INSTYTUT INFORMATYKI
UNIWERSYTETU WROCŁAWSKIEGO

INSTITUTE OF COMPUTER SCIENCE
UNIVERSITY OF WROCŁAW

ul. Joliot-Curie 15
50–383 Wrocław
Poland

Report 01/08

Tomasz Jurdziński

Leftist Grammars are Nonprimitive Recursive

March 2008

Leftist Grammars are Nonprimitive Recursive*

Tomasz Jurdziński

Institute of Computer Science, University of Wrocław, Poland
tju@ii.uni.wroc.pl

Abstract. Leftist grammars were introduced by Motwani et. al. [6], as a tool to show decidability of the accessibility problem in certain general protection systems. It is shown that the membership problem for languages defined by leftist grammars is nonpritive recursive. Therefore, by the reduction of Motwani et. al., the accessibility problem in the appropriate protection systems is nonprimitive recursive as well.

1 Introduction

A protection system is a set of policies that prescribes the ways in which *objects* interact with each other in a computer system. By objects we mean users, processes or other entities. Interactions can include access rights, information sharing privileges and so on. The accessibility problem for a protection system is formulated in the form "Can object p gain (illegal) access to object q by a series of legal moves (as prescribed by the policy)?". A formal treatment of accessibility was first presented by Harrison, et. al. [3] who showed that the accessibility problem is undecidable for a general access-matrix model. This result prompted a broad research on tradeoffs between expressibility and verifiability in protection systems. We consider the model proposed in [2, 7] in the context of Java virtual worlds, called here the *Saraswat's model.* The accessibility problem is decidable for this model. This was obtained by a polynomial time reduction of the accessibility problem to the intersection problem for leftist grammars [6]. Further refinement and applications of this model are presented in [8].

Leftist grammars can be characterized in terms of rules of the form $a \to ba$ and $cd \to d$, where a, b, c, d belong to the finite alphabet Σ. A symbol $x \in \Sigma$ is called the final symbol and a word $w \in \Sigma^*$ belongs to the language defined by the grammar G iff there exists a derivation which starts at wx and ends at x. The membership problem for (the language defined by) a leftist grammar G is, given a word w, to decide whether w belongs to the language of G. Let us note here that each leftist grammar is actually a semi-Thue system with leftist rewriting rules. So, the membership problem is actually a special case of the word problem for semi-Thue systems with leftist rewriting rules.

It is known that the membership problem problem for leftist grammars is decidable [6]. The result from [6] implies also that any lower bound for complexity of the membership problem for leftist grammars induces an analog lower bound for the accessibility problem in the Saraswat's model (up to a polynomial factor). Simplicity of leftist grammars led to the conjecture that the actual complexity of the membership problem is small. As shown in [5, 1], quite natural restrictions imply context-freenes or even regularity of languages defined by leftist grammars.

* Partially supported by MNiSW grant number N206 024 31/3826, 2006-2008.

However, it has been shown that the membership problem for general leftist grammars is PSPACE-hard [4]. In this paper, we strengthen this result in the following way. We prove that, if a grammar and an input word form the input for the problem, the membership problem is nonprimitive recursive. In the case that a grammar is fixed, we show that, for each primitive recursive function f, there exists a grammar $\mathcal{G}_f$ such that the membership problem for $\mathcal{G}_f$ is not in space $O(f(n))$, where n is the length of an input word.

In Section 2 we provide basic definitions and notations. In Sections 3 and 4 we show how leftist grammars can "compute" certain families of functions related to the Ackermann's function. Finally, in Section 5, we present the main result of the paper, using the tools from Sections 3 and 4.

2 Definitions

Throughout the paper λ denotes the empty word, $\mathbb{N}$ denotes the set of non-negative integers. For a word x, let $|x|$, $x[i]$ and $x[i,j]$ denote the length of x, the ith symbol of x and the factor $x[i]\cdots x[j]$ respectively, where $0 < i \leq j \leq |x|$. Moreover, let $[i,j] = \{l \in \mathbb{N} \mid i \leq l \leq j\}$, let $i\%j$ be equal the remainder of the division i/j and let $\bar{i} = 1 - i$ for $i \in [0,1]$. Furthermore, we identify regular expressions with languages defined by them.

We say that the sets $A_0, \ldots, A_p$ form the partition $\mathcal{A} = (A_0, \ldots, A_p)$ of the set $A = \bigcup_{i=0}^{p} A_i$ iff $A_i \neq \emptyset$ and $A_i \cap A_j = \emptyset$ for each $i, j \in [0,p]$, $i \neq j$. A word $w \in A^*$ is an ***alternating word*** with respect to the partition $\mathcal{A}$ if:

- $w \in A_i^*$ for $i \in [0,p]$, or
- $w = w_1ab$, where w_1a is an alternating word with respect to $\mathcal{A}$, $a \in A_i$ and $b \in A_i \cup A_{(i+1)\%(p+1)}$ for $i \in [0,p]$.

Note that, if $p = 1$ then each word over A is an alternating word wrt $\mathcal{A}$. Let w be an alternating word wrt to $\mathcal{A}$. Then, $\|w\|_{\mathcal{A}}$ is defined as follows:

- $\|\lambda\|_{\mathcal{A}} = 0$;
- $\|w\|_{\mathcal{A}} = 1$ if $w \in A_i^+$ for $i \in [0,p]$;
- $\|w\|_{\mathcal{A}} = \|w_2\|_{\mathcal{A}} + 1$ if $w = w_1w_2$ for $w_1 \in A_i^+$, $w_2 \neq \lambda$ and $w_2[1] \notin A_i$ for $i \in [0,p]$.

If w is not an alternating word wrt $\mathcal{A}$, $\|w\|_{\mathcal{A}}$ is not defined.

A *leftist grammar* $\mathcal{G} = (\Sigma, P, x)$ consists of a finite alphabet Σ, a final symbol $x \in \Sigma$, and a set of production rules P of the following two types,

$$ab \to b \text{ (Delete Rule)}, \; c \to dc \text{ (Insert Rule)}$$

where $a, b, c, d \in \Sigma$. We will denote the above productions as $b \to^{\mathsf{del}} a$ and $c \to^{\mathsf{ins}} d$. We say that $u \Rightarrow_{\mathcal{G}} v$ (or shortly $u \Rightarrow v$) is a derivation step for $u, v \in \Sigma^*$, if $u = u_1yu_2$ and $v = u_1zu_2$ such that $y \to z$ is a production rule in P. The sequence of derivation steps $u_1 \Rightarrow u_2 \Rightarrow \ldots \Rightarrow u_p$ is called a derivation. The word u_i for $i \in [1,p]$ is called a sentential form in this derivation. Finally, the language of $\mathcal{G}$ is defined to be $L(\mathcal{G}) = \{w \in \Sigma^* \mid wx \Rightarrow^* x\}$. The ***membership problem*** for a fixed leftist grammar $\mathcal{G} = (\Sigma, P, x)$ is, given a word $w \in \Sigma^*$ as an input, to decide whether $w \in L(\mathcal{G})$. The ***variable membership problem*** is, given a word $w \in \Sigma^*$ and a leftist grammar $\mathcal{G} = (\Sigma, P, x)$ as an input, to decide whether $w \in L(\mathcal{G})$.

We think of symbols of sentential forms as objects which can insert/delete other symbols and can be inserted/deleted. In order to simplify notations, we identify the particular occurrence of a symbol a with its value a.

We say that the symbol b in the delete rule $ab \to b$ is ***active***. Similarly, the symbol c is ***active*** in the insert rule $c \to dc$. Let $u \Rightarrow v$, where $u = u_1 y u_2$ and $v = v_1 z v_2$ such that $y \to z$ is a production rule in P. Then, we say that the rightmost symbol of the prefix $u_1 y$ of $u_1 y u_2$ is *active* in the derivation step $u \Rightarrow v$. Though it might be in general the case that the choice of the active symbol for a step $u \Rightarrow v$ is not unique, one can avoid this ambiguity [5]. So, we assume that the active symbol can be determined uniquely for each derivation step. Let $U \equiv (u_1 \Rightarrow u_2 \Rightarrow \cdots \Rightarrow u_p)$ be a derivation. We say that $u_1[i]$ is ***idle*** in u_1 with respect to U if it is *not* active in each derivation step of U.

