing regular ω-languages is not lost.

Example 1. Let us consider the ω-language $L = \{a^{2^i}.\{b\}^\omega : i \geqslant 0\}$. The language L is clearly non-regular and it is recognized by the deterministic systolic tree automaton $\mathscr{A} = \langle \Sigma, Q, in, f, F\rangle$ defined as follows:

- $\Sigma = \{a, b\}$; $Q = \{q_1, q_2, q_3\}$; $F = \{q_3\}$;
- $in = \{\langle a, q_1\rangle, \langle b, q_2\rangle\}$;
- $f = \{\langle q_1, q_1, q_1\rangle, \langle q_2, q_2, q_2\rangle, \langle q_1, q_2, q_3\rangle, \langle q_3, q_2, q_3\rangle\}$.

Another example of non-regular ω-language is

$$L' = \{b.b.a^{2^1-1}.b.a^{2^2-1}.b\ldots b.a^{2^i-1}.b.\ldots\}.$$

The language L' is recognized by a deterministic systolic tree automaton $\mathscr{A} = \langle \Sigma, Q, in, f, F\rangle$ defined as follows:
- $\Sigma = \{a, b\}$; $Q = \{q_1, q_2, q_3, q_4\}$; $F = \{q_3\}$;
- $in = \{\langle a, q_2\rangle, \langle b, q_1\rangle\}$;
- $f = \{\langle q_1, q_1, q_3\rangle, \langle q_2, q_1, q_4\rangle, \langle q_3, q_4, q_3\rangle, \langle q_2, q_2, q_2\rangle, \langle q_2, q_4, q_4\rangle\}$.

4. Closure properties and characterization

In this section we prove that the class of systolic tree ω-languages is closed under union, intersection and complementation. Then, we provide a characterization of systolic tree ω-languages in terms of (finitary) systolic tree languages. The proof of closure of the class $\mathscr{L}_\omega(SBTA)$ under union and intersection is similar to that given in the case of regular ω-languages (see [9]).

Theorem 4. *The class $\mathscr{L}_\omega(SBTA)$ is (effectively) closed under union and intersection.*

Proof. Let $\mathscr{A}_i = \langle \Sigma, Q_i, in_i, f_i, F_i\rangle$ (for $i \in \{1, 2\}$) be two *SBTA*. Assume, without loss of generality, that Q_1 and Q_2 are disjoint. The automaton

$$\mathscr{A} = \langle \Sigma, Q_1 \cup Q_2, in_1 \cup in_2, f_1 \cup f_2, F_1 \cup F_2\rangle$$

recognizes the ω-language $\mathscr{L}_\omega(\mathscr{A}) = \mathscr{L}_\omega(\mathscr{A}_1) \cup \mathscr{L}_\omega(\mathscr{A}_2)$.

An automaton $\mathscr{A} = \langle \Sigma, Q, in, f, F\rangle$ recognizing the ω-language $\mathscr{L}_\omega(\mathscr{A}_1) \cap \mathscr{L}_\omega(\mathscr{A}_2)$ is defined as follows:
- $Q = Q_1 \times Q_2 \times \{0, 1, 2\}$; $F = Q_1 \times Q_2 \times \{2\}$;
- $in = \{\langle a, \langle p, q, 0\rangle\rangle : \langle a, p\rangle \in in_1, \langle a, q\rangle \in in_2\}$;
- $f = \{\langle\langle p_1, q_1, 0\rangle, \langle p_2, q_2, i\rangle, \langle p, q, 0\rangle\rangle :$
  $\langle p_1, p_2, p\rangle \in f_1, \langle q_1, q_2, q\rangle \in f_2, p \notin F_1, i \in \{0, 1, 2\}\} \cup$
  $\{\langle\langle p_1, q_1, 0\rangle, \langle p_2, q_2, i\rangle, \langle p, q, 1\rangle\rangle :$
  $\langle p_1, p_2, p\rangle \in f_1, \langle q_1, q_2, q\rangle \in f_2, p \in F_1, i \in \{0, 1, 2\}\} \cup$
  $\{\langle\langle p_1, q_1, 1\rangle, \langle p_2, q_2, i\rangle, \langle p, q, 1\rangle\rangle :$
  $\langle p_1, p_2, p\rangle \in f_1, \langle q_1, q_2, q\rangle \in f_2, q \notin F_2, i \in \{0, 1, 2\}\} \cup$
  $\{\langle\langle p_1, q_1, 1\rangle, \langle p_2, q_2, i\rangle, \langle p, q, 2\rangle\rangle :$
  $\langle p_1, p_2, p\rangle \in f_1, \langle q_1, q_2, q\rangle \in f_2, q \in F_2, i \in \{0, 1, 2\}\} \cup$
  $\{\langle\langle p_1, q_1, 2\rangle, \langle p_2, q_2, i\rangle, \langle p, q, 0\rangle\rangle :$
  $\langle p_1, p_2, p\rangle \in f_1, \langle q_1, q_2, q\rangle \in f_2, i \in \{0, 1, 2\}\}$. □

Proving that $\mathscr{L}_\omega(SBTA)$ is closed under complementation requires a bit more effort. We exploit the technique adopted by Sistla et al. in [8] for achieving the explict constructive complementation of Büchi automata. The idea is that, given a systolic tree ω-language L, we can find a finite set of systolic tree ω-languages $L_1 \ldots L_n$ such that $L=\bigcup_{i=1}^{n} L_i$, and a finite set of systolic tree ω-languages $L_{n+1}\dots L_m$ such that $\Sigma^\omega - L=\bigcup_{i=n+1}^{m} L_i$ (this technique is known as "saturation"). In particular, each systolic tree ω-language L_i (for $1\leqslant i\leqslant m$) can be obtained by suitably concatenating the elements of a pair of systolic finitary languages. For that purpose we define a notion of *restricted concatenation*.

Throughout this section $2^{[m,n]}$ (for $m\leqslant n$) is a short notation for $\sum_{j=m}^{n} 2^j$.

Definition 5. The binary operation $\diamond$ on $\Sigma^\star$ is defined for $w,w'\in\Sigma^\star$ iff

$$|w|=2^m \text{ and } |w'|=2^{[m,n]}, \text{ with } m<n,$$

and in this case $w\diamond w'=w.w'$. For $U,V\subseteq\Sigma^\star$:
1. $U\diamond V=\{w\diamond w' : w\in U, w'\in V\}$;
2. $U^{\diamond\omega}=\{w_1\diamond w_2\diamond w_3\diamond\dots : w_i\in U, \text{ for } i\geqslant 1\}$;
3. $U\diamond V^{\diamond\omega}=\{w_1\diamond w_2\diamond w_3\diamond\dots : w_1\in U \text{ and } w_i\in V, \text{ for } i\geqslant 2\}$;

Remark. The operation $\diamond$ is not associative. So, when we write $w_1\diamond w_2\diamond w_3\diamond\dots$, we actually mean $((w_1\diamond w_2)\diamond w_3)\diamond\dots$. The class of systolic ω-languages is closed under restricted (ω-)concatenation of (finitary) systolic sets.

