

Reducing the time complexity of testing for local threshold testability

A.N. Trahtman

Bar-Ilan University, Dep. of Math., 52900, Ramat Gan, Israel

Abstract

A locally threshold testable language L is a language with the property that for some nonnegative integers k and l and for some word u from L , a word v belongs to L iff

- (1) the prefixes [suffixes] of length $k - 1$ of words u and v coincide,
- (2) the numbers of occurrences of every factor of length k in both words u and v are either the same or greater than $l - 1$.

A deterministic finite automaton is called locally threshold testable if the automaton accepts a locally threshold testable language for some l and k .

New necessary and sufficient conditions for a deterministic finite automaton to be locally threshold testable are found. On the basis of these conditions, we modify the algorithm to verify local threshold testability of the automaton and to reduce the time complexity of the algorithm. The algorithm is implemented as a part of the C/C^{++} package TESTAS. <http://www.cs.biu.ac.il/~trakht/Testas.html>.

Key words: automaton, threshold locally testable, graph, algorithm

1 Introduction

The locally threshold testable languages introduced by Beauquier and Pin [1] now have various applications [8], [15], [16]. In particular, stochastic locally threshold testable languages, also known as $n - grams$, are used in pattern recognition and in speech recognition, both in acoustic-phonetics decoding and in language modelling [15]. These languages generalize the concept of local testability [2], [7], which can be considered as a special case of local l -threshold testability for $l = 1$.

Email address: trakht@macs.biu.ac.il (A.N. Trahtman).

Necessary and sufficient conditions of local testability [5] form a basis of polynomial-time algorithms for the local testability problem [5], [13]. The algorithms were implemented [3], [4], [13].

Necessary and sufficient conditions of local threshold testability for deterministic finite automata (DFA) found in [1] are based on a syntactic characterization of locally threshold testable languages [10]. A polynomial-time algorithm of order $O(|\Gamma|^5|\Sigma|)$ for the local threshold testability problem based on some other kind of necessary and sufficient conditions was described in [13] and implemented. We modify the last necessary and sufficient conditions and reduce in that way the order of the algorithm for local threshold testability to $O(|\Gamma|^4|\Sigma|)$. The algorithm was successfully implemented.

2 Notation and definitions

Let Σ^+ [Σ^*] denote the free semigroup [monoid] over an alphabet Σ .

If $w \in \Sigma^+$, let $|w|$ denote the length of w . Let k be a positive integer. Let $i_k(w)$ [$t_k(w)$] denote the prefix [suffix] of w of length k or w if $|w| < k$. Let $F_{k,j}(w)$ denote the set of factors of w of length k with at least j occurrences. A language L is called *l -threshold k -testable* if there is an alphabet Σ such that for all $u, v \in \Sigma^+$, if $i_{k-1}(u) = i_{k-1}(v)$, $t_{k-1}(u) = t_{k-1}(v)$ and $F_{k,j}(u) = F_{k,j}(v)$ for all $j \leq l$, then either both u and v are in L or neither is in L .

An automaton is *l -threshold k -testable* if the automaton accepts a l -threshold k -testable language. A language L [an automaton] is *locally threshold testable* if it is l -threshold k -testable for some k and l .

Let us now consider the transition graph of a DFA.

The action of a word $v \in \Sigma^*$ on a state \mathbf{q} is denoted by $\mathbf{q}v$. Thus $\mathbf{q}v$ is the state reached by the unique path of label v starting at \mathbf{q} .

A state \mathbf{p} is a *cycle state* if, for some $e \in \Sigma^+$, $\mathbf{p}e = \mathbf{p}$.

A maximal strongly connected component of a directed graph will be denoted for brevity by *SCC*.

We shall write $\mathbf{p} \succeq \mathbf{q}$ if \mathbf{q} is reachable from \mathbf{p} (that is, if $\mathbf{p}v = \mathbf{q}$ for some word $v \in \Sigma^*$) and $\mathbf{p} \sim \mathbf{q}$ if $\mathbf{p} \succeq \mathbf{q}$ & $\mathbf{q} \succeq \mathbf{p}$ (that is, if \mathbf{p} and \mathbf{q} are in the same *SCC*).

The number of vertices of a graph Γ is denoted by $|\Gamma|$.

An oriented labelled graph is *complete* if any of its vertex has outgoing edge with any label from the alphabet of labels. A non-complete graph can be completed by adding a sink state and then adding lacking edges from corresponding vertices to the sink state.

The direct product Γ^k of k copies of a directed labelled graph Γ over an alphabet Σ consists of vertices $(\mathbf{p}_1, \dots, \mathbf{p}_k)$ and edges $(\mathbf{p}_1, \dots, \mathbf{p}_k) \rightarrow (\mathbf{p}_1\sigma, \dots, \mathbf{p}_k\sigma)$ labelled by σ . Here \mathbf{p}_i are vertices from Γ , $\sigma \in \Sigma$.

3 The necessary and sufficient conditions of local threshold testability

Let us formulate the result of Beauquier and Pin [1] in the following form:

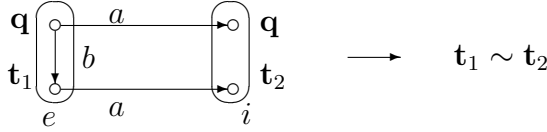
Theorem 1 [1] *A language L is locally threshold testable if and only if the syntactic semigroup S of L is aperiodic and for any two idempotents e, f and elements a, u, b of S , we have*

$$eafuebf = ebfueaf. \quad (1)$$

We now consider a fixed locally threshold testable DFA with state transition graph Γ and transition semigroup S .