Now, let b, d be symbols which appear in some sentential forms of a derivation U. Then, d is a ***descendant*** of b with respect to U if (b, d) belongs to the reflexive and transitive closure of the relation

$$\{(e, f) \mid v\underline{e}w \Rightarrow v\underline{f}ew \text{ is the derivation step in } U \text{ for } v, w \in \Sigma^*,\ e, f \in \Sigma\}.$$

Let $\mathcal{G} = (\Sigma, P, x)$ be a leftist grammar, where $\Sigma = \{a_i\}_{i=1}^p$. The *Insert Graph* of $\mathcal{G}$ is $G(V, E)$, where $V = \{v_i\}_{i=1}^p$ and $E = \{(v_i, v_j) \mid (a_i \to a_j a_i) \in P\}$. Similarly, the *Delete Graph* of $\mathcal{G}$ is $G(V, E)$, where $V = \{v_i\}_{i=1}^p$, and $E = \{(v_i, v_j) \mid (a_j a_i \to a_i) \in P\}$.

Let $\mathsf{Del}(a) = \{b \mid b \to^{\mathsf{del}} a\}$ for $a \in \Sigma$ and $\mathcal{G} = (\Sigma, P, x)$, and let $\mathsf{Del}(A)$ for $A \subseteq \Sigma$ be $\bigcup_{a \in A} \mathsf{Del}(a)$. That is, $\mathsf{Del}(a)$ is the set of in-neighbours of a in the insert graph of $\mathcal{G}$. A set $A \subseteq \Sigma$ is ***homogeneous*** if a is not active in any production rule of $\mathcal{G}$ for each $a \in A$, and $\mathsf{Del}(a) = \mathsf{Del}(b)$ for each $a, b \in A$.

The derivation $u_1 \Rightarrow u_2 \Rightarrow \ldots \Rightarrow u_p$ is a ***leftmost derivation*** if all symbols located to the left of the symbol active in $u_i \Rightarrow u_{i+1}$ are idle with respect to $u_i \Rightarrow^* u_p$ for every $i \in [1, p-1]$. For each $u, v \in \Sigma^*$ such that $u \Rightarrow^*_{\mathcal{G}} v$, there exists a leftmost derivation which starts at u and ends at v [5].

We say that a word u ***eliminates*** a word $w \in \Sigma^+$ in the derivation $z_1 w u z_2 \Rightarrow^* z'$, if all elements of w are idle with respect to this derivation, and all elements of w are deleted by the elements of u and their descendants. Moreover, we say that $u \in \Sigma^+$ is ***able to eliminate*** $v \in \Sigma^+$ if there exists a derivation $vu \Rightarrow^* v'$, where u eliminates v.

Let $U \equiv (w \Rightarrow^* w_1' v a w_2')$ and let v consist of all descendants of a in $w_1' v a w_2'$. Then, the ***trace*** of a in U, $\mathsf{tr}_U(a)$ or shortly $\mathsf{tr}(a)$, is equal to va.

Definition 1 (Greedy derivation). *A derivation $U \equiv (u_1 \Rightarrow u_2 \Rightarrow \cdots \Rightarrow u_p)$ is greedy if:*

(a) U is a leftmost derivation;
(b) a symbol a is allowed to become idle wrt to $u_j \Rightarrow^* u_p$ only if it is not able to apply a delete rule in u_j (that is, if the grammar contains the rule $a \to^{\mathsf{del}} b$ then a is not idle wrt $u_j \Rightarrow^* u_p$, if $u_j = zbav$ for $z, v \in \Sigma^*$);
(c) there is no derivation step $u\underline{a}v \Rightarrow u\underline{ba}v$ in U such that b does not eliminate any element of u during U.

It is known that there exists a greedy derivation $wx \Rightarrow^*_{\mathcal{G}} x$ for each leftist grammar $\mathcal{G}$ and $w \in L(\mathcal{G})$ [4]. So, we will consider only greedy derivations.

For a grammar $\mathcal{G}$, let size of $\mathcal{G}$, $|\mathcal{G}|$, be equal to the size of its alphabet.

2.1 Ackermann's function

For a function $f : \mathbb{N} \to \mathbb{N}$, let $f^{(0)}(n) = n$ and $f^{(k)}(n) = f(f^{(k-1)}(n))$ for each $n \in \mathbb{N}$ and $k \geq 1$. Let the functions $\phi_n : \mathbb{N} \to \mathbb{N}$ be defined as follows:

$$\begin{array}{ll} \phi_2(k) = 2 \cdot k & \text{for } k > 0, \\ \phi_n(k) = \phi_{n-1}^{(k)}(1) & \text{for } n > 2 \text{ and } k > 0. \end{array}$$

Moreover, let $\phi_n(0) = 1$ for each $n \geq 2$. One of possible definitions of Ackermann's function is $Ack(n) = \phi_n(3)$, see [9]. We define inverse functions $(\phi_n^{-1})_{n \geq 2}$ as

$$\phi_n^{-1}(m) := \min\{k \mid \phi_n(k) \geq m\}.$$

Notice that $\phi_n^{-1}(m) = k$ if and only if $(\phi_{n-1}^{-1})^{(k)}(m) = 1$ for $n > 2$. Moreover, observe that ϕ_ns are monotone, i.e., $\phi_n(k+1) \geq \phi_n(k)$.
It is well-known that $Ack(n)$ dominates any primitive recursive function of n.

3 Expansion

Our first goal is to build a grammar which somehow "computes" a function ϕ_p for $p \in \mathbb{N}$. The idea is as follows. Let I (the "input alphabet") and O (the "output alphabet") be disjoint finite sets and let $g \notin I \cup O$. Given a word $w \in I^*$, g is able to eliminate w only in a derivation $wg \Rightarrow^* w'g$, such that $w' \in O^*$ and $|w'| = \phi_p(|w|)$. Due to limitations of leftist grammars, the following definition is more complicated than the above scenario. In particular, the result of a "computation" is approximate in such a way that $|w'|$ might be larger than $\phi_p(|w|)$, but not smaller. Since we will eventually use these "computational" grammars as parts of larger grammars accepting words in a standard way, we can assume that only greedy derivations are considered.

Definition 2. *Let I, B, O, F and $\{g\}$ be disjoint, finite and nonempty sets of symbols and let $\mathcal{I} = (I_0, I_1)$, $\mathcal{O} = (O_0, O_1)$ be partitions of I and O, respectively. Moreover, let the following conditions be satisfied in a grammar $\mathcal{G}$:*

1. *For each $v \in B^+$ and $w \in I^*$, there exists a derivation $vwg \Rightarrow^* v'w'g$ in which g eliminates vw, such that $v' \in F^+$, $w' \in O^+$ and $\|w'\|_{\mathcal{O}} = \phi_p(\|w\|_{\mathcal{I}})$.*
2. *Let $v \in B^*$, $w \in I^*$ such that $|vw| > 0$ and let $U \equiv (vwg \Rightarrow^* zg)$ be a derivation in which g eliminates vw; then, either*
   - *$z = v'w'g$ such that $v' \in F^+$, $w' \in O^+$ and $\|w'\|_{\mathcal{O}} \geq \phi_p(\|w\|_{\mathcal{I}})$;* ***or***
   - *$z \notin (F \cup O)^*$.*
3. *$\mathsf{Del}(B) \subseteq F$, the elements of $\mathsf{Del}(B)$ are able to eliminate only the elements of B and g is not able to eliminate elements of $\mathsf{Del}(B)$.*
4. *The sets I_0, I_1 and B are homogeneous in $\mathcal{G}$.*

Then, we say that $\mathcal{G}$ satisfies p-expansion wrt $\mathcal{I}$, B, $\mathcal{O}$, F and $\{g\}$, we denote this fact by $\mathsf{ex}_p(\mathcal{I}, B, \mathcal{O}, F, g)$.

The statement 1. of the above definition corresponds to the property that it is possible to "compute" the function ϕ_p precisely, while the second statement denotes the property that each "computation" is such that the result is not smaller than the actual value of the function ϕ_p. In both cases, the "computation" stands for a subderivation, where g eliminates a word over I and rewrites it with a word over O. The values of arguments and results of computations correspond to the measure of alternation of the deleted word and the inserted word wrt the partitions $\mathcal{I}$ and $\mathcal{O}$, respectively. The subalphabets B (the "border alphabet") and F (the "frontier alphabet") are introduced for technical reasons. Their roles are essential in the inductive construction in Lemma 2. The third statement of Def. 2 guarantees that g is not able to eliminate the elements of I which appear to the left of a symbol from B. This property helps in the proof that the inductive construction from Lemma 2 works.