Lemma 6. *For any $U,V\in\mathscr{L}(SBTA)$, $U\diamond V^{\diamond\omega}\in\mathscr{L}_\omega(SBTA)$.*

Proof. Let $\mathscr{A}_i=\langle\Sigma,Q_i,in_i,f_i,F_i\rangle$, $i\in\{1,2\}$, be two *SBTA* such that $\mathscr{L}(\mathscr{A}_1)=U$ and $\mathscr{L}(\mathscr{A}_2)=V$. Assume without loss of generality that $\mathscr{A}_1$ and $\mathscr{A}_2$ are deterministic and Q_1, Q_2, and $Q_2\times\{0,1\}$ are pairwise disjoint sets. An automaton $\mathscr{A}=\langle\Sigma,Q,in,f,F\rangle$ such that $\mathscr{L}_\omega(\mathscr{A})=U\diamond V^{\diamond\omega}$ is defined as follows:
- $Q=Q_1\cup Q_2\cup(Q_2\times\{0,1\})$;
- $F=Q_2\times\{0\}$;
- $in=in_1\cup in_2$;
- $f=f_1\cup f_2\cup\{\langle q_1,q_2,\langle q_3,0\rangle\rangle : q_1\in F_1, q_2\in Q_2, \langle\#,q_2,q_3\rangle\in f_2\}\cup$
  $\{\langle\langle q_1,0\rangle,q_2,\langle q_3,1\rangle\rangle : \langle q_1,q_2,q_3\rangle\in f_2\}\cup$
  $\{\langle\langle q_1,1\rangle,q_2,\langle q_3,1\rangle\rangle : \langle q_1,q_2,q_3\rangle\in f_2\}\cup$
  $\{\langle\langle q_1,1\rangle,q_2,\langle q_3,0\rangle\rangle : \langle\#,q_2,q_3\rangle\in f_2, q_1\in F_2\}$.

We prove that $\mathscr{L}_\omega(\mathscr{A})=U\diamond V^{\diamond\omega}$.

($\subseteq$) If $\alpha\in\mathscr{L}_\omega(\mathscr{A})$, then there exists a successful run σ and an infinite (increasing) sequence of integers $m_1,m_2,\dots$ such that $\sigma(m_i)\in F$, for $i\geqslant 1$ and such that $\sigma(k)\in F$ implies $k=m_j$, for some j. By a simple induction on $i\in\mathbb{N}$ it can be proved that $\alpha(0,2^{m_i-1}-1)=u\diamond v_1\diamond\dots\diamond v_i$, for some $u\in U$ and $v_j\in V$, $1\leqslant j\leqslant i$.

($\supseteq$) Let $\alpha=u\diamond v_1\diamond v_2\diamond\dots$, $u\in U$, $v_i\in V$, $i\geqslant 1$. Consider the infinite sequence of integers $m_0,\dots,m_i\dots$ such that $|v_0.v_1.v_2\dots v_i|=2^{m_i}$, $i\geqslant 0$. It is easy to prove by induction on i, that there exists a run σ on α such that $\sigma(m_j+1)\in F$, for all $1\leqslant j\leqslant i$. $\square$

For an *SBTA* $\mathscr{A}$, we can determine now the set of ω-languages such that $\mathscr{L}_\omega(\mathscr{A})=\bigcup_{i=1}^{n} L_i$ and $\Sigma^\omega-\mathscr{L}_\omega(\mathscr{A})=\bigcup_{i=n+1}^{m} L_i$. Each language L_i takes the form $U_i\diamond V_i^{\diamond\omega}$, where U_i and V_i are (finitary) systolic languages. To define languages U_i and V_i ($1\leqslant i\leqslant m$), we fit for *SBTA* the *generalized subset construction* given in [8] in the case of Büchi automata. Assume that $Q=\{s_1,\dots,s_n\}$ is the set of states of $\mathscr{A}$. We construct the structure $Q_{\mathscr{A}} = (2^{Q\times\{0,1\}})^n$ and for each element $p \in Q_{\mathscr{A}}$ we construct an automaton $\overline{\mathscr{A}_p}$. The set of states of $\overline{\mathscr{A}_p}$ is $Q_{\mathscr{A}}$ and p is the final state. The automaton $\overline{\mathscr{A}_p}$ has n components (i.e. each state has n components corresponding to states $s_1 \dots s_n$ of $\mathscr{A}$) each of which simulates the behaviour of $\mathscr{A}$ but in a special case. If, during the processing of a word w a node of the tree of $\overline{\mathscr{A}_p}$ receives the special symbol # from the left-hand side son and the state $q \neq \#$ from the right-hand side son (i.e. $2^i < |w| < 2^{i+1}$), then the kth component of $\overline{\mathscr{A}_p}$ behaves as if the node had received the state s_k instead of #. In the particular case when a word w with $|w| = 2^{[m,l]}$ (with $m < l$) is processed, then the kth component of $\overline{\mathscr{A}_p}$ processes w as if it was processing a word $w' \diamond w$ such that $\langle w', s_k\rangle \in O_{\mathscr{A}}$. In this sense we can say that the computation on w *starts from* state s_k. Now, two words w and w' belonging to $\mathscr{L}(\overline{\mathscr{A}_p})$ enjoy the property that a computation on w starting from s_k leads to a state $\bar{s}_k$ iff there is a computation on w' starting from s_k and leading to $\bar{s}_k$, for all $1 \leqslant k \leqslant n$.