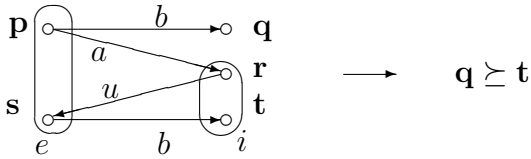
Lemma 2 [5] [13] *Let (\mathbf{p}, \mathbf{q}) be a cycle state of Γ^2 . If $\mathbf{p} \sim \mathbf{q}$, then $\mathbf{p} = \mathbf{q}$.*

Lemma 3 *Let $(\mathbf{q}, \mathbf{t}_1)$ and $(\mathbf{q}, \mathbf{t}_2)$ be cycle states of Γ^2 . If $(\mathbf{q}, \mathbf{t}_1) \succeq (\mathbf{q}, \mathbf{t}_2)$ and $\mathbf{q} \succeq \mathbf{t}_1$ then $\mathbf{t}_1 \sim \mathbf{t}_2$.*



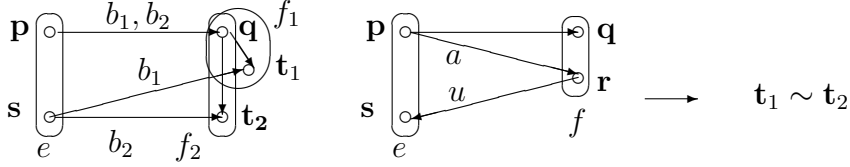
Proof. One has $(\mathbf{q}, \mathbf{t}_1)e = (\mathbf{q}, \mathbf{t}_1)$, $(\mathbf{q}, \mathbf{t}_2)i = (\mathbf{q}, \mathbf{t}_2)$, $(\mathbf{q}, \mathbf{t}_1)a = (\mathbf{q}, \mathbf{t}_2)$, $\mathbf{q}b = \mathbf{t}_1$ for some idempotents e, i and elements a, b from S . The substitution ai in place of a and e in place of f and u in (1) implies $eaiebe = ebeaie$. Therefore $\mathbf{t}_2e = \mathbf{t}_2ie = \mathbf{t}_1eaie = \mathbf{q}ebeaie = \mathbf{q}eaiebe$. Thus $\mathbf{t}_2e = \mathbf{q}eaiebe = \mathbf{q}ebe = \mathbf{t}_1e = \mathbf{t}_1$. So $\mathbf{t}_2 \succeq \mathbf{t}_1$. We have $\mathbf{t}_1a = \mathbf{t}_2$, whence $\mathbf{t}_1 \sim \mathbf{t}_2$.

Lemma 4 *Let $\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}$ be states such that (\mathbf{p}, \mathbf{s}) and (\mathbf{r}, \mathbf{t}) are cycle states of Γ^2 . If $(\mathbf{p}, \mathbf{s}) \succeq (\mathbf{q}, \mathbf{t})$ and $\mathbf{p} \succeq \mathbf{r} \succeq \mathbf{s}$, then $\mathbf{q} \succeq \mathbf{t}$.*



Proof. One has $(\mathbf{p}, \mathbf{s})e = (\mathbf{p}, \mathbf{s})$ and $(\mathbf{r}, \mathbf{t})i = (\mathbf{r}, \mathbf{t})$ for some idempotents $e, i \in S$. Furthermore, $(\mathbf{p}, \mathbf{s})b = (\mathbf{q}, \mathbf{t})$, $\mathbf{p}a = \mathbf{r}$ and $\mathbf{r}u = \mathbf{s}$ for some elements $a, u, b \in S$. In view of (1), $\mathbf{t} = \mathbf{p}eaiuebi = \mathbf{p}ebiueai$. Thus $\mathbf{t} = \mathbf{p}ebiueai = \mathbf{q}iuebi$, whence $\mathbf{q} \succeq \mathbf{t}$.

Lemma 5 *Let (\mathbf{q}, \mathbf{r}) , (\mathbf{p}, \mathbf{s}) , $(\mathbf{q}, \mathbf{t}_1)$ and $(\mathbf{q}, \mathbf{t}_2)$ be cycle states of the graph Γ^2 such that $(\mathbf{p}, \mathbf{s}) \succeq (\mathbf{q}, \mathbf{t}_i)$, $\mathbf{q} \succeq \mathbf{t}_i$ for $i = 1, 2$ and $\mathbf{p} \succeq \mathbf{r} \succeq \mathbf{s}$. Then $\mathbf{t}_1 \sim \mathbf{t}_2$.*



Proof. One has $(\mathbf{p}, \mathbf{s})e = (\mathbf{p}, \mathbf{s})$, $(\mathbf{q}, \mathbf{r})f = (\mathbf{q}, \mathbf{r})$, $(\mathbf{q}, \mathbf{t}_1)f_1 = (\mathbf{q}, \mathbf{t}_1)$, $(\mathbf{q}, \mathbf{t}_2)f_2 = (\mathbf{q}, \mathbf{t}_2)$ for some idempotents $e, f, f_1, f_2 \in S$. Furthermore, $(\mathbf{p}, \mathbf{s})b_1 = (\mathbf{q}, \mathbf{t}_1)$, $(\mathbf{p}, \mathbf{s})b_2 = (\mathbf{q}, \mathbf{t}_2)$, $\mathbf{p}a = \mathbf{r}$ and $\mathbf{r}u = \mathbf{s}$ for some elements $a, u, b_1, b_2 \in S$.