For $\mathcal{G}$ satisfying $\mathsf{ex}_p(\mathcal{I}, B, \mathcal{O}, F, g)$, we use the notation $\mathcal{I}(\mathcal{G}) = \mathcal{I}$, $B(\mathcal{G}) = B$, $\mathcal{O}(\mathcal{G}) = \mathcal{O}$, and $F(\mathcal{G}) = F$. According to Def. 2 (statement 4), the construction of a grammar satisfying p-expansion is independent of $\mathcal{I}$ and B. Therefore, if $\mathcal{I}, B$ do not matter, we denote p-expansion by $\mathsf{EX}_p(\mathcal{O}, F, g)$.
Given I, B and a partition $\mathcal{I} = (I_0, I_1)$ of I, we build a grammar $\mathcal{G}_{\phi_2}$ with the following production rules:

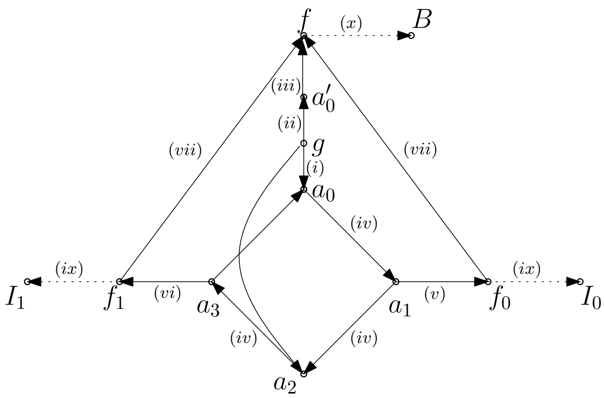

Fig. 1. The insert graph and a part of the delete graph of $\mathcal{G}_{\phi_2}$. Dotted edges belong to the delete graph. Vertices labeled with sets denote here sets of vertices corresponding to sets of homogenous symbols.

(i)	$g \to^{\mathsf{ins}} a_i$	for $i \in \{0, 2\}$	(vi)	$a_3 \to^{\mathsf{ins}} f_1$	
(ii)	$g \to^{\mathsf{ins}} a'_0$		(vii)	$f_j \to^{\mathsf{ins}} f$	for $j \in [0, 1]$
(iii)	$a'_0 \to^{\mathsf{ins}} f$		$(viii)$	$f_j \to^{\mathsf{del}} f_{1-j}$	for $j \in [0, 1]$
(iv)	$a_i \to^{\mathsf{ins}} a_{(i+1)\%4}$	for $i \in [0, 3]$	(ix)	$f_j \to^{\mathsf{del}} \imath$	for $j \in [0, 1], \imath \in I_j$
(v)	$a_1 \to^{\mathsf{ins}} f_0$		(x)	$f \to^{\mathsf{del}} b$	for $b \in B$

where a'_0, a_j, f_i, $g \notin I \cup B$ for each $i \in [0, 1]$ and $j \in [0, 3]$ (see Figure 1).

Lemma 1. *The grammar $\mathcal{G}_{\phi_2}$ satisfies $\mathsf{ex}_2(\mathcal{I}, B, \mathcal{O}, F, g)$, where $\mathcal{O} = (O_0, O_1)$, $F = \{f_0, f_1, f\}$, $O_0 = \{a_0, a'_0, a_2\}$, and $O_1 = \{a_1, a_3\}$.*

Proof. The idea of the proof is based on the following observations:

- each greedy derivation in which g and its descendants are active corresponds to a path in the insert graph of $\mathcal{G}_{\phi_2}$ which starts with g, and goes either through a'_0 or through one of the cycles $a_0 a_1 a_2 a_3$, $a_2 a_3 a_0 a_1$ (possibly many times);

– the cycle $a_0a_1a_2a_3$ ($a_2a_3a_0a_1$) describes the fact that the factor $a_3a_2a_1a_0$ ($a_1a_0a_3a_2$) is added to the trace of g and each subword from $I_1^*I_0^*$ can be deleted.

In the above description, we ignored the roles of symbols f_0, f_1, and f. However, the rules defining insertions/deletions of these symbols and their activity play the key role in the construction. Below, we prove the lemma by checking all statements of Definition 2.

Observe that $\mathsf{Del}(B) = \{f\}$. This fact, combined with the constraints of (i)-(x), implies that the statements 3 and 4 of Def. 2 are satisfied (note that f cannot be deleted, the elements of I_j may be deleted only using (ix)).

Now, let $v \in B^+$ and $w \in I^*$. We show that there exists a derivation $vwg \Rightarrow^* v'w'g$ which satisfies the statement 1 of Def. 2. If $w = \lambda$, then

$$vwg = vg \Rightarrow_{(ii)} va_0'g \Rightarrow_{(iii)} vfa_0'g \Rightarrow_{(x)}^* fa_0'g = v'w'g,$$

where $v' = f \in F$, $w' = a_0' \in O$. Thus, $\|w'\|_{\mathcal{O}} = 1 = \phi_2(0) = \phi_2(\|w\|_{\mathcal{I}})$.

Now, assume that $|w| > 0$. Let $w = w_{n-1} \dots w_0$ such that $w_i \in I_i^+$ for $i \in [0, n-1]$, n is even and $n > 0$. That is, $\|w\|_{\mathcal{O}} = n$. One can build a derivation described by the following algorithm:

1. $vwg \Rightarrow_{(i,iv,v)}^* vw_{n-1} \dots w_0f_0a_1a_0g \Rightarrow_{(ix)}^* vw_{n-1} \dots w_1f_0a_1a_0g$.
2. For $j = 1, 2, \dots, n-1$ do: **[delete w_j]**
   (a) if j is even: $w_jf_1(a_3a_2a_1a_0)^{j/2}g \Rightarrow_{(iv,v,viii,ix)}^* f_0a_1a_0(a_3a_2a_1a_0)^{j/2}g$.
   (b) if j is odd: $w_jf_0a_1a_0(a_3a_2a_1a_0)^{\lfloor j/2 \rfloor}g \Rightarrow_{(iv,vi,viii,ix)}^* f_1(a_3a_2a_1a_0)^{(j+1)/2}g$.
3. $vf_1(a_3a_2a_1a_0)^{n/2}g \Rightarrow_{(vii,x)}^* ff_1(a_3a_2a_1a_0)^{n/2}g$.

So, we obtain the derivation

$$vwg \Rightarrow^* ff_1(a_3a_2a_1a_0)^{n/2} = v'w'g,$$

where $v' = ff_1 \in F^+$, and $\|w'\|_{\mathcal{O}} = 2n = \phi_2(\|w\|_{\mathcal{I}})$. One can design an analog derivation for the case that n is odd. If $w_j \in I_{(j+1)\%2}^+$ for $w = w_{n-1} \cdots w_0$, we use the production $g \rightarrow^{\mathsf{ins}} a_2$ instead of $g \rightarrow^{\mathsf{ins}} a_0$ in the item 1.

Finally, we show that the statement 2 of Definition 2 is satisfied. Note that

(a) a_i is not able to eliminate a_j for each $i \neq j$;
(b) g is not able to eliminate any element of the set $\{a_0, \dots, a_3, a_0'\}$.

These restrictions and the requirements for greedy derivations imply that each derivation $g \Rightarrow^* \mathsf{tr}(g)$ is such that the elements of O are inserted from right to left by the production (iv) ((i) or (ii) can be applied only once), they are *not* deleted, and no symbol from F appears between or to the right of the elements of O. Moreover, by (a) and greediness, only the leftmost element of O can be active in each derivation step. In other words, each derivation $g \Rightarrow^* \mathsf{tr}(g)$ corresponds to a path in the insert graph of $\mathcal{G}_{\phi_2}$ which starts in g and goes through the vertices labeled by the elements of O, i.e., through the cycle $a_0a_1a_2a_3$ or the cycle $a_2a_3a_0a_1$ or through the vertex a_0'. Moreover, each occurrence of a_1 or a_3 gives a possibility to delete *at most* one of w_j's. These restrictions imply that $\mathsf{tr}(g) = v'w'g$ for $v' \in F^+$ and $w' \in O^+$. According to the constraints of $(i) - (x)$, w' belongs to the language of suffixes of the set of words $(a_3a_2a_1a_0)^*+$

$(a_1a_0a_3a_2)^*$. Since the elements of w' are inserted from right to left, the elements of O cannot be deleted in $\mathcal{G}_{\phi_2}$, and only a_1, a_3 and a_0' can insert elements of F (note that $\mathsf{Del}(I) \subseteq F$), each word vw for $v \in B^+$, $w \in O^+$ eliminated by g in a derivation $vwg \Rightarrow v'w'g$ satisfies $\|w'\|_{\mathcal{O}} \geq 2\|w\|_{\mathcal{I}} = \phi_2(\|w\|_{\mathcal{I}})$.