Definition 7. Let $\mathscr{A} = \langle \Sigma, Q, in, f, F\rangle$ be a *SBTA* with $Q = \{s_1, \dots, s_n\}$. The *generalized set of states* of $\mathscr{A}$ is $Q_{\mathscr{A}} = (2^{Q\times\{0,1\}})^n$. For $p \in Q_{\mathscr{A}}$, the *generalized automaton* $\mathscr{A}$ *under* p is $\overline{A}_p = \langle \Sigma, Q_{\mathscr{A}}, in', f', \{p\}\rangle$ defined as follows:

– $in' = \{\langle a, \langle X_1, \dots, X_n\rangle\rangle : a \in \Sigma,$
$$X_i = \{\langle q, 0\rangle : \langle a, q\rangle \in in\} \cup \{\langle q, 1\rangle : \langle a, q\rangle \in in, q \in F\},$$
$$1 \leqslant i \leqslant n\} \cup \{\langle \#, \#\rangle\};$$

– $f' = \{\langle\langle X_1, \dots, X_n\rangle, \langle Y_1, \dots, Y_n\rangle, \langle Z_1, \dots, Z_n\rangle\rangle :$
$$Z_i = \{\langle q, 0\rangle : \langle q_1, q_2, q\rangle \in f, \langle q_1, j\rangle \in X_i, \langle q_2, j\rangle \in Y_i, j \in \{0,1\}\} \cup$$
$$\{\langle q, 1\rangle : \langle q_1, q_2, q\rangle \in f, q \in F, \langle q_1, j\rangle \in X_i, \langle q_2, j\rangle \in Y_i, j \in \{0,1\}\} \cup$$
$$\{\langle q, 1\rangle : \langle q_1, q_2, q\rangle \in f, \langle q_1, 1\rangle \in X_i, \langle q_2, j\rangle \in Y_i, j \in \{0,1\}\},$$
$$1 \leqslant i \leqslant n\} \cup$$
$$\{\langle \#, \langle X_1, \dots, X_n\rangle, \langle Z_1, \dots, Z_n\rangle\rangle :$$
$$Z_i = \{\langle q, 0\rangle : \langle s_i, q_1, q\rangle \in f, \langle q_1, j\rangle \in X_i, j \in \{0,1\}\} \cup$$
$$\{\langle q, 1\rangle : \langle s_i, q_1, q\rangle \in f, q \in F, \langle q_1, j\rangle \in X_i, j \in \{0,1\}\} \cup$$
$$\{\langle q, 1\rangle : \langle s_i, q_1, q\rangle \in f, \langle q_1, j\rangle \in X_i, s_i \in F, j \in \{0,1\}\}$$
$$1 \leqslant i \leqslant n\}.$$

The following properties of the language recognized by the generalized automaton $\mathscr{A}$ under p immediately derive from Definition 7.

Proposition 8. *Let $\mathscr{A} = \langle \Sigma, Q, in, f, F\rangle$ be an SBTA with $Q = \{s_1, \dots, s_n\}$ and $p = (S_1, \dots, S_n) \in Q_{\mathscr{A}}$. For a word w with $|w| = 2^{[l,m]}$ and $l < m$, $w \in \mathscr{L}(\overline{\mathscr{A}_p})$ iff for each $v \in \Sigma^\star$ such that $|v| = 2^l$ it holds*

1. *$\langle \sigma(m), 0\rangle \in S_k$, for each run σ on $v \diamond w$ with $\sigma(l) = s_k$;*
2. *$\langle \sigma(m), 1\rangle \in S_k$, for each run σ on $v \diamond w$ such that $\sigma(l) = s_k$ and $\sigma(j) \in F$, for some $l \leqslant j \leqslant m$.*

Lemma 9. *For an SBTA $\mathscr{A}$,*

$$\Sigma^\omega = \bigcup_{p,q \in Q_{\mathscr{A}}} \mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega}.$$

Proof. The proof is a refinement of the proof of Lemma 2.3 in [8]. Let us consider an ω-word α. There is a state $p \in Q_{\mathscr{A}}$ and an infinite set of natural numbers $D = \{i_0, i_1, i_2, \ldots\}$ such that $\alpha(0, 2^{i_j} - 1) \in \mathscr{L}(\overline{\mathscr{A}_p})$ for all $i_j \in D$ ($Q_{\mathscr{A}}$ is a finite set). Without loss of generality we can assume that

$$i_j + 1 < i_{j+1} \text{ for all } i_j \in D. \tag{1}$$

For $s \in Q_{\mathscr{A}}$, let D_s be the set of unordered pairs

$$\{\{i_m, i_n\} : i_m, i_n \in D,\ i_m < i_n,\ \alpha(2^{i_m}, 2^{i_n} - 1) \in \mathscr{L}(\overline{\mathscr{A}_s})\}.$$

Since each automaton $\overline{\mathscr{A}_p}$ is deterministic, the finite set of (non-empty) sets D_s (with $s \in Q_{\mathscr{A}}$) is a partion of all unordered distinct pairs of the infinite set D. By Ramsey's Theorem, there are an infinite subset $D' = \{k_0, k_1, k_2, \ldots\}$ of D and a set D_q such that $\{k_i, k_j\} \in D_q$ for all $k_i, k_j \in D'$ with $k_i < k_j$. This implies that α can be partitioned into words

$$u = \alpha(0, 2^{k_0} - 1),\ \ v_1 = \alpha(2^{k_0}, 2^{k_1} - 1),\ \ v_2 = \alpha(2^{k_1}, 2^{k_2} - 1), \ldots$$

where $u \in \mathscr{L}(\overline{\mathscr{A}_p})$ and $v_i \in \mathscr{L}(\overline{\mathscr{A}_q})$, for all $i \geqslant 1$. Eq. (1) ensures that $\alpha = u \diamond v_1 \diamond v_2 \diamond \ldots$ thus proving that $\alpha \in \mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega}$. $\square$

The languages having the form $\mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega}$, with $p, q \in \mathscr{A}_p$ saturate the language $\mathscr{L}_\omega(\mathscr{A})$.

Lemma 10. *For an SBTA $\mathscr{A}$ and $p, q \in Q_{\mathscr{A}}$,*

$$\text{either}\quad \mathscr{L}_\omega(\mathscr{A}) \cap \mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega} = \emptyset \quad \text{or} \quad \mathscr{L}_\omega(\mathscr{A}) \supseteq \mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega}.$$

Proof. We prove that if $\alpha, \alpha' \in \mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega}$ and $\alpha \in \mathscr{L}_\omega(\mathscr{A})$, then $\alpha' \in \mathscr{L}_\omega(\mathscr{A})$. If $\alpha, \alpha' \in \mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega}$, then $\alpha = u \diamond v_1 \diamond v_2 \diamond \ldots$ and $\alpha' = u' \diamond v_1' \diamond v_2' \diamond \ldots$ with $u, u' \in \mathscr{L}(\overline{\mathscr{A}_p})$ and $v_i, v_i' \in \mathscr{L}(\overline{\mathscr{A}_q})$, for all $i \geqslant 1$. Let l_0 (resp.: k_0) be the number such that $|u| = 2^{l_0}$ (resp.: $|u'| = 2^{k_0}$) and l_i (resp.: k_i) be the number such that $|u \diamond v_1 \diamond \ldots \diamond v_i| = 2^{l_i}$ (resp.: $|u' \diamond v_1' \diamond \ldots \diamond v_i'| = 2^{k_i}$). By Proposition 8(1), if σ is a systolic run of $\mathscr{A}$ on α, then there exists a systolic run σ' of $\mathscr{A}$ on α' such that $\sigma(l_j) = \sigma'(k_j)$, for all $j \geqslant 0$. Moreover, by Proposition 8(2), σ' can be chosen in such a way that if $\sigma(h) \in F$ for some $l_j \leqslant h \leqslant l_{j+1}$, then $\sigma'(h') \in F$ for some $k_j \leqslant h' \leqslant k_{j+1}$, for all $j \geqslant 0$. So, if σ is a successful run of $\mathscr{A}$ on α, then σ' is a successful run of $\mathscr{A}$ on α'. $\square$