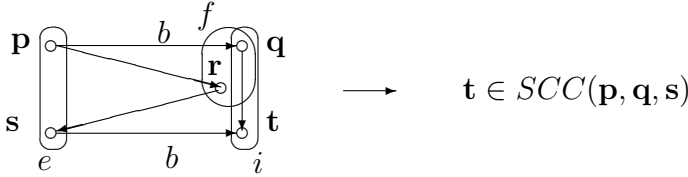
Let us consider the state $\mathbf{t}_i f$ ($i = 1, 2$) where the idempotent f is right unit for (\mathbf{q}, \mathbf{r}) . The states $(\mathbf{q}, \mathbf{t}_i)$ and $(\mathbf{q}, \mathbf{t}_i f)$ are cycle states, $\mathbf{q} \succeq \mathbf{t}_i$, $(\mathbf{q}, \mathbf{t}_i) \succeq (\mathbf{q}, \mathbf{t}_i f)$, whence by Lemma 3, $\mathbf{t}_i f \sim \mathbf{t}_i$ for any such f . So $\mathbf{t}_1 f \sim \mathbf{t}_1$ and $\mathbf{t}_2 f \sim \mathbf{t}_2$.

The equality of local threshold testability (1) implies $\mathbf{t}_1 f = \mathbf{s}b_1 f = \mathbf{r}ue b_1 f = \mathbf{p}e a f u e b_1 f = \mathbf{p}e b_1 f u e a f$. Furthermore, $\mathbf{p}b_1 = \mathbf{p}b_2 = \mathbf{q}$, whence $\mathbf{t}_1 f = \mathbf{p}e b_1 f u e a f = \mathbf{p}e b_2 f u e a f = \mathbf{p}e a f u e b_2 f = \mathbf{t}_2 f$. So $\mathbf{t}_2 f = \mathbf{t}_1 f$. We have $\mathbf{t}_1 f \sim \mathbf{t}_1$ and $\mathbf{t}_2 f \sim \mathbf{t}_2$, whence $\mathbf{t}_2 \sim \mathbf{t}_1$. \square

If $\mathbf{p}, \mathbf{q}, \mathbf{s}$ are states of Γ and there exists some state \mathbf{r} such that (\mathbf{q}, \mathbf{r}) and (\mathbf{p}, \mathbf{s}) are cycle states of Γ^2 , $\mathbf{p} \succeq \mathbf{q}$, and $\mathbf{p} \succeq \mathbf{r} \succeq \mathbf{s}$, then the non-empty set

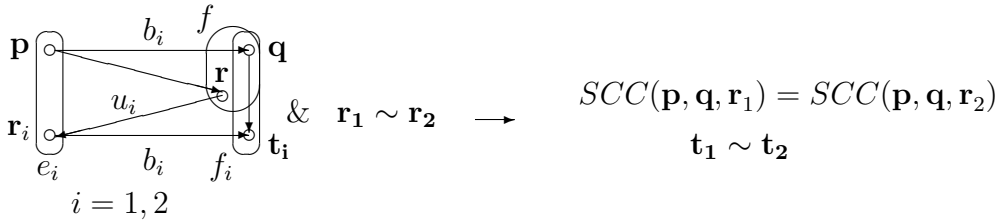
$$T = \{t \mid (\mathbf{p}, \mathbf{s}) \succeq (\mathbf{q}, \mathbf{t}), \mathbf{q} \succeq \mathbf{t} \text{ and } (\mathbf{q}, \mathbf{t}) \text{ is a cycle state}\}$$

by the previous Lemma is a subset of some *SCC* from transition graph Γ of locally threshold testable automaton. This *SCC* will be denoted by $SCC(\mathbf{p}, \mathbf{q}, \mathbf{s})$. In the case that \mathbf{r} does not exist or T is empty, let $SCC(\mathbf{p}, \mathbf{q}, \mathbf{s})$ be empty.



By Lemma 5, $SCC(\mathbf{p}, \mathbf{q}, \mathbf{s})$ is well-defined for transition graphs of locally threshold testable automata.

Lemma 6 Let $(\mathbf{p}, \mathbf{r}_1)$ and $(\mathbf{p}, \mathbf{r}_2)$ be cycle states of the graph Γ^2 . Suppose that $\mathbf{r}_1 \sim \mathbf{r}_2$, $\mathbf{q} \succeq \mathbf{t}_i$, $\mathbf{p} \succeq \mathbf{r} \succeq \mathbf{r}_i$ ($i = 1, 2$) for some \mathbf{r} such that (\mathbf{q}, \mathbf{r}) is a cycle state. Then $\mathbf{t}_1 \sim \mathbf{t}_2$ and $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_2)$.



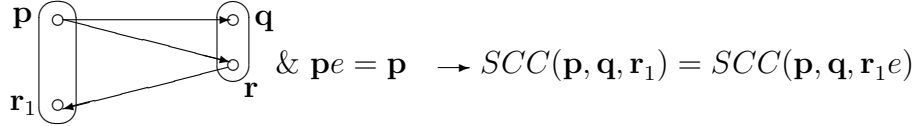
Proof. One has $(\mathbf{p}, \mathbf{r}_1)e_1 = (\mathbf{p}, \mathbf{r}_1)$, $(\mathbf{p}, \mathbf{r}_2)e_2 = (\mathbf{p}, \mathbf{r}_2)$, $(\mathbf{q}, \mathbf{r})f = (\mathbf{q}, \mathbf{r})$ for some idempotents $e_1, e_2, f \in S$, $\mathbf{p} \succeq \mathbf{r}$ and $(\mathbf{p}, \mathbf{r}_1)b_1 = (\mathbf{q}, \mathbf{t}_1)$, $(\mathbf{p}, \mathbf{r}_2)b_2 = (\mathbf{q}, \mathbf{t}_2)$, $\mathbf{r}u_i = \mathbf{r}_i$ for some elements $u_i, b_1, b_2 \in S$.