□

Lemma 2. *Assume that a grammar $\mathcal{G}_{\phi_i}$ satisfies $\mathsf{EX}_i(\widehat{\mathcal{O}}, \widehat{F}, \widehat{g})$. Then, one can build a grammar $\mathcal{G}_{\phi_{i+1}}$ of size at most $5|\mathcal{G}_{\phi_i}|$ which satisfies $\mathsf{EX}_{i+1}(\mathcal{O}, F, g)$ for some sets of symbols O, F, $\{g\}$ and the partition $\mathcal{O} = (O_0, O_1)$ of O.*

Proof. Let $\mathcal{G}_0, \ldots, \mathcal{G}_3$ be copies of the grammar $\mathcal{G}_{\phi_i}$ such that:

- $\mathcal{G}_j$ satisfies $\mathsf{EX}_i(\widehat{\mathcal{O}}_j,\ \widehat{F}_j, \widehat{g}_j)$ for $j \in [0, 3]$,
- $\Sigma_j \cap \Sigma_l = \emptyset$ for $j \neq l$, where Σ_k is the set of symbols accessible from $\widehat{g}_k$ in the insert graph of $\mathcal{G}_k$.

Let $\widehat{\mathcal{O}}_j = (\widehat{O}_{j,0}, \widehat{O}_{j,1})$ be a partition of $\widehat{O}_j$ for $j \in [0, 3]$. The partition $\mathcal{I}(\mathcal{G}_j)$ and the set $B(\mathcal{G}_j)$ for $j \in [0, 3]$ are as follows:

$$\mathcal{I}(\mathcal{G}_j) = \widehat{\mathcal{I}}_j := \widehat{\mathcal{O}}_{(j+1)\%2}, \text{ and}$$
$$B(\mathcal{G}_j) = \widehat{B}_j := \widehat{F}_{(j+1)\%2}$$

That is,

$$\mathcal{G}_j \text{ satisfies } \mathsf{ex}_i(\widehat{\mathcal{O}}_{j\%2}, \widehat{F}_{j\%2}, \widehat{\mathcal{O}}_j, \widehat{F}_j, \widehat{g}_j) \qquad \mathbf{(e(\mathit{j}))}$$

for $j \in [0, 3]$. So, the "output" subalphabet of the subgrammar $\mathcal{G}_j$ for $j \in [0, 1]$ is the "input" subalphabet of $\mathcal{G}_{(j+1)\%2}$ and $\mathcal{G}_{1+(j+1)\%2}$ (see Figure 2). Given disjoint nonempty sets I, B and a partition $\mathcal{I} = (I_0, I_1)$ of I, the grammar $\mathcal{G}_{\phi_{i+1}}$ is obtained by combining $\mathcal{G}_0, \ldots, \mathcal{G}_3$, adding new symbols g, f_0, f_1, f_2, and f_3 to the alphabet and the following production rules (see Figure 2):

$$\begin{array}{rlll}
(i) & e \rightarrow^{\mathsf{ins}} f_j & \text{for each } e \in \mathsf{Del}(\widehat{B}_j) \cap \widehat{F}_j, j \in [0,3] \\
(ii) & f_j \rightarrow^{\mathsf{del}} \imath & \text{for each } \imath \in I_j, j \in [0,1] \\
(iii) & f_j \rightarrow^{\mathsf{del}} \imath & \text{for each } \imath \in B, j \in [2,3] \\
(iv) & f_j \rightarrow^{\mathsf{del}} f_l & \text{for each } j \in [0,3], l = (j+1)\%2 \\
(v) & g \rightarrow^{\mathsf{ins}} \widehat{g}_j & \text{for } j \in [0,3] \\
(vi) & g \rightarrow^{\mathsf{del}} \widehat{g}_j & \text{for } j \in [0,3]
\end{array}$$

Our goal is to show that $\mathcal{G}_{\phi_{i+1}}$ defined in this way satisfies

$$\mathsf{ex}_{i+1}(\mathcal{I},\ B,\ \mathcal{O},\ \widehat{F}_2 \cup \widehat{F}_3 \cup \{f_2, f_3\},\ g),$$

where $\mathcal{O} = (O_0, O_1)$ for $O_k = \widehat{O}_{2,k} \cup \widehat{O}_{3,k}$ is a partition of $O := \widehat{O}_2 \cup \widehat{O}_3$. That is,

$$\mathcal{O}(\mathcal{G}_{\phi_{i+1}}) := (O_0, O_1), \text{ and } F(\mathcal{G}_{\phi_{i+1}}) := \widehat{F}_2 \cup \widehat{F}_3 \cup \{f_2, f_3\}.$$

The idea of the above construction is as follows:

- since $\mathcal{G}_0$ and $\mathcal{G}_1$ can delete elements of I_0 an I_1, respectively, it is necessary to apply these grammars alternately;

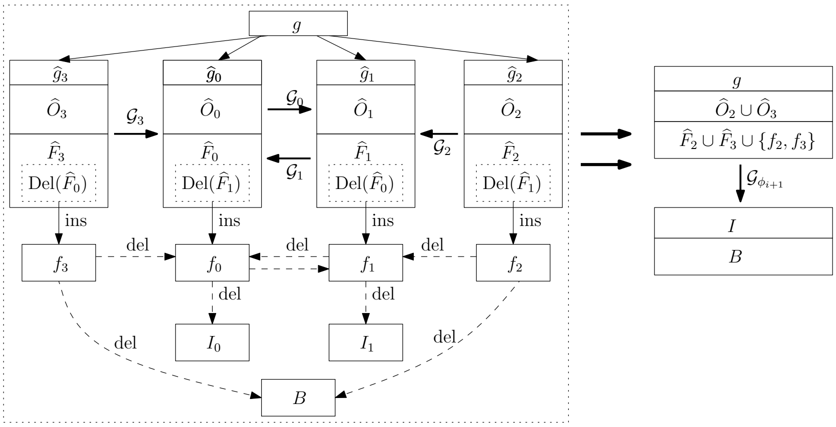

Fig. 2. An illustration for the construction of $\mathcal{G}_{\phi_{i+1}}$. An edge labeled by a grammar $\mathcal{G}$, going from a box $[\mathcal{O}, F]$ to a box $[\mathcal{I}, B]$ denotes that $\mathsf{ex}_i(g, \mathcal{I}, B, \mathcal{O}, F)$.

– each application of $\mathcal{G}_j$ deleting an element of I_j requires that the sequence over $\widehat{O}_{1-j}$ following the element of I_j is deleted as well; however, since $\mathcal{I}(\mathcal{G}_j) = \mathcal{O}(\mathcal{G}_{1-j})$, this subderivation replaces a sequence $w \in \widehat{O}^*_{1-j}$ with $w' \in \widehat{O}^*_j$ such that $\|w'\|_{\widehat{\mathcal{O}}_j} \geq \|w\|_{\widehat{\mathcal{O}}_{1-j}}$.

Below, we prove the lemma formally by checking whether $\mathcal{G}_{\phi_{i+1}}$ satisfies the statements of Def. 2.

First, observe that the statements 3 and 4 of Def. 2 hold for $\mathcal{G}_{\phi_{i+1}}$ by the fact that $\mathsf{Del}(B) = \{f_2, f_3\}$ and by the constraints of (ii)-(iii).

Next, we show that the stat. 1 of Def. 2 holds for $\mathcal{G}_{\phi_{i+1}}$. Let $v \in B^+$, $w \in I^*$ and let let $w = w_1 \cdots w_n$, where $w_l \in I^+_{l\%2}$ for each $l \in [1, n]$, i.e. $\|w\|_{\mathcal{I}} = n$.