Theorem 11. *The class $\mathscr{L}_\omega(SBTA)$ is (effectively) closed under complementation.*

Proof. Let $\mathscr{A}$ be an *SBTA*. By Lemmata 9 and 10 we have that

$$\Sigma^\omega - \mathscr{L}_\omega(\mathscr{A}) = \bigcup_{p, q \in Q_{\mathscr{A}}} \{\mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega} : \mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega} \cap \mathscr{L}_\omega(\mathscr{A}) = \emptyset\}.$$

By Lemma 6, an *SBTA* $\mathscr{A}'$ such that $\mathscr{L}(\mathscr{A}') = \mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega}$ can be effectively constructed, and by Theorem 13 the emptiness of $\mathscr{L}(\overline{\mathscr{A}_p}) \diamond \mathscr{L}(\overline{\mathscr{A}_q})^{\diamond\omega} \cap \mathscr{L}(\mathscr{A})$ can be decided (by Theorem 4, the class $\mathscr{L}_\omega(SBTA)$ is closed under intersection). The thesis follows from the closure of $\mathscr{L}_\omega(SBTA)$ under union (see Theorem 4). □

The saturation technique, exploited for proving closure under complementation, allows also to characterize systolic tree ω-languages. In [1], it has been shown that any regular ω-language can be expressed as the finite union of languages having the form $U.V^\omega$, where U and V are regular sets. Systolic tree ω-languages allow a similar characterization, namely, they can be expressed as the finite union of languages having the form $U \diamond V^{\diamond\omega}$, where U and V are (finitary) systolic languages.

Theorem 12. *An ω-language L belongs to $\mathscr{L}_\omega(SBTA)$ iff L is the finite union of sets having the form $U \diamond V^{\diamond\omega}$ with $U, V \in \mathscr{L}(SBTA)$.*

Proof. ($\Rightarrow$) For a *SBTA* $\mathscr{A}$, we have, by Lemmata 9 and 10, that

$$\mathscr{L}_\omega(\mathscr{A}) = \bigcup_{p,q \in Q_\mathscr{A}} \{\mathscr{L}(\overline{\mathscr{A}}_p) \diamond \mathscr{L}(\overline{\mathscr{A}}_q)^{\diamond\omega} : \mathscr{L}(\overline{\mathscr{A}}_p) \diamond \mathscr{L}(\overline{\mathscr{A}}_q)^{\diamond\omega} \cap \mathscr{L}_\omega(\mathscr{A}) \neq \emptyset\}.$$

($\Leftarrow$) If $U, V \in \mathscr{L}(SBTA)$, then $U \diamond V^{\diamond\omega} \in \mathscr{L}_\omega(SBTA)$ (see Lemma 6). Thus, the thesis follows since $\mathscr{L}_\omega(SBTA)$ is closed under union (see Theorem 4). □

5. Decidability and complexity of emptiness problem

In [6] the decidability of emptiness problem for *SBTA* has been proved by showing that the set of successful systolic runs of an automaton $\mathscr{A}$ is an ω-regular language. The emptiness problem for *SBTA* can then be reduced to the problem of emptiness for Büchi automata which is known to be decidable. In this section we provide also the exact complexity of the problem. It has been proved in [4] that the emptiness for Büchi automata is solvable in linear time. Unfortunately, the emptiness for systolic tree ω-automata turns out to be logspace complete for **PSPACE**. So, the class $\mathscr{L}_\omega(SBTA)$ extends (see Section 6) the class of ω-regular languages preserving the closure and decidability properties enjoyed by the class of ω-regular languages, but unfortunately the greater expressive power is obtained at the expense of complexity.

In the following we define an algorithm which checks the emptiness problem for systolic tree ω-automata in polynomial space.

For a systolic automaton $\mathscr{A} = \langle \Sigma, Q, in, f, F \rangle$, let $(S_i)_{i \geqslant 0}$ be the family of sets of states defined as follows:
- $S_0 = \{q : \langle x, q \rangle \in in\}$;
- $S_{i+1} = \{q : q \in \langle q_1, q_2, q \rangle \in f, q_1, q_2 \in S_i\}$.

For $p \in F$ and $X \subseteq Q$, let $(T(p,X)_i)_{i \geqslant 0}$ be the sequence of pairs of sets of states defined as follows:
- $T(p,X)_0 = \langle X_0, Y_0 \rangle = \langle \{p\}, X \rangle$;
- $T(p,X)_i = \langle X_i, Y_i \rangle$, with

$$X_i = \{q : \langle q_1, q_2, q \rangle \in f,\ q_1 \in X_{i-1},\ q_2 \in Y_{i-1}\} \text{ and}$$
$$Y_i = \{q : \langle q_1, q_2, q \rangle \in f,\ q_1, q_2 \in Y_{i-1}\}, \text{ for } i \geqslant 1.$$

Intuitively, S_i is the set of states that might occur at the ith position in a run generated by $\mathscr{A}$. The pair $T(p,S_i)_j$ is equal to $\{X,S_{i+j}\}$, where X is the set of states that might occur at the $(i+j)$th position in a run having p at its ith position. That Q is finite implies that $(S_i)_{i\geqslant 0}$ and $(T(p,X)_i)_{i\geqslant 0}$ are eventually periodic. The sum of the preperiod and period of $(S_i)_{i\geqslant 0}$ is at most $2^{|Q|}$ and the one of $(T(p,X)_i)_{i\geqslant 0}$ is at most $2^{2|Q|}$. Now, $L_\omega(\mathscr{A})\neq\emptyset$ iff there exist two integers i and j and a final state p such that $i<2^{|Q|}$, $p\in S_i$, $j<2^{2|Q|}$ and $T(p,S_i)_j=\{X,S_i\}$.