From $(\mathbf{p}, \mathbf{r}_1)e_2 = (\mathbf{p}, \mathbf{r}_1 e_2)$, by Lemma 3, it follows that $\mathbf{r}_1 \sim \mathbf{r}_1 e_2$. Notice

that $\mathbf{r}_2 e_2 = \mathbf{r}_2 \sim \mathbf{r}_1$, whence $\mathbf{r}_2 \sim \mathbf{r}_1 e_2$. Therefore, by Lemma 2, $\mathbf{r}_2 = \mathbf{r}_1 e_2$. Furthermore, $(\mathbf{p}, \mathbf{r}_1) e_2 b_2 = (\mathbf{q}, \mathbf{r}_1 e_2 b_2) = (\mathbf{q}, \mathbf{r}_2 b_2) = (\mathbf{q}, \mathbf{t}_2)$. Thus $(\mathbf{p}, \mathbf{r}_1) \succeq (\mathbf{q}, \mathbf{t}_2)$. Then $(\mathbf{p}, \mathbf{r}_1) \succeq (\mathbf{q}, \mathbf{t}_1)$ in view of $(\mathbf{p}, \mathbf{r}_1) b_1 = (\mathbf{q}, \mathbf{t}_1)$. Now by Lemma 5, the states $\mathbf{t}_1, \mathbf{t}_2$ belong to $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1)$ and $\mathbf{t}_1 \sim \mathbf{t}_2$. Let us notice that the state \mathbf{t}_2 belongs to $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_2)$ too. Hence by Lemma 5, $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_2)$.

Lemma 7 *If*

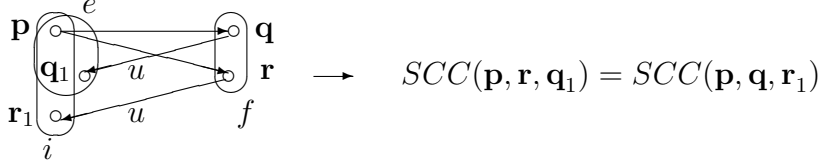
$\mathbf{p} \succeq \mathbf{q}$, $\mathbf{p} \succeq \mathbf{r}$ and $\mathbf{p}e = \mathbf{p}$ for an idempotent $e \in S$,
the state (\mathbf{q}, \mathbf{r}) is a cycle state of the graph Γ^2 ,
there exists a state \mathbf{r}_1 such that $(\mathbf{p}, \mathbf{r}_1)$ is a cycle state and $\mathbf{r} \succeq \mathbf{r}_1$,
then $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1 e) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1)$.



Proof. One has $\mathbf{p} \succeq \mathbf{r} \succeq \mathbf{r}_1$. Then $(\mathbf{p}, \mathbf{r}_1) \succeq (\mathbf{p}, \mathbf{r}_1) e = (\mathbf{p}, \mathbf{r}_1 e)$ and both these states are cycle states. Therefore, by Lemma 3, $\mathbf{r}_1 e \sim \mathbf{r}_1$. Lemma 6 for $\mathbf{r}_2 = \mathbf{r}_1 e$ implies now $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1 e) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1)$.

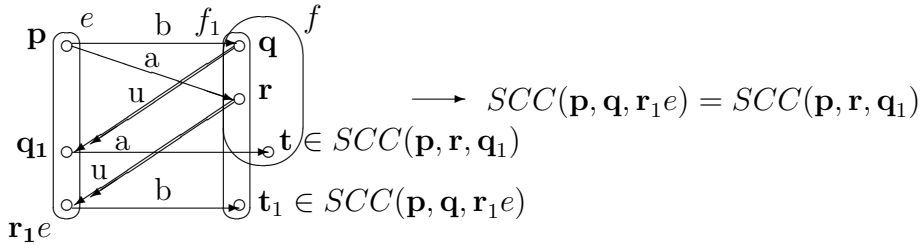
Theorem 8 *DFA A with state transition complete graph Γ (or completed by a sink state) is locally threshold testable iff*

- 1) for every cycle state (\mathbf{p}, \mathbf{q}) of Γ^2 , $\mathbf{p} \sim \mathbf{q}$ implies $\mathbf{p} = \mathbf{q}$,
- 2) for every states $\mathbf{p}, \mathbf{q}, \mathbf{t}, \mathbf{s}$ of Γ such that
 - (\mathbf{p}, \mathbf{s}) is a cycle state,
 - $(\mathbf{p}, \mathbf{s}) \succeq (\mathbf{q}, \mathbf{t})$,
 - $\mathbf{p} \succeq \mathbf{r} \succeq \mathbf{s}$ and (\mathbf{r}, \mathbf{t}) is a cycle state for some \mathbf{r} ,
it holds $\mathbf{q} \succeq \mathbf{t}$. (see figure to Lemma 4)
- 3) for every states $\mathbf{p}, \mathbf{q}, \mathbf{r}$, $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r})$ is well defined,
- 4) for every four states $\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{q}_1$ such that
 - $(\mathbf{p}, \mathbf{q}_1)$ and (\mathbf{q}, \mathbf{r}) are cycle states of the graph Γ^2 ,
 - $\mathbf{p} \succeq \mathbf{q}$ and $\mathbf{p} \succeq \mathbf{r}$,
 - for some state \mathbf{r}_1 such that $(\mathbf{p}, \mathbf{r}_1)$ is a cycle state and $(\mathbf{q}, \mathbf{r}) \succeq (\mathbf{q}_1, \mathbf{r}_1)$,
it holds $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1) = SCC(\mathbf{p}, \mathbf{r}, \mathbf{q}_1)$.