For technical reasons, we extend the grammar $\mathcal{G}_j$ for $j \in [0, 3]$ by adding to it the productions (i)-(vi) in which $\widehat{g}_j$ is inserted/deleted or the elements of $\{f_j\} \cup \widehat{F}_j$ are active. Below, we describe a derivation $U \equiv (vwg \Rightarrow^* v'w'g)$ which consists of the stages $U_n, U_{n-1}, \ldots, U_1, U_0$ such that the stage U_k for $k > 0$ applies (productions of) the grammar $\mathcal{G}_{k\%2}$ and w_k is deleted in U_k. W.l.o.g.,

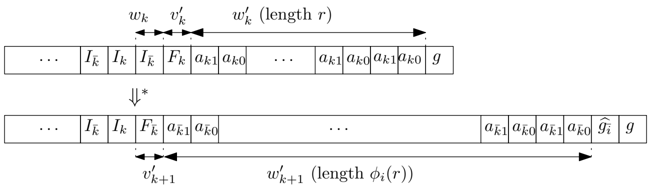

Fig. 3. One stage of the derivation $vwg \Rightarrow^* v'w'g$ (here, $a_{j,l} \in O_{j\%2,l}$).

assume that we are going to eliminate the supersequence vwv'_n of vw in U, where $v'_n \in f_{\overline{n\%2}}\widehat{F}_{\overline{n\%2}}$. The following algorithm describes the derivation U (see Figure 3):

1. $w'_n := \lambda$, $v'_n := f_{\overline{n\%2}} a$ for some $a \in \widehat{F}_{\overline{n\%2}}$;
2. For $k = n, n-1, \ldots, 1$ do

$$\textbf{stage } U_k\textbf{:} \quad w_1 \cdots \underline{w_k v'_k w'_k} g \Rightarrow^* w_1 \cdots w_{k-1} \underline{v'_{k-1} w'_{k-1}} g$$

   (a) Given $v'_k \in f_{\overline{k\%2}} \widehat{F}^+_{\overline{k\%2}}$ and $w'_k \in \widehat{O}^*_{\overline{k\%2}}$, we apply the subgrammar $\mathcal{G}_{k\%2}$ in order to eliminate $v'_k w'_k$ (without the leftmost $f_{\overline{k\%2}}$):

$$w_k v'_k w'_k g \Rightarrow_{(v)} w_k v'_k w'_k \widehat{g}_{k\%2} g \Rightarrow^*_{\text{ind.ass.}} w_k f_{\overline{k\%2}} v''_k w''_k \widehat{g}_{k\%2} g,$$

   where $v''_k \in \widehat{F}^+_{k\%2}$, $w''_k \in \widehat{O}^+_{k\%2}$ and $\|w''_k\|_{\widehat{\mathcal{O}}_{k\%2}} = \phi_i(\|w'_k\|_{\widehat{\mathcal{O}}_{\overline{k\%2}}}) = \phi_i^{(n-k+1)}(0)$, by the assumption $e(k\%2)$ on page 7. (Note that the leftmost symbol of v''_k belongs to $\widehat{F}_{k\%2} \cap \mathsf{Del}(\widehat{F}_{(k+1)\%2})$, by the stat. 3 of Def. 2.)

   (b) Eliminate $w_k f_{\overline{k\%2}}$:

$$\begin{aligned} w_k f_{\overline{k\%2}} v''_k w''_k \widehat{g}_{k\%2} g &\Rightarrow^*_{(i,iv)} w_k f_{k\%2} v''_k w''_k \widehat{g}_{k\%2} g \Rightarrow^*_{(ii)} f_{k\%2} v''_k w''_k \widehat{g}_{k\%2} g \\ &\Rightarrow_{(vi)} \quad f_{k\%2} v''_k w''_k g \quad = \quad v'_{k-1} w'_{k-1} g, \end{aligned}$$

   where $v'_{k-1} = f_{k\%2} v''_k \in f_{k\%2} \widehat{F}^+_{k\%2}$ and $w'_{k-1} = w''_k \in \widehat{O}^+_{k\%2}$.
3. Apply $\mathcal{G}_2$ in order to eliminate v:

$$v v'_0 w'_0 g \Rightarrow^*_{\text{ind.ass.}} v f_1 v'' w'' g \Rightarrow^*_{(i,iv)} v f_2 v'' w'' g \Rightarrow^*_{(iii)} f_2 v'' w'' g = v' w' g$$

   where $f_2 v'' = v' \in f_2 \widehat{F}^+_2 \subset F^+$, $w' = w'' \in \widehat{O}^+_2$ and

$$\|w'\|_{\widehat{\mathcal{O}}_2} = \phi_i(\|w'_0\|)_{\widehat{\mathcal{O}}_1} = \phi_i(\phi_i^{(n)}(0)) = \phi_i^{(n)}(1).$$

Finally, we have obtained a derivation $vwg \Rightarrow^* v v'_0 w'_0 g \Rightarrow^* v' w' g$, where $v' \in F^+$, $w' \in O^+$ and $\|w'\|_{\mathcal{O}} = \phi_{i+1}(\|w\|_{\mathcal{I}})$. This shows that the grammar $\mathcal{G}_{\phi_{i+1}}$ satisfies the statement 1 of Definition 2. (It is easy to build an analog derivation for the case that $w_j \in I^+_{\overline{j\%2}}$ for each j; then, we use $\mathcal{G}_3$ instead of $\mathcal{G}_2$ in the item 3.)

The final step is to show that the statement 2 of Def. 2 holds for $\mathcal{G}_{\phi_{i+1}}$. Let

$$U \equiv vwg \Rightarrow^* \mathsf{tr}(g) = z'g,$$

where $v \in B^*$, $w \in I^*$, $|vw| > 0$, and g eliminates vw in U. Moreover, assume that

$$z' \in (O \cup F)^* = (\widehat{O}_2 \cup \widehat{O}_3 \cup \widehat{F}_2 \cup \widehat{F}_3 \cup \{f_2, f_3\})^* \tag{3.1}$$

W.l.o.g. assume that $w = w_1 \cdots w_n$, where $w_k \in I^+_{k\%2}$ for each $k \in [1, n]$ and n is even. As before, one can split U into *stages* $U_m, U_{m-1}, \ldots, U_0$ such that each stage applies one of the grammars $\mathcal{G}_0, \ldots, \mathcal{G}_3$ and each two consecutive stages apply different grammars from $\mathcal{G}_0, \ldots, \mathcal{G}_3$. Observe that

- $\widehat{g}_2$ and $\widehat{g}_3$ are not able to eliminate the elements of I;
- $\widehat{g}_2$ ($\widehat{g}_3$, resp.) leaves a trace which contains the elements of Σ_2 (Σ_3, resp.); and the elements of Σ_2 (Σ_3, resp.) cannot be eliminated by $\widehat{g}_0, \widehat{g}_1$ and $\widehat{g}_3$ ($\widehat{g}_2$, resp.).

The second of these observations implies that none of the grammars $\mathcal{G}_0$, $\mathcal{G}_1$, $\mathcal{G}_3$ ($\mathcal{G}_2$, resp.) can be applied after $\mathcal{G}_2$ ($\mathcal{G}_3$, resp.) is applied. Therefore, the partition of U is such that U_j uses the productions of $\mathcal{G}_0$ or $\mathcal{G}_1$ for $j > 0$ and U_0 uses $\mathcal{G}_2$ or $\mathcal{G}_3$. That is, $U_m, \ldots, U_1$ eliminate $w \in I^+$ and U_0 eliminates $v \in B^*$. Thus, let

$$vwg = z_m g \Rightarrow^*_{U_m} z_{m-1} g \Rightarrow^*_{U_{m-1}} \cdots \Rightarrow^*_{U_1} z_0 g \Rightarrow^*_{U_0} z' g,$$

i.e., $U_j \equiv (z_j g \Rightarrow^* z_{j-1} g)$ for $j > 0$ and $U_0 \equiv (z_0 g \Rightarrow^* z')$. We make use of the following claim.

Claim 3 *Assume that the assumption (3.1) is satisfied for the derivation*

$$U \equiv (vwg = z_m g \Rightarrow^*_{U_m} z_{m-1} g \Rightarrow^*_{U_{m-1}} \cdots \Rightarrow^*_{U_1} z_0 g \Rightarrow^*_{U_0} z' g).$$

Then, for each $j \in [0, m-1]$, $z_j g$ *is equal to* $vw_1 \cdots w_l v'_j w'_j g$ *such that*

(a) $v'_j \in (\widehat{F}_0 \cup \widehat{F}_1)^* \widehat{F}^+_{\gamma(j)}$, $w'_j \in \widehat{O}^+_{\gamma(j)}$;
(b) $j \geq l$;
(c) $\|w'_j\|_{\widehat{\mathcal{O}}_{\gamma(j)}} \geq \phi_i^{(m-j)}(0)$;

where $\gamma(j) = (m - j + 1)\%2$ *and* $U_j \equiv (z_j g \Rightarrow^* z_{j-1} g)$.