A simple procedure requiring polynomial space is the following:

$i=0$; compute S_0
<u>while</u> $i<2^{|Q|}$
    <u>for</u> each $p\in S_i\cap F$ <u>do</u>
        Let $j=0$; $X_0=\{p\}$ and $Y_0=S_i$
        <u>while</u> $j<2^{2|Q|}$
            compute X_{j+1} and Y_{j+1}
            <u>if</u> ($p\in X_{j+1}$ and $Y_{j+1}=X_i$) <u>then</u> <u>print</u>("$L_\omega(\mathscr{A})\neq\emptyset$"), <u>stop</u>
            j=j+1
        <u>end</u>
    <u>end</u>
i=i+1
<u>end</u>
<u>print</u>("$L_\omega(\mathscr{A})=\emptyset$"), <u>stop</u>

Theorem 13. *The emptiness problem for systolic tree ω-languages is logspace complete for* **PSPACE**.

Proof. The procedure defined above shows that the problem is in **PSPACE**. We can reduce the emptiness problem for *SBTA* as acceptors of finitary languages, which is proved in [7] to be complete for **PSPACE**, to the emptiness problem for *SBTA* as acceptors of ω-languages.

For an *SBTA* $\mathscr{A}=\langle\Sigma,Q,in,f,F\rangle$, we define an *SBTA* $\mathscr{A}'=\langle\Sigma,Q',in',f',F'\rangle$ such that $\mathscr{L}(\mathscr{A})=\emptyset$ iff $\mathscr{L}_\omega(\mathscr{A}')=\emptyset$. Assume that Q, $\{q' : q\in Q\}$ and $\{acc\}$ are pairwise disjoint sets. The automaton $\mathscr{A}'$ is as follows:
- $Q'=Q\cup\{q' : q\in Q\}\cup\{acc\}$; $F'=\{acc\}$;
- $in'=\{\langle a,q\rangle : \langle a,q\rangle\in in, q\notin F\}\cup\{\langle a,acc\rangle : \langle a,q\rangle\in in, q\in F\}$;
- $f'=\{\langle q_1,q_2,q_3\rangle : \langle q_1,q_2,q_3\rangle\in f, q_3\notin F\}\cup$
  $\{\langle q_1,q_2,acc\rangle : \langle q_1,q_2,q_3\rangle\in f, q_3\in F\}\cup$
  $\{\langle acc,q,acc\rangle : q\in Q\}\cup$
  $\{\langle q_1,q_2,q_3'\rangle : \langle \#,q_2,q_3\rangle\in f\}\cup$
  $\{\langle q_1,q_2',q_3\rangle : \langle \#,q_2,q_3\rangle\in f\}\cup$
  $\{\langle q_1',q_2,q_3'\rangle : \langle q_1,q_2,q_3\rangle\in f, q_3\notin F\}\cup$
  $\{\langle q_1',q_2,acc\rangle : \langle q_1,q_2,q_3\rangle\in f, q_3\in F\}$.

The size of $\mathscr{A}'$ is polynomial in the size of $\mathscr{A}$ and $\mathscr{A}'$ can be constructed in logarithmic space. An ω-word α belongs to $\mathscr{L}_\omega(\mathscr{A}')$ iff there exists a successful run σ on α where the state *acc* occurs infinitely often. (Assume that k is the first position of σ where *acc* occurs). Then, $\sigma(k)=acc$ iff there exists m with $2^{k-1}<m\leqslant 2^k$ such that $\alpha(2^k-m, 2^k-1)$ belongs to $\mathscr{L}(\mathscr{A})$. $\square$

6. Expressive power

We investigate now the relationship between systolic tree ω-languages and regular ω-languages. We prove that the class of systolic tree ω-languages properly contains the class of regular ω-languages and that the class of deterministic systolic ω-languages properly contains that of deterministic regular ω-languages.

For a set of finite words $W\subseteq\Sigma^*$, let $\overrightarrow{W}$ be the set

$$\{\alpha\in\Sigma^\omega : \alpha(0,n)\in W \text{ for infinitely many } n\in\mathbb{N}\}.$$

It is well known (see [9]) that if V is a regular set, then $\overrightarrow{V}$ is a deterministic ω-language and that an ω-language L (on Σ) is a regular ω-language iff it is the finite union of ω-languages having the form $\overrightarrow{V}\cap(\Sigma^\omega-\overrightarrow{W})$ with V and W regular sets. We exploit this characterization for proving that the class of systolic tree ω-languages extends the class of ω-regular languages.

Lemma 14. *If U is a regular set, then $\overrightarrow{U}\in\mathscr{L}_\omega(DSBTA)$.*

Proof. For a regular set U, we construct a *DSBTA* $\mathscr{A}=\langle\Sigma, Q, in, f, F\rangle$ such that $\mathscr{L}(\mathscr{A})=\overrightarrow{U}$. Let $\mathscr{B}=\langle\Sigma,\{q_0\dots q_{n-1}\}, q_0, \Delta, \mathscr{F}\rangle$ be a deterministic finite state automaton[3] recognizing U and, without loss of generality, assume that for any q_i $(0\leqslant i\leqslant n-1)$ and for any $a\in\Sigma$, $\langle q_i, a, q_j\rangle\in\Delta$, for some q_j. The automaton $\mathscr{A}$ is defined as follows:

- $Q=(\{q_0\dots q_{n-1}\}\times\{0,1\}^n)^n$;
- $F=\{\langle\langle q_{i_0}, w_0\rangle,\dots,\langle q_{i_{n-1}}, w_{n-1}\rangle\rangle : w_i\in\{0,1\}^n,$
  there is $0\leqslant t\leqslant n-1$ s.t. $q_t\in\mathscr{F}$ and $w_0(t)=1\}$;
- $in=\{\langle a,\langle\langle q_{j_0}, w_0\rangle,\dots,\langle q_{j_{n-1}}, w_{n-1}\rangle\rangle\rangle : \langle q_i, a, q_{j_i}\rangle\in\Delta, w_i\in\{0,1\}^n,$
  $w_i(t)=1$ iff $t=i$, for all $0\leqslant i,t\leqslant n-1\}$;
- $f=\{\langle\langle\langle q_{j_0}, w_0\rangle,\dots,\langle q_{j_{n-1}}, w_{n-1}\rangle\rangle,\langle\langle q_{l_0}, u_0\rangle,\dots,\langle q_{l_{n-1}}, u_{n-1}\rangle\rangle,$
  $\langle\langle q_{l_{j_0}}, v_0\rangle,\dots,\langle q_{l_{j_{n-1}}}, v_{n-1}\rangle\rangle\rangle :$
  $v_i(t)=u_i(t)$ for $t\neq l_{j_i}, v_i(t)=1$ for $t=l_{j_i}$, for all $0\leqslant i,t\leqslant n-1\}$.