Proof. Let \mathbf{A} be a locally threshold testable DFA. Condition 1 follows in this case from Lemma 2. Condition 2 follows from Lemma 4. Condition 3 follows from Lemma 5.

Let us check the last condition. For some idempotent e , it holds $(\mathbf{p}, \mathbf{q}_1)e = (\mathbf{p}, \mathbf{q}_1)$. By Lemma 7, $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1e) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1)$. Therefore let us compare $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1e)$ and $SCC(\mathbf{p}, \mathbf{r}, \mathbf{q}_1)$.



One has $\mathbf{t}_1f = \mathbf{r}_1ebf = \mathbf{p}eafuebf$. Then by (1) $\mathbf{p}eafuebf = \mathbf{p}ebfueaf = \mathbf{q}ueaf = \mathbf{q}_1af = \mathbf{t}$. So $\mathbf{t}_1 \succeq \mathbf{t}$. Analogously, $\mathbf{t} \succeq \mathbf{t}_1$. Therefore $\mathbf{t}_1 \sim \mathbf{t}$, whence $SCC(\mathbf{p}, \mathbf{r}, \mathbf{q}_1) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1e)$. Consequently, $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1) = SCC(\mathbf{p}, \mathbf{r}, \mathbf{q}_1)$. Conversely, suppose that all four conditions of the theorem hold. Our aim is to prove the local threshold testability of DFA. For this aim, let us consider an arbitrary state \mathbf{v} , arbitrary elements a, u, b and idempotents e, f from the syntactic semigroup S of the automaton. We must to prove that $\mathbf{v}eafuebf = \mathbf{v}ebfueaf$ (Theorem 1).

Let us denote $\mathbf{p} = \mathbf{v}e$, $\mathbf{q} = \mathbf{v}bf$, $\mathbf{q}_1 = \mathbf{v}bfue$, $\mathbf{t} = \mathbf{v}ebfueaf$, $\mathbf{r} = \mathbf{v}eaf$, $\mathbf{r}_1 = \mathbf{v}eafue$, $\mathbf{t}_1 = \mathbf{v}eafuebf$.

We have $(\mathbf{p}, \mathbf{r}_1) \succeq (\mathbf{q}, \mathbf{t}_1)$, the states $(\mathbf{p}, \mathbf{r}_1)$, $(\mathbf{q}, \mathbf{t}_1)$ and $(\mathbf{r}, \mathbf{t}_1)$ are cycle states, $\mathbf{p} \succeq \mathbf{r} \succeq \mathbf{r}_1$. Therefore by condition 2, for $\mathbf{r}_1 = \mathbf{s}$, $\mathbf{q} \succeq \mathbf{t}_1$. Now $\mathbf{t}_1 \in SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1)$. Analogously $\mathbf{t} \in SCC(\mathbf{p}, \mathbf{r}, \mathbf{q}_1)$. The state $(\mathbf{p}, \mathbf{q}_1)$ is a cycle state and $(\mathbf{q}, \mathbf{r})ue = (\mathbf{q}_1, \mathbf{r}_1)$. Hence condition 4 implies $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1) = SCC(\mathbf{p}, \mathbf{r}, \mathbf{q}_1)$. These sets are well-defined, whence by condition 3, $\mathbf{t}_1 \sim \mathbf{t}$. Both these states have common right unit f . Consequently, $(\mathbf{t}, \mathbf{t}_1)$ is a cycle state. Now by condition 1, $\mathbf{t}_1 = \mathbf{t}$. Thus $\mathbf{v}eafuebf = \mathbf{v}ebfueaf$ is true for an arbitrary state \mathbf{v} and the identity $eafuebf = ebfueaf$ of local threshold testability holds.

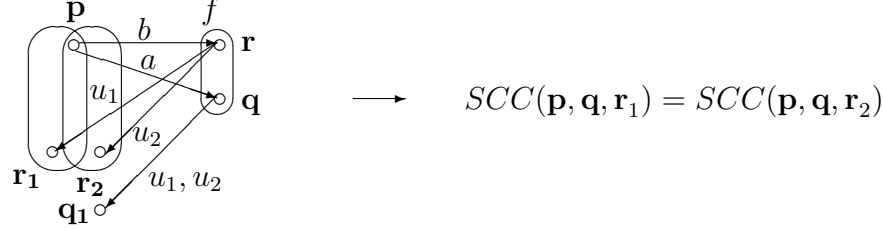
It remains now to prove the aperiodicity of S . Let \mathbf{p} be an arbitrary state and let s be an arbitrary element of S . The semigroup S is finite, whence for some integers k and m , it holds $s^k = s^{k+m}$. Let us consider the states $\mathbf{p}s^k$ and $\mathbf{p}s^{k+1}$. We have $\mathbf{p}s^k \succeq \mathbf{p}s^{k+1}$ and, in view $s^k = s^{k+m} = s^{k+1}s^{m-1}$, it holds $\mathbf{p}s^{k+1} \succeq \mathbf{p}s^k$. Thus $\mathbf{p}s^{k+1} \sim \mathbf{p}s^k$. Some power of s is an idempotent and a right unit for both these states. Then by condition 1, $\mathbf{p}s^k = \mathbf{p}s^{k+1}$. Therefore

S is aperiodic, and thus the automaton is locally threshold testable.