The proof of Claim 3 is fairly technical, we present it at the end of this section. Claim 3 implies that $U_m, \ldots, U_1$ form a subderivation $vwg \Rightarrow^* z_0 g = vv'_0 w'_0 g$ of U, where $v'_0 \in (\widehat{F}_0 \cup \widehat{F}_1)^+ \widehat{F}^+_{\gamma(0)}$, $w'_0 \in \widehat{O}^+_{\gamma(0)}$, and $\|w'_0\|_{\widehat{\mathcal{O}}} \geq \phi_i^{(m)}(0) = \phi_i^{(m-1)}(1)$. Finally, $z_0 = vv'_0 w'_0$ has to be eliminated in the stage U_0 using $\mathcal{G}_2$ or $\mathcal{G}_3$, since the final sentential form $z'g$ of U should satisfy the assumption (3.1) on page 9. That is, by the assumption $e(3 - \gamma(0))$ on page 7, $U_0 \equiv (vv'_0 w'_0 g \Rightarrow^* v'w'g)$, where $v' \in \widehat{F}_{3-\gamma(0)}$, $w' \in \widehat{O}^+_{3-\gamma(0)}$ and

$$\|w'\|_{\widehat{\mathcal{O}}_{3-\gamma(0)}} \geq \phi_i(\|w'_0\|_{\widehat{\mathcal{O}}_{\gamma(0)}}) \geq_{(\text{Claim 3(c)})} \phi_i^{(m)}(1) = \phi_{i+1}(m) \geq \phi_{i+1}(n),$$

since $m \geq n$, which follows from the fact that each application of $\mathcal{G}_j$ for $j \in [0, 3]$ allows to delete at most one block of w. Thus, we obtain the statement 2 of Definition 2 for the case $w \neq \lambda$. One can easily check that this statement holds for $w = \lambda$ as well. □

3.1 Proof of Claim 3

The main idea of the proof is as follows. First, we show that the construction of $\mathcal{G}_{\phi_{i+1}}$ and the assumptions $e(0) - e(3)$ on page 7 combined with the statement 3 of Def. 2 guarantee that either the assumption (3.1) on page 9 is not satisfied or all elements from $\widehat{O}_0 \cup \widehat{O}_1$ inserted in U_{j+1} should be deleted in U_j for each $j > 0$. If elements from $\widehat{O}_0 \cup \widehat{O}_1$ inserted in U_{j+1} are deleted in U_j, the claim holds by the assumptions $e(0) - e(3)$ on page 7.

In order to simplify the formal proof, we first make the following observation. W.l.o.g. one can assume that the sentential form at the beginning of U is equal to

$$z_m = vw_1 \cdots w_n v'_m w'_m g, \text{ where } v'_m \in \widehat{F}_1 \text{ and } w'_m = \lambda \in \widehat{O}^*_1. \quad (3.2)$$

Indeed, the suffix w_n of w has to be deleted in the stage $U_m \equiv vw_1 \cdots w_n g \Rightarrow^* z_{m-1}g$, by the assumption that we consider greedy derivations. Since the elements of $w_n \in I_0^+$ can be deleted only by f_0 and f_0 can be inserted only by the elements of $\mathsf{Del}(\widehat{F}_1) \cap \widehat{F}_0$, one can assume that a symbol from $\widehat{F}_0$ follows w_n.

Next, we show that no symbols from $\Sigma_0 \cup \Sigma_1 \setminus (\widehat{F}_0 \cup \widehat{F}_1)$ inserted *before* the stage U_j can be deleted in the stage U_{j-1}.

Claim 4 *Let $z_{j-1} = uyg$ be a sentential form at the end of U_j for $j \in [1, m]$, where y consists of symbols inserted in U_j. Moreover, let $\mathcal{G}_k$ be applied in U_j for $k \in [0, 1]$. Then,*

(A) $y \in \widehat{F}_k^+ \widehat{O}_k^+$ or $y \notin (\widehat{F}_k \cup \widehat{O}_k)^*$;
(B) no symbol from u belonging to $\Sigma_0 \cup \Sigma_1 \setminus (\widehat{F}_0 \cup \widehat{F}_1)$ will be deleted in U_{j-1}.

Proof. First, we prove (A) by induction wrt the number of the stage. Using (3.2), we combine the base step and the inductive step of the proof. By the inductive hypothesis or (3.2), we may assume that $z_j = u'y'g$ for $j \in [1, m]$, where $y' \in \widehat{F}_{k'}^+ \widehat{O}_{k'}^+$ or $y' \notin (\widehat{F}_{k'} \cup \widehat{O}_{k'})^*$ for $k' \in [0, 1]$. Thus, the grammar $\mathcal{G}_k$ is applied in U_j for $k = (k' + 1)\%2$ or $k = 2 + (k' + 1)\%2$. Moreover, the rightmost symbol of y' belongs to $\widehat{F}_{k'} \cup \widehat{O}_{k'}$. Indeed, otherwise no grammar from $(\bigcup_{p=0}^{3} \mathcal{G}_p) \setminus \{\mathcal{G}_{k'}\}$ allows to delete the rightmost symbol of y' what implies that the stage U_j does not occur in a greedy derivation at all. So, (A) holds by the statement 2 of Def. 2, since $\mathcal{G}_k$ satisfies the assumption $(e(k))$ on page 7 (see the statement 2 of Def. 2).

Now, we prove (B). Let $j \in [1, m]$, let $z_{j-1} = uyg$ as defined in the claim and let $\mathcal{G}_k$ be applied in U_j. Thus, $\mathcal{G}_{k'}$ is applied in U_{j-1} for $k' = (k + 1)\%2$ or $k' = 2 + (k + 1)\%2$. The statement (A) implies that y contains a symbol b such that $b \in \widehat{F}_k$ or $b \notin \widehat{F}_k \cup \widehat{O}_k$. Let b be the leftmost symbol which satisfies this condition. If $b \in \widehat{F}_k$, then the claim holds by the fact that $B(\mathcal{G}_{k'}) = \widehat{F}_k$, and $\mathcal{G}_{k'}$ satisfies the stat. 3 of Definition 2 (by the assumption $e(k')$ on page 7). If $b \notin \widehat{F}_k$ then the symbols from $\mathcal{G}_{k'}$ are not able to eliminate b, so no element located to the left of y is deleted in U_{j-1}, either. □

Below, we state corollary from the statement (B) of the above claim.

Corollary 1. *If a symbol $b \in \Sigma_k \setminus \widehat{F}_k$ for $k \in [0, 1]$ is inserted in U_j for $j \in [1, m]$ and it is not deleted during U_j and U_{j-1}, it is not deleted in U at all.*

We prove Claim 3(a) by contradiction. Assume that j is the minimal value such that Claim 3(a) is not satisfied. Thus, by (3.2) and the choice of j, U_{j+1} has a form

$$w_1 \cdots w_l v'_{j+1} w'_{j+1} g \Rightarrow^*_{U_{j+1}} uyg = z_j$$

where $l \in [1, n]$, u is a (possibly empty) prefix of $w_1 \cdots w_l v'_{j+1} w'_{j+1}$, y is inserted in U_{j+1}, $v'_{j+1} \in \widehat{F}^+_{\gamma(j+1)}$ and $w'_{j+1} \in \widehat{O}^*_{\gamma(j+1)}$. Since the derivation is greedy, the grammar $\mathcal{G}_k$ is applied in U_{j+1}, where $k = \overline{\gamma(j+1)} = \gamma(j)$. So, we have the following cases:

- w'_{j+1} is deleted in U_{j+1} and no element from $\Sigma_k \setminus (\widehat{O}_k \cup \widehat{F}_k)$ appears in y; then, (a) holds by Claim 4(A). This contradicts the choice of j.

- w'_{j+1} is deleted in U_{j+1} and y contains $c \in \Sigma_k \setminus (\widehat{O}_k \cup \widehat{F}_k)$; then, c is not deleted in $U_j, \ldots, U_0$ by Corollary 1. This contradicts the assumption (3.1) on page 9.
- w'_{j+1} is not deleted in U_{j+1}: then, the nonempty prefix w''_{j+1} of w'_{j+1} is not deleted in $U_j, \ldots, U_0$, by Corollary 1. This contradicts the assumption (3.1) on page 9.

If the statement (a) of Claim 3 is true, the statement (b) follows directly from the fact that symbols of the grammar $\mathcal{G}_l$ are not able to delete elements of $I_{(l+1)\%2}$. (So, at most one block from $w_1, \ldots, w_n$ can be deleted in one stage.)