By construction, for each $w\in\Sigma^\star$ with $|w|=2^i$, $(i\geqslant 0)$, it holds that

$$\langle w,\langle\langle q_{j_0}, w_0\rangle,\dots,\langle q_{j_{n-1}}, w_{n-1}\rangle\rangle\rangle\in O_{\mathscr{A}}$$

iff the two following conditions hold for each $0\leqslant j,t\leqslant n-1$:

(1) the finite run, of the deterministic automaton $\mathscr{B}$, on w that starts from the state q_i leads to the state q_{j_i};

[3] The definition of finite state automata differs from that of Büchi automata only on the acceptance condition. A word w of length k is accepted if there exists a (finite) run σ on w such that $\sigma(k)$ is a final state.

(2) $w_j(t)=1$ iff there exists a finite run, of the deterministic automaton $\mathscr{B}$, on a prefix w' of w with $|w'| > \lfloor \frac{w}{2} \rfloor$ that starts from the state q_j and leads to the state q_t.

Therefore, $\alpha \in \mathscr{L}_\omega(\mathscr{A})$ iff $\alpha(0,n) \in U$, for infinitely many $n \in N$. $\square$

Theorem 15. *The following relations hold*:
1. $\mathscr{L}_\omega(DBA) \subset \mathscr{L}_\omega(DSBTA)$;
2. $\mathscr{L}_\omega(BA) \subset \mathscr{L}_\omega(SBTA)$;
3. $\mathscr{L}_\omega(DSBTA) \subset \mathscr{L}_\omega(SBTA)$.

Proof. (1) As stated in [9], $L \in \mathscr{L}_\omega(DBA)$ iff there is a regular set $V \in \Sigma^*$ such that $L = \overrightarrow{V}$. Now, by Lemma 14 we have that $L \in \mathscr{L}_\omega(DSBTA)$. The languages given in Example 1 show that the inclusion is proper.

(2) As stated in [9] L is a regular ω-language iff L is the finite union of sets having the form $\overrightarrow{W} \cap (\Sigma^\omega - \overrightarrow{W'})$ with regular $W, W' \subseteq \Sigma^*$. By Lemma 14 we have that $\overrightarrow{W}, \overrightarrow{W'} \in \mathscr{L}_\omega(DSBTA)$. The thesis follows since, by Theorems 4 and 11, the class $\mathscr{L}_\omega(SBTA)$ is closed under intersection and complement. The languages given in Example 1 show that the inclusion is proper.

(3) Relation $\mathscr{L}_\omega(DSBTA) \subseteq \mathscr{L}_\omega(SBTA)$ holds by definition. Now, it can be shown that the ω-language $L = (a+b)^\star . b^\omega$ belongs to $\mathscr{L}_\omega(BA)$ and that $L \notin \mathscr{L}_\omega(DSBTA)$. The inclusion is proper by the previous point. $\square$

7. Systolic tree ω-languages: the logical view

In this section we generalize Büchi Theorem to the class of systolic tree ω-languages. The sequential calculus is enriched by a unary function, called *power function*, which gives for a natural number $x > 0$ the natural number $x - x'$, where x' is the least power of 2 (with non-null coefficient) in the binary representation of x.

Definition 16. *Power function* $\overset{2}{\leftarrow}: \mathbb{N}^+ \to \mathbb{N}$ is the function such that

$$y = \overset{2}{\leftarrow}(x) \quad \text{iff} \quad x = \sum_{j=0}^{n} 2^{i_j}, \quad \text{with} \quad i_n > i_{n-1} > \cdots > i_0 \geqslant 0, \qquad \text{and } y = x - 2^{i_0}.$$

Note that x is a power of 2 iff $0 = \overset{2}{\leftarrow}(x)$ and then the function $\overset{2}{\leftarrow}$ is more expressive than the predicate "is a power of 2". The power function allows to associate, with each natural number, a set of natural numbers structured as a perfectly balanced binary tree.

For a number $z = 2^{k_n} + \cdots + 2^{k_1} + 2^{k_0}$, with $k_n > \cdots > k_1 > k_0 > 0$, the *left son* of z is the number

$$2^{k_n} + \cdots + 2^{k_1} + 2^{k_0 - 1}$$

and the *right son* of z is the number

$$2^{k_n} + \cdots + 2^{k_1} + 2^{k_0} + 2^{k_0 - 1}.$$

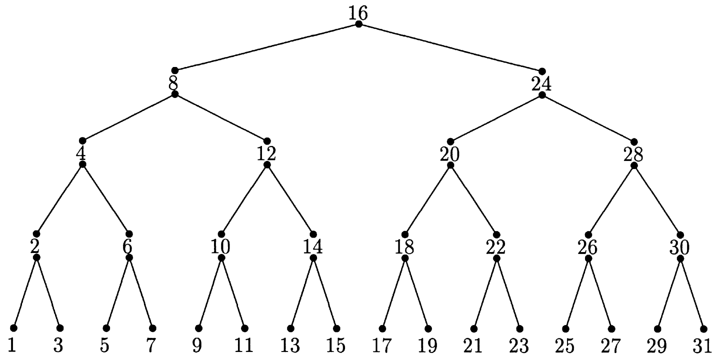

Fig. 3. The tree associated with number 16.

The tree associated with number 16 is shown in Fig. 3. In general, with a number $z = 2^k$, a tree is associated on the set of numbers $\{1, \ldots, 2^{k+1} - 1\}$ and odd numbers in that set are the leaves of the tree. The structure imposed in this way on natural numbers is analogous to the "structure" of a systolic computation (i.e. a systolic run, see Fig. 2) on an ω-word.

Now, it is easy to see that y is the right son of z iff

$$y = \max\{w : z = \overset{2}{\leftarrow} (w)\}$$

and y is the left son of z iff

$$y = \max\{w : w < y, \overset{2}{\leftarrow} (w) = \overset{2}{\leftarrow} (z)\}.$$

Let us consider the model-theoretic structures under which the formulas of the extended sequential calculus are interpreted.

Definition 17. An ω-word $\alpha \in \Sigma^\omega$ *induces* a model-theoretic structure $\underline{\alpha}$ (the *canonical interpretation under* α) having the form

$$\underline{\alpha} = \langle \omega, 0, +1, \overset{2}{\leftarrow}, <, (Q_a)_{a \in \Sigma} \rangle,$$

where $\langle \omega, 0, +1, < \rangle$ is the structure of natural numbers with 0, the successor function and the usual ordering, $\overset{2}{\leftarrow}$ is the power function, and Q_a (for $a \in \Sigma$) is the set $\{i \in \omega : \alpha(i) = a\}$.

We consider here the sequential calculus $S1S$ in the style of [9]. The extended sequential calculus $S1S^+$ differs from $S1S$ only in that it allows a reacher set of terms (i.e. terms are freely constructed by exploiting also the power function besides successor function).