Lemma 9 *Let $P(\mathbf{q}, \mathbf{r})$ be a non-empty set of cycle states of a locally threshold testable DFA such that $\mathbf{p} \succeq \mathbf{q}$ and $\mathbf{p} \succeq \mathbf{r}$ for a cycle state (\mathbf{q}, \mathbf{r}) .*

By $\mathbf{r}_2 \rho_r \mathbf{r}_1$ we denote the case that for a pair of cycle states $(\mathbf{p}, \mathbf{r}_1)$ and $(\mathbf{p}, \mathbf{r}_2)$, it holds $(\mathbf{q}, \mathbf{r}) \succeq (\mathbf{q}_1, \mathbf{r}_1)$ and $(\mathbf{q}, \mathbf{r}) \succeq (\mathbf{q}_1, \mathbf{r}_2)$.

Then $\mathbf{r}_1 \rho_r \mathbf{r}_2$ implies $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_2)$ for any $\mathbf{p} \in P(\mathbf{q}, \mathbf{r})$.



Proof. One has $(\mathbf{q}, \mathbf{r})f = (\mathbf{q}, \mathbf{r})$, $(\mathbf{q}, \mathbf{r})u_1 = (\mathbf{q}_1, \mathbf{r}_1)$, $(\mathbf{q}, \mathbf{r})u_2 = (\mathbf{q}_1, \mathbf{r}_2)$, $\mathbf{p}a = \mathbf{q}$, $\mathbf{p}b = \mathbf{r}$, $\mathbf{p}e = \mathbf{p}$ for some idempotents e, f and elements u_i, a, b from S . So $\mathbf{q}_1 = \mathbf{p}eafu_2 = \mathbf{p}eafu_1$, $\mathbf{p}ebfu_1 = \mathbf{r}_1$, $\mathbf{p}ebfu_2 = \mathbf{r}_2$. For the state \mathbf{r}_1eaf from $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1e)$, it holds $\mathbf{r}_1eaf = \mathbf{p}ebfu_1eaf = \mathbf{p}eafu_1ebf = \mathbf{p}eafu_2ebf = \mathbf{p}ebfu_2eaf = \mathbf{r}_2eaf \in SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_2e)$. So $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1e) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_2e)$. Thus $\mathbf{r}_2 \rho_r \mathbf{r}_1$ implies $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1e) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_2e)$. By Lemma 7, $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_ie) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_i)$, whence $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1) = SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_2)$.

Corollary 10 *Let $P(\mathbf{q}, \mathbf{r})$ be a non-empty set of cycle states \mathbf{p} of a locally threshold testable DFA such that $\mathbf{p} \succeq \mathbf{q}$ and $\mathbf{p} \succeq \mathbf{r}$ for cycle state (\mathbf{q}, \mathbf{r}) .*

Then non-empty $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1)$ does not depend on \mathbf{r}_1 for any $\mathbf{p} \in P(\mathbf{q}, \mathbf{r})$.

4 An algorithm for local threshold testability

A linear depth-first search algorithm which finds all SCC (see [9]) will be used. The algorithm is based on Theorem 8 for a complete transition graph Γ (or Γ which is completed by sink state). The measures of complexity of the transition graph Γ are here $|\Gamma|$ (state complexity), the sum of the numbers of the states and the transitions k and the size of the alphabet g of the labels. Let us notice that $|\Gamma|(g+1) \geq k$.

Let us find all SCC of the graphs Γ and Γ^2 and all their cycle states. Furthermore we recognize the reachability on the graph Γ and form the table of reachability for all pairs of states. The step uses $O(|\Gamma|^2g)$ time and space.

The first condition of Theorem 8. For every cycle state (\mathbf{p}, \mathbf{q}) ($\mathbf{p} \neq \mathbf{q}$) from Γ^2 , let us check the condition $\mathbf{p} \sim \mathbf{q}$. A negative answer for any considered cycle state (\mathbf{p}, \mathbf{q}) implies the validity of the condition. In the opposite case, the automaton is not locally threshold testable. The time of the step is $O(|\Gamma|^2)$.

The second condition of Theorem 8. For every cycle state (\mathbf{p}, \mathbf{s}) , we form the set T of states $\mathbf{t} \in \Gamma$ such that $\mathbf{s} \succeq \mathbf{t}$ and for some state \mathbf{r} holds: (\mathbf{r}, \mathbf{t}) is a

cycle state and $\mathbf{p} \succeq \mathbf{r} \succeq \mathbf{s}$. If there exists a state \mathbf{q} such that $(\mathbf{p}, \mathbf{s}) \succeq (\mathbf{q}, \mathbf{t})$ for $\mathbf{t} \in T$ and $\mathbf{q} \not\succeq \mathbf{t}$, then the automaton is not threshold locally testable. It is a step of worst case asymptotic cost $O(|\Gamma|^4 g)$ with space complexity $O(|\Gamma|^3)$.