Finally, we show the statement (c) by induction. Assume that z_{j+1} satisfied (c). The assumptions $(e(0)) - (e(3))$ and the fact that the whole w'_{j+1} is eliminated in U_j imply that

$$\|w'_j\|_{\widehat{\mathcal{O}}_{\gamma(j)}} \geq \phi_i(\|w'_{j+1}\|)_{\widehat{\mathcal{O}}_{\gamma(j+1)}} \geq_{\text{ind.ass.(c)}} \phi_i(\phi_i^{(m-j-1)}(0)) = \phi_i^{(m-j)}(0).$$

□

4 Shrinking

Similarly to the expansion property, we define the shrinking property which describes a way in which the functions ϕ_p^{-1} are "computed" by leftist grammars.

Definition 5. *Let I, O, F and $\{g\}$ be disjoint, finite and nonempty sets of symbols and let $\mathcal{I} = (I_0, \ldots, I_3)$, $\mathcal{O} = (O_0, \ldots, O_3)$ be partitions of I and O, respectively. Moreover, let the following conditions be satisfied in a grammar $\mathcal{G}$:*

1. *Let $w \in I^+$ be an alternating word wrt $\mathcal{I}$. Then, there exists a derivation $wg \Rightarrow^* v'w'g$ in which g eliminates w, such that $v' \in F^+$, $w' \in O^*$ is an alternating word wrt $\mathcal{O}$ and $\|w'\|_{\mathcal{O}} = \phi_p^{-1}(\|w\|_{\mathcal{I}})$.*
2. *Let $w \in I^+$ and let $U \equiv (wg \Rightarrow^* zg)$ be a derivation in which g eliminates w. Then, either*
   - *$z = v'w'$ such that $v' \in F^+$, $w' \in O^+$, and $\|w'\|_{\mathcal{O}} \geq \phi_p^{-1}(\|w\|_{\mathcal{I}})$;* ***or***
   - *$z \notin (F \cup O)^*$.*
3. *The sets $I_0, \ldots, I_3$ are homogeneous in $\mathcal{G}$.*

Then, we say that $\mathcal{G}$ satisfies p-shrinking wrt $\mathcal{I}$, $\mathcal{O}$, F and $\{g\}$, we denote this fact by $\mathsf{sh}_p(\mathcal{I}, \mathcal{O}, F, g)$.

For a grammar $\mathcal{G}$ satisfying $\mathsf{sh}_p(\mathcal{I}, \mathcal{O}, F, g)$, we use the notation $\mathcal{I}(\mathcal{G}) := \mathcal{I}$, $\mathcal{O}(\mathcal{G}) := \mathcal{O}$, and $F(\mathcal{G}) := F$. Since the construction of a grammar satisfying p-shrinking is independent of $\mathcal{I}$, we denote p-shrinking by $\mathsf{SH}_p(\mathcal{O}, F, g)$.

Let $\mathcal{G}_{\phi_2^{-1}}$ be a grammar which contains the following productions:

$$\begin{array}{rll}
(i) & g \to^{\mathsf{ins}} a_l & \text{for each } l \in [0,1] \\
(ii) & a_i \to^{\mathsf{ins}} a_{(i+1)\%4} & \\
(iii) & a_i \to^{\mathsf{ins}} f_i & \\
(iv) & a_i \to^{\mathsf{ins}} f'_i & \\
(v) & g \to^{\mathsf{ins}} f''_i & \\
(vi) & f_i \to^{\mathsf{del}} f_{(i-1)\%4} & \\
(vii) & f'_i \to^{\mathsf{del}} f'_{(i-1)\%4} & \\
(viii) & f_i \to^{\mathsf{del}} \imath & \text{for each } \imath \in I_j \text{ such that } i\%\,2 = \lfloor j/2 \rfloor \\
(ix) & f'_i \to^{\mathsf{del}} \imath & \text{for each } \imath \in I_j \text{ such that } i\%\,2 = |\lfloor (j-1)/2 \rfloor| \\
(x) & f''_i \to^{\mathsf{del}} \imath & \text{for each } \imath \in I_i.
\end{array}$$

for each $i \in [0,3]$.

Lemma 3. *The grammar* $\mathcal{G}_{\phi_2^{-1}}$ *satisfies* $\mathsf{sh}_2(\mathcal{O}, F, g)$, *where* $F = \bigcup_{i=0}^{3}\{f_i, f_i', f_i''\}$ *and* $\mathcal{O} = (O_0, \ldots, O_3)$ *is a partition of* $O = \bigcup_{j=0}^{3} O_j$ *and* $O_i = \{a_i\}$ *for* $i \in [0,3]$.

Proof. First, note that the insert graph of $\mathcal{G}_{\phi_2^{-1}}$ presented at Figure 4 exhibits similarities in the construction of $\mathcal{G}_{\phi_2}$ and $\mathcal{G}_{\phi_2^{-1}}$. As in the construction of $\mathcal{G}_{\phi_2}$, the cycle $a_0a_1a_2a_3$ plays the key role in the construction. The main difference between $\mathcal{G}_{\phi_2}$ and $\mathcal{G}_{\phi_2^{-1}}$ is that here each of $a_0, \ldots, a_3$ is able to eliminate symbols of two sets from $\widehat{O}_0, \ldots, \widehat{O}_3$ while in $\mathcal{G}_{\phi_2}$, only each second symbol in $a_0, \ldots, a_3$ is able to eliminate elements of one of the sets $\widehat{O}_0, \widehat{O}_1$. Below, we prove the lemma

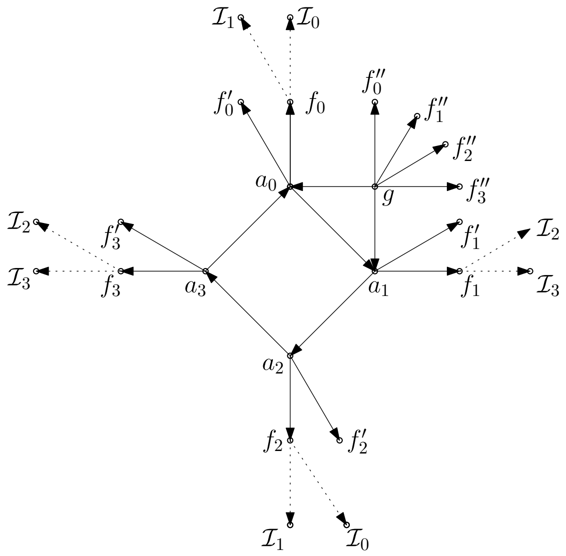

Fig. 4. The insert graph and a part of the delete graph of $\mathcal{G}_{\phi_2^{-1}}$. Dotted edges belong to the delete graph.

by checking whether $\mathcal{G}_{\phi_2^{-1}}$ satisfies the statements of Def. 5.

First, observe that restrictions of $(viii) - (x)$ imply that the statement 3 of Definition 5 is satisfied.

Now, let $w = w_1 \cdots w_n$, where $n > 1$, $w_i \in I^+_{(k+n-i)\%4}$ for some $k \in [0,3]$ and each $i \in [1,n]$. Assume that $k = 0$. Then, g can eliminate w in the following derivation:

1. $w_1 \cdots w_n g \Rightarrow_{(i)} w_1 \cdots w_n a_0 g \Rightarrow_{(iii)} w_1 \cdots w_n f_0 a_0 g \Rightarrow_{(viii)} w_1 \cdots w_{n-2} f_0 a_0 g$.
2. For $j = 1, 2, \ldots, \lceil \frac{n-2}{2} \rceil$ do: **[delete** $w_{n-2j-1}w_{n-2j}$**]**

$$w_1 \cdots w_{n-2j} u_j y^{\lfloor j/4 \rfloor} g \Rightarrow^*_{(ii,iii,viii)} w_1 \cdots w_{n-2(j+1)} u_{j+1} y^{\lfloor (j+1)/4 \rfloor} g,$$

where $u_j = \begin{cases} f_0 a_0 & \text{if } j\%4 = 1 \\ f_1 a_1 a_0 & \text{if } j\%4 = 2 \\ f_2 a_2 a_1 a_0 & \text{if } j\%4 = 3 \\ f_3 & \text{if } j\%4 = 0 \end{cases}$ and $y = a_3a_2a_1a_0$.

Thus, we obtain a derivation $wg \Rightarrow^* v'w'g$, where $v' \in F$, w' is an alternating word over $(O_0, \ldots, O_3)$ and $\|w'\|_{\mathcal{O}} = \lceil \|w\|_{\mathcal{I}}/2 \rceil = \phi_2^{-1}(\|w\|_{\mathcal{I}})$.