Definition 18. For an alphabet Σ, the interpreted system $S1S^+_\Sigma$ is built up as follows:
- *Terms* are freely constructed from the constant 0 and the (first order) variables $x, y, z, \ldots$ by application of $+1$ (successor function) and $\overset{2}{\leftarrow}$ (power function);

– *Atomic formulas* are of the form $t=t'$, $t<t'$, $t\in X$, $t\in Q_a$ (for $a\in\Sigma$) where t and t' are terms and X is a set variable;
– $S1S^+_\Sigma$-*Formulas* are freely constructed from atomic formulas by using the connectives $\wedge$, $\vee$, $\neg$, $\Rightarrow$ and $\Leftrightarrow$ and by using the quantifiers $\exists$ and $\forall$ acting on either kind of variables.

We write $\phi(X_1,\ldots,X_n)$ to indicate that at most the variables $X_1,\ldots,X_n$ occur free in ϕ (i.e. they are not in the scope of a quantifier). Formulas without free variables are called *sentences*.

Given $\alpha\in\Sigma^\omega$ and a sentence ϕ, we write $\underline{\alpha}\models\phi$ if ϕ is satisfied[4] in $\underline{\alpha}$. The ω-*language defined by* a $S1S^+_\Sigma$-sentence ϕ is $\mathscr{L}(\phi)=\{\alpha\in\Sigma^\omega : \underline{\alpha}\models\phi\}$.

The calculus $S1S^+$ gives the logical counterpart of systolic ω-languages which have been defined in an operational way in Section 3.

Theorem 19. *An ω-language is definable in $S1S^+$ iff it belongs to $\mathscr{L}_\omega(SBTA)$.*

Proof. ($\subseteq$) Let $\mathscr{A}=\langle\Sigma,Q,in,f,F\rangle$ be a systolic automaton and assume, without loss of generality, that $Q=\{1,\ldots,m\}$. We prove that there exists an $S1S^+_\Sigma$-sentence ϕ such that $\mathscr{L}_\omega(\mathscr{A})=\mathscr{L}(\phi)$. In order to simulate a systolic run, we associate an element of Q with each natural number. For any state $1\leqslant i\leqslant m$, the set $Y_i\subseteq\mathbb{N}$ is the set of natural numbers the state i is associated with. Let us consider the systolic run on an ω-word α. Odd numbers (i.e. the leaves of the structure on $\mathbb{N}$) act as input nodes and a state i is associated with an odd x if $\langle\alpha(x-1),q_1\rangle\in in$, $\langle\alpha(x),q_2\rangle\in in$ and $\langle q_1,q_2,i\rangle\in f$. (See Eq. (3) in the definition of the sentence). If numbers x and y are the left and right sons, respectively, of z and states i and j are associated with x and y, respectively, then k such that $\langle i,j,k\rangle\in f$ is associated with z. (See Eq. (4) in the definition of the sentence). Finally, if node x is a power of 2, then the state associated with x is a state which results from processing the prefix $\alpha(0,2^{x+1}-1)$ of α. So, the acceptance condition can be expressed by requiring that an infinite number of powers of two are related with a final state. (See Eq. (5) in the definition of the sentence).

We introduce some short notations:

$x\xrightarrow{\mathrm{R}} i$ (i.e. the right son of x is associated with i) stands for

$$\exists y(x=\overset{2}{\leftarrow}(y)\wedge y\in Y_i\wedge\forall z((z\neq y\wedge x=\overset{2}{\leftarrow}(z))\Rightarrow z<y));$$

$x\xrightarrow{\mathrm{L}} i$ (i.e. the left son of x is associated with i) stands for

$$\exists y(x>y\wedge\overset{2}{\leftarrow}(y)=\overset{2}{\leftarrow}(x)\wedge y\in Y_i\wedge$$
$$\forall z((z\neq y\wedge x>z\wedge\overset{2}{\leftarrow}(z)=\overset{2}{\leftarrow}(x))\Rightarrow z<y));$$

$Odd(X)$ stands for

$$0\notin X\wedge 1\in X\wedge\forall x\,(x>1\Rightarrow(x\in X\Leftrightarrow\exists y\,(y+2=x\wedge y\in X))).$$

[4] The notion "ϕ is satisfied in $\underline{\alpha}$" is standard and is not formally given here.

A sentence ϕ such that $\mathscr{L}(\phi)=\mathscr{L}_\omega(\mathscr{A})$ is the following:
$\exists Y_1,\ldots,Y_m$ (

$$\bigwedge_{i\neq j}\neg\exists y(y\in Y_i\wedge y\in Y_j)\wedge \tag{2}$$

$$\exists X\left(Odd(X)\wedge\forall x\left(x\in X\Rightarrow\left(\bigwedge_{a,b\in\Sigma}(x-1\in Q_a\wedge x\in Q_b)\Rightarrow \bigvee_{\{i:\langle j,k,i\rangle\in f,\langle a,j\rangle,\langle b,k\rangle\in in\}} x\in Y_i\right)\right)\right)\wedge \tag{3}$$

$$\forall z\left(\bigwedge_{i,j\in Q}\left((z\xrightarrow{L}i\wedge z\xrightarrow{R}j)\Rightarrow\bigvee_{\{k:\langle i,j,k\rangle\in f\}}z\in Y_k\right)\right)\wedge \tag{4}$$

$$\bigvee_{i\in F}\forall x\exists y\,(x<y\wedge 0=\overset{2}{\leftarrow}(y)\wedge y\in Y_i) \tag{5}$$

).

($\supseteq$) The implication of the theorem is proved by exploiting the same technique used in [9]. For technical convenience, predicate symbols Q_a (for $a\in\Sigma$) are cancelled and free set variables are used in their place. So, formulas $\phi(X_1,\ldots,X_n)$ are considered, where no symbol Q_a occurs, and they are interpreted over ω-words over the special alphabet $\{0,1\}^n$. If $\alpha\in(\{0,1\}^n)^\omega$, then $\underline{\alpha}\models x\in X_k$ iff the xth symbol of α has 1 in its kth component. For a suitable n, symbols in Σ can be binary encoded and instead of atomic formula $x\in Q_a$, the corresponding conjunction consisting of formulas $x\in X_k$ and $\neg x\in X_k$ can be written. So, given a $S1S^+_\Sigma$-sentence ϕ, the thesis is proved for the corresponding formula $\phi(X_1,\ldots,X_n)$ interpreted in ω-words over $\{0,1\}^n$ (we also say that $\phi(X_1,\ldots,X_n)$ is an $S1S^+$-formula). An $S1S^+$-formula $\phi(X_1,\ldots,X_n)$ of $S1S^+$ can be reduced to an equivalent formula of a formalism simpler than $S1S^+$ (denoted by $S1S^+_0$) where the only possible terms are second order variables and the atomic formulas have the form:

1. $X_i\subseteq X_j$ ("X_i is a subset of X_j");
2. $Succ(X_i,X_j)$ ("X_i, X_j are singletons $\{x\}$, $\{y\}$, resp., with $x+1=y$");
3. $Power(X_i,X_j)$ ("X_i, X_j are singletons $\{y\}$, $\{x\}$, resp., with $\overset{2}{\leftarrow}(x)=y$").