The condition 3 of Theorem 8. For every three states $\mathbf{p}, \mathbf{q}, \mathbf{s}$ of the automaton such that (\mathbf{p}, \mathbf{s}) is a cycle state, $\mathbf{p} \succeq \mathbf{s}$ and $\mathbf{p} \succeq \mathbf{q}$, let us find a state \mathbf{r} such that $\mathbf{p} \succeq \mathbf{r} \succeq \mathbf{s}$ and then let us find $SCC(\mathbf{p}, \mathbf{q}, \mathbf{s})$. In the case that this set is not well-defined (for $\mathbf{t}_1, \mathbf{t}_2$ from $SCC(\mathbf{p}, \mathbf{q}, \mathbf{s})$ $\mathbf{t}_1 \not\succeq \mathbf{t}_2$), the automaton is not threshold locally testable (Lemma 5). The time required for this step in the worst case is $O(|\Gamma|^4 g)$. The space complexity is $O(|\Gamma|^3)$.

Before checking condition 4, let us check the assertion of Lemma 9. For every cycle state (\mathbf{q}, \mathbf{r}) of the graph Γ^2 , let us form the set $P(\mathbf{q}, \mathbf{r})$ of cycle states \mathbf{p} such that $\mathbf{p} \succeq \mathbf{q}$ and $\mathbf{p} \succeq \mathbf{r}$. We continue for non-empty set $P(\mathbf{q}, \mathbf{r})$. For every state \mathbf{q}_1 , let us form the set $R(\mathbf{q}_1)$ of states \mathbf{r}_1 such that $(\mathbf{q}, \mathbf{r}) \succeq (\mathbf{q}_1, \mathbf{r}_1)$ and the state $(\mathbf{q}_1, \mathbf{r}_1)$ is a cycle state. Let us consider two states $\mathbf{r}_1, \mathbf{r}_2$ from the set $R(\mathbf{q}_1)$ such that the states $(\mathbf{p}, \mathbf{r}_1)$ and $(\mathbf{p}, \mathbf{r}_2)$ are cycle states for some \mathbf{p} from $P(\mathbf{q}, \mathbf{r})$. If $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1) \neq SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_2)$, then the automaton is not locally threshold testable.

The condition 4 of Theorem 8. For every cycle state (\mathbf{q}, \mathbf{r}) of Γ^2 , let us form the set $P(\mathbf{q}, \mathbf{r})$ of cycle states \mathbf{p} such that $\mathbf{p} \succeq \mathbf{q}$ and $\mathbf{p} \succeq \mathbf{r}$. We continue for non-empty set $P(\mathbf{q}, \mathbf{r})$. By Corollary 10, $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1)$ for given \mathbf{r} depends only on the states \mathbf{p}, \mathbf{q} , and $SCC(\mathbf{p}, \mathbf{r}, \mathbf{q}_1)$ for given \mathbf{q} depends only on \mathbf{p}, \mathbf{r} . If $SCC(\mathbf{p}, \mathbf{q}, \mathbf{r}_1)$ and $SCC(\mathbf{p}, \mathbf{r}, \mathbf{q}_1)$ exist and are not equal, then the automaton is not locally threshold testable according to condition 4. The time required for these last two steps in the worst case is $O(|\Gamma|^4 g)$ with $O(|\Gamma|^3)$ space.

A positive answer for all the cases implies the local threshold testability of the automaton. The time complexity of the algorithm is no more than $O(|\Gamma|^4 g)$. The space complexity is $\max(O(|\Gamma|^2 g), O(|\Gamma|^3))$. In more conventional formulation, we have $O(k^4)$ time and $O(k^3)$ space.

5 Conclusion. The package TESTAS

The considered algorithm is now implemented as a part of the C/C^{++} package TESTAS replacing the old version of the algorithm and reducing the time of execution. The program worked essentially faster in many cases we have studied because of the structure of the algorithm. A part of branches of the algorithm have only $O(|\Gamma|^2 g)$ or $O(|\Gamma|^3 g)$ time and space complexity.

The maximal size of the considered graphs on an ordinary PC was about several hundreds states with an alphabet of several dozen letters. The program in such case used memory on hard disc and works slower.

The package realizes, besides the considered algorithm for local threshold testability, a set of algorithms for checking local testability, left local testability, right local testability, piecewise testability and some other programs.

The package checks also the synchronizeability of the automaton and finds synchronizing words. The programs of the package TESTAS analyze:

1) an automaton of the language presented by oriented labelled graph. The automaton is given by the matrix:

states X labels

The non-empty (i,j) cell contains the state from the end of the transition with label from the j-th column and beginning in the i-th state.

2) An automaton of the language presented by its syntactic semigroup. The semigroup is presented by the matrix (Cayley graph)

elements X generators

where the i-th row of the matrix is a list of products of the i-th element on all generators. The set of generators is not necessarily minimal, therefore the multiplication table of the semigroup (Cayley table) is acceptable, too.

Some auxiliary programs of the package find direct products of the objects and build the syntactic semigroup of the automaton on the base of the transition graph.

The space complexity of the algorithms which consider the transition graph of an automaton is not less than $|\Gamma|g$ because of the structure of the input. The graph programs use usually a table of reachability defined on the states of the graph. The table of reachability is a square table and so we have $|\Gamma|^2$ space complexity.

The number of the states of Γ^n is $|\Gamma|^n$, the alphabet is the same as in Γ . So the sum of the numbers of the states and the transitions of the graph Γ^n is not greater than $(g + 1)|\Gamma|^n$. Some algorithms of the package use the powers Γ^2 , and Γ^3 . So the space complexity of the algorithms reaches in these cases $|\Gamma|^2g$ or $|\Gamma|^3g$.

An algorithm for the local testability problem for the transition graph ([13]) of $O(k^2)$ (or $O(|\Gamma|^2g)$) time and space is implemented in the package. An algorithm of $O(|\Gamma|^2g)$ time and of $O(|\Gamma|^2g)$ space is used for finding the bounds on the order of local testability for a given transition graph of the automaton [12]. An algorithm of worst case $O(|\Gamma|^3g)$ time complexity and of $O(|\Gamma|^2g)$ space complexity checked the 2-testability [12]. The 1-testability is verified using an algorithm [6] of order $O(|\Gamma|g^2)$. For checking the n -testability [12], we use an algorithm of worst case asymptotic cost $O(|\Gamma|^3g^{n-1})$ of time complexity with $O(|\Gamma|^2g)$ space. The time complexity of the last algorithm grows with n and in this way we obtain a non-polynomial algorithm for finding the order of local testability. However, $n \leq \log_g M$ where M is the maximal size of the integer in the computer memory.

The time complexity of the algorithm to verify piecewise testability of DFA is $O(|\Gamma|^2g)$. The space complexity of the algorithm is $O(k)$ [13].

The algorithms for right and left local testability for the transition graph are essentially distinct, moreover, the time complexity of the algorithms differs. The graph algorithm for the left local testability problem needs in the worst case $O(|\Gamma|^3g)$ time and $O(|\Gamma|^3g)$ space and the algorithm for the right local testability problem for transition graph of the deterministic finite automaton

needs $O(|\Gamma|^2g)$ time and space [14].

The main measure of complexity for semigroup S is the size of the semigroup $|S|$. We use also the number of generators (size of alphabet) g and the number of idempotents i .

Algorithms of the package dealing with the transition semigroup of an automaton use the multiplication table of the semigroup of $O(|S|^2)$ space. Other arrays used by the package present subsemigroups or subsets of the transition semigroup. So we usually have $O(|S|^2)$ space complexity.

We implement in the package TESTAS a polynomial-time algorithms of $O(|S|^2)$ time complexity for the local testability problem and for finding the order of local testability for a given semigroup [11].

The time complexity of the semigroup algorithm for both left and right local testability is $O(|S|i)$ [14]. The time complexity of the semigroup algorithm for local threshold testability is $O(|S|^3)$. Piecewise testability is verified in $O(|S|^2)$ time [13].

Acknowledgments

I am grateful to the anonymous referees for helpful and detailed comments that proved very useful in improving the presentation and style of the paper.

References

- [1] D. Beauquier, J.E. Pin, Factors of words, Lect. Notes in Comp. Sci., Springer, 372 (1989) 63-79.
- [2] J.A. Brzozowski, I. Simon, Characterizations of locally testable events, Discrete Math. 4 (1973) 243-271.
- [3] P. Caron, LANGAGE: A Maple package for automaton characterization of regular languages, Lect. Notes in Comp. Sci., Springer, 1436 (1998) 46-55.
- [4] P. Caron, Families of locally testable languages, Theoret. Comput. Sci., 242 (2000) 361-376.
- [5] S. Kim, R. McNaughton, R. McCloskey, A polynomial time algorithm for the local testability problem of deterministic finite automata, IEEE Trans. Comput., **10**, 40 (1991) 1087-1093.
- [6] S. Kim, R. McNaughton, Computing the order of a locally testable automaton, SIAM J. Comput. 23 (1994) 1193-1215.
- [7] R. McNaughton, S. Papert, Counter-free Automata M.I.T. Press. Mass. 1971.

- [8] J. Ruiz, S. Espana, P. Garcia, Locally threshold testable languages in strict sense: Application to the inference problem. *Lect. Notes in Comp. Sci.*, Springer, 1433 (1998) 150-161.
- [9] R.E. Tarjan, Depth first search and linear graph algorithms, *SIAM J. Comput.*, 1 (1972) 146-160.
- [10] D. Thérien, A. Weiss, Graph congruences and wreath product. *J. Pure Appl. Alg.*, 35 (1985) 205-215.
- [11] A.N. Trahtman, A polynomial time algorithm for local testability and its level. *Int. J. of Algebra and Comp.* **9**, 1 (1998) 31-39.
- [12] A.N. Trahtman, Algorithms finding the order of local testability of deterministic finite automaton and estimation of the order, *Theoret. Comput. Sci.*, 235 (2000) 183-204.
- [13] A.N. Trahtman, Piecewise and local threshold testability of DFA. *Lect. Notes in Comp. Sci.*, Springer, 2138 (2001) 347-358.
- [14] A.N. Trahtman, A polynomial time algorithm for left [right] local testability. *Lect. Notes in Comp. Sci.*, Springer, 2608 (2003) 203-212.
- [15] E. Vidal, F. Casacuberta, P. Garcia, Grammatical inference and automatic speech recognition. In *Speech Recognition and Coding* (Springer, 1995) 175-191.
- [16] Th. Wilke, Locally threshold testable languages of infinite words, *Lect. Notes in Comp. Sci.*, Springer, 665 (1993) 63-79.
- [17] Th. Wilke, An algebraic theory for regular languages of finite and infinite words, *Int. J. Alg. and Comput.*, **4**, 3 (1993) 447-489.