(The reduction from $S1S^+$ to $S1S^+_0$ is a trivial extension of the reduction from $S1S$ to $S1S_0$: we refer to [9] for the details). We show now by induction on the structure of the $S1S^+_0$-formula $\phi(X_1,\ldots,X_n)$ that there exists an *SBTA* $\mathscr{A}$ such that $\mathscr{L}_\omega(\mathscr{A})=\mathscr{L}(\phi(X_1,\ldots,X_n))$.

Basic case. Let us consider the atomic formula $\phi(X_1,X_2)=Power(X_1,X_2)$. An automaton $\mathscr{A}=\langle\Sigma,Q,in,f,F\rangle$ such that $\mathscr{L}_\omega(\mathscr{A})=\mathscr{L}(\phi)$ is defined as follows:

– $\Sigma=\{0,1\}^2$; $Q=\{0,1\}^2\cup\{1\}$ and $F=\{1\}$;
– *in* is the identity function;

– $f = \{\langle\langle 1,0\rangle,\langle 0,0\rangle,\langle 1,0\rangle\rangle,\langle\langle 0,1\rangle,\langle 0,0\rangle,\langle 0,1\rangle\rangle,\langle\langle 0,0\rangle,\langle 0,0\rangle,\langle 0,0\rangle\rangle,$
$\langle\langle 1,0\rangle\langle 0,1\rangle,1\rangle,\langle 1,\langle 0,0\rangle,1\rangle,\langle\langle 0,0\rangle,1,1\rangle\}$.

It is easy to check that $\alpha \in \mathscr{L}_\omega(\mathscr{A})$ iff $\alpha(i) = \langle 1,0\rangle$ and $\alpha(j) = \langle 0,1\rangle$ for i and j such that $i = \overset{2}{\leftarrow}(j)$ and $\alpha(k) = \langle 0,0\rangle$, for all k such that $k \neq i$ and $k \neq j$. Any other kind of atomic $S1S_0^+$-formula ϕ, is a $S1S_0$-formula. Therefore, by Büchi Theorem, $\mathscr{L}(\phi) \in \mathscr{L}_\omega(BA)$ and by Theorem 15 we have that $\mathscr{L}(\phi) \in \mathscr{L}_\omega(SBTA)$.

Induction step. For the induction step it suffices to treat connectives $\neg$, $\vee$ and $\exists$. Cases $\neg$ and $\vee$ are apparent by closure of $\mathscr{L}_\omega(SBTA)$ under complementation and union. Concerning $\exists$, we have to show closure of $\mathscr{L}_\omega(SBTA)$ under projection. Let us assume that, for an $S1S_0^+$-formula $\phi(X_1,\dots,X_n)$, there exists an $SBTA$ $\mathscr{A} = (\{0,1\}^n, Q, in, f, F\rangle$ such that $\mathscr{L}(\phi) = \mathscr{L}_\omega(\mathscr{A})$. Consider the formula $\phi'(X_1,\dots,X_{i-1},X_{i+1},\dots,X_n) = \exists X_i \phi(X_1,\dots,X_n)$. An automaton $\mathscr{A}' = (\{0,1\}^{n-1}, Q, in', f, F\rangle$ such that $\mathscr{L}_\omega(\mathscr{A}') = \mathscr{L}(\phi'(X_1,\dots,X_{i-1},X_{i+1},\dots,X_n)$ is the one having in' defined as follows:

$$in' = \{\langle\langle x_1,\dots,x_{i-1},x_{i+1},\dots,x_n\rangle,q\rangle : \langle\langle x_1,\dots,x_{i-1},0,x_{i+1},\dots,x_n\rangle,q\rangle \in in\} \cup$$
$$\{\langle x_1,\dots,x_{i-1},x_{i+1},\dots,x_n\rangle,q\rangle : \langle\langle x_1,\dots,x_{i-1},1,x_{i+1},\dots,x_n\rangle,q\rangle \in in\}. \qquad \square$$

The decidability of $S1S^+$ immediately follows from the decidability of emptiness problem for systolic tree ω-languages.

Corollary 20. *Truth of sentences of $S1S^+$ is decidable.*

The extension of $S1S$ by the predicate "is a power of 2" has been proved decidable by Elgot and Rabin in [3]. We believe that $S1S^+$ is not equivalent to that extension and that the method proposed by Elgot and Rabin cannot be applied for obtaining $S1S^+$.

References

[1] J.R. Büchi, On a decision method in restricted second order arithmetic, in: Proc. Internat. Congress on Logic, Methodology and Philosophy of Science, Stanford Univ. Press, Stanford, 1960, pp. 1–11.
[2] K. Culik II, A. Salomaa, D. Wood, Systolic tree acceptors, R.A.I.R.O Informatique Théorique 18 (1984) 53–69.
[3] C.C. Elgot, M.O. Rabin, Decidability and undecidability of extensions of second (first) order theory of (generalized) successor, J. Symbol. Logic 31 (1966) 169–181.
[4] E.A. Emerson, C.D. Lei, Modalities for model checking: branching time strikes back, in: Proc. 12th ACM Symp. on Principles of Programming Languages, New Orleans, 1985, pp. 84–96.
[5] J. Gruska, Synthesis, structure and power of systolic computations, Theoret. Comput. Sci. 71 (1990) 47–78.
[6] A. Monti, A. Peron, Systolic tree ω-languages, in: Proc. STACS'95, Lecture Notes in Computer Science, vol. 900, Springer, Berlin, 1995, pp. 131–142.
[7] A. Monti, A. Roncato, On the complexity of some reachability problems, Inform. Process. Lett. 53 (1995) 11–16.
[8] A.P. Sistla, M.Y. Vardi, P. Wolper, The complementation problem for Büchi automata with application to temporal logic, Theoret. Comput. Sci. 49 (1987) 217–237.
[9] W. Thomas, Automata on infinite objects, in: J. van Leeuwen (Ed.), Handbook of Theoretical Computer Science, vol. B, Elsevier, Amsterdam, 1990, pp. 133–191.