If $k \neq 0$, then one can design an analog derivation. That is, if $k \in [2,3]$, we start with the application of the rule $g \rightarrow^{\mathsf{ins}} a_1$, instead of $g \rightarrow^{\mathsf{ins}} a_0$. Moreover, if $k\%2 \neq 0$, the symbols $f'_0, \ldots, f'_3$ are inserted instead of $f_0, \ldots, f_3$.
Finally, if $w \in I_j^+$ for $j \in [0,3]$, i.e. $\|w\|_{\mathcal{I}} = 1$, w can be eliminated in the following derivation

$$wg \Rightarrow_{(v)} wf''_j g \Rightarrow^*_{(x)} f''_j g = v'w'g,$$

where $v' = f''_j \in F$, $w' = \lambda$. That is, $\|w'\|_{\mathcal{O}} = 0 = \phi_2^{-1}(1) = \phi_2^{-1}(\|w\|_{\mathcal{I}})$. So, the statement 1 of Def. 5 holds.

It remains to show that $\mathcal{G}_{\phi_2^{-1}}$ satisfies the statement 2 of Definition 2. Let $w = w_1 \cdots w_n$, where $w_i \in I^+_{(j+n-i)\%4}$ for each $i \in [1,n]$ and some $j \in [0,3]$. That is, $\|w\|_{\mathcal{I}} = n$. Assume that $j = 0$. Let $U \equiv (wg \Rightarrow^* zg)$ be a derivation in which g eliminates w.

First, assume that symbols from the set $\{f''_i\}_{i=0}^3$ are not inserted during U. Similarly as in Lemma 1, the restrictions of $\mathcal{G}_{\phi_2^{-1}}$ imply that the (greedy) derivation U satisfies the following conditions:

– elements of O are inserted from left to right using (ii) ((i) can be applied only once);
– elements of O are not deleted;
– elements of F can appear only to the left of all elements of O;
– only the leftmost element from O can be active in each derivation step of U.

These restrictions and the constraints of $(i) - (x)$ imply that $\mathsf{tr}(g) = v'w'g$ for $v' \in F^+$, $w' \in O^+$ and w' is a suffix of a word from the set

$$(a_3a_2a_1a_0)^* + (a_0a_3a_2a_1)^*.$$

Let $F_0 = \{f_i\}_{i\in[0,3]}$, $F_1 = \{f'_i\}_{i\in[0,3]}$. Note that elements of F_i are not able to eliminate elements of F_{1-i} for $i \in [0,1]$. Therefore, if a symbol from F_i is inserted during U, no symbol from F_{1-i} is inserted during U for $i \in [0,1]$. Thus, each symbol $a \in O$ inserted during U inserts at most one element of F. As each element of F can delete at most two blocks of $w = w_1 \cdots w_n$, we obtain

$$wg \Rightarrow^* \mathsf{tr}(g) = v'w'g,$$

such that $v' \in F^+$, $w' \in O^+$ is an alternating word over $(O_0, \ldots, O_3)$ and

$$\|w'\|_{\mathcal{O}} = |w'| \geq \lceil \|w\|_{\mathcal{I}}/2 \rceil = \phi_2^{-1}(\|w\|_{\mathcal{I}}).$$

Finally, consider the case that an element of $F_2 = \{f''_0, \ldots, f''_3\}$ is inserted during U. Observe that

– the elements of F_2 can be inserted only by g;
– g is not able to eliminate any element of F_2;
– elements of F_2 cannot eliminate elements of $O \cup F$.

Therefore, if g inserts an element of F_2, no other element is inserted during U. Thus,

$$U \equiv (wg \Rightarrow_{(v)} wf''_i g \Rightarrow^*_{(x)} f''_i g),$$

for $i \in [0,3]$. The above derivation is possible only in the case that $w \in I_i^+$. That is $f''_i g = v'w'g$, where $v' = f''_i$, $w' = \lambda$ and $\|w'\|_{\mathcal{O}} = 0 = \phi_2^{-1}(1) = \phi_2^{-1}(\|w\|_{\mathcal{I}})$. □

Lemma 4. *Assume that a grammar $\mathcal{G}_{\phi_i^{-1}}$ satisfies $\mathsf{SH}_i(\widehat{\mathcal{O}}, \widehat{F}, \widehat{g})$. Then, one can build a grammar $\mathcal{G}_{\phi_{i+1}^{-1}}$ of size at most $6|\mathcal{G}_{\phi_i^{-1}}|$ which satisfies $\mathsf{SH}_{i+1}(\mathcal{O}, F, g)$ for some sets O, F, $\{g\}$ and the partition $\mathcal{O} = (O_0, \dots, O_3)$ of O.*

Proof. Let $\mathcal{G}_0, \dots, \mathcal{G}_4$ be copies of the grammar $\mathcal{G}_{\phi_i^{-1}}$ such that:

- $\mathcal{G}_j$ satisfies $\mathsf{SH}_{\phi_i^{-1}}(\widehat{\mathcal{O}}_j, \widehat{F}_j, \widehat{g}_j)$,
- $\widehat{\mathcal{O}}_j = (\widehat{O}_{j,0}, \dots, \widehat{O}_{j,3})$ is a partition of $\widehat{O}_j$;
- $\Sigma_j \cap \Sigma_l = \emptyset$ for $j \neq l$, where Σ_k is the set of symbols accessible from $\widehat{g}_k$ in the insert graph of $\mathcal{G}_k$.

Let $\mathcal{I}(\mathcal{G}_4) = \widehat{\mathcal{I}}_4 := \mathcal{I}$ and $\mathcal{I}(\mathcal{G}_j) = \widehat{\mathcal{I}}_j := \widehat{\mathcal{O}}_{(j-1)\%4} \cup \widehat{\mathcal{O}}_4$ for $j \in [0,3]$. That is,

$$\begin{array}{lll} \mathcal{G}_j & \text{satisfies} & \mathsf{sh}_i(\widehat{\mathcal{O}}_{(j-1)\%4} \cup \widehat{\mathcal{O}}_4, \widehat{\mathcal{O}}_j, \widehat{F}_j, \widehat{g}_j) \text{ for } j \in [0,3] \quad \mathbf{(s(\boldsymbol{j}))} \\ \mathcal{G}_4 & \text{satisfies} & \mathsf{sh}_i(\mathcal{I}, \widehat{\mathcal{O}}_4, \widehat{F}_4, \widehat{g}_4) \quad \mathbf{(s(4))} \end{array}$$

We add a new symbol g and the production rules:

$$\begin{array}{l} (i)\ g \rightarrow^{\mathsf{ins}} \widehat{g}_j \text{ for } j \in [0,4] \quad \text{and} \\ (ii)\ g \rightarrow^{\mathsf{del}} \widehat{g}_j \text{ for } j \in [0,4]. \end{array}$$

Given a set I and a partition $\mathcal{I} = (I_0, \dots, I_3)$ of I, the grammar $\mathcal{G}_{\phi_{i+1}}$ is obtained by combining the alphabets and the rules of $\mathcal{G}_0, \dots, \mathcal{G}_4$, adding g to the alphabet and the rules $(i) - (ii)$. We claim that $\mathcal{G}_{\phi_{i+1}^{-1}}$ defined in this way satisfies $\mathsf{sh}_{i+1}(\mathcal{I}, \mathcal{O}, F, g)$, where $\mathcal{O} = (\widehat{F}_0, \dots, \widehat{F}_3)$ is a partition of $O = \bigcup_{j=0}^3 \widehat{F}_j$ and $F = \widehat{F}_4$. That is,

$$\begin{array}{l} \mathcal{O}(\mathcal{G}_{\phi_{i+1}^{-1}}) := \mathcal{O} = (\widehat{F}_0, \dots, \widehat{F}_3) \text{ and} \\ F(\mathcal{G}_{\phi_{i+1}^{-1}}) := \widehat{F}_4. \end{array}$$

The construction is based on the following observations:

- the elements of $w \in I^*$ can be deleted only by productions of $\mathcal{G}_4$;
- if $\widehat{g}_4$ eliminates $w \in I^*$ leaving the trace in $(\widehat{F}_4 \cup \widehat{O}_4)^* \widehat{g}_4$, then the trace of $\widehat{g}_4$ is equal to $v'w'$, where $v' \in \widehat{F}_4$, $\|w'\|_{\widehat{\mathcal{O}}_4} \geq \phi_p^{-1}(\|w\|_{\mathcal{I}})$;
- each alternating word $w' \in \widehat{O}_{j'}^*$ for $j' \in [0,4]$ can be eliminated by $\widehat{g}_{j''}$ such that $j'' = (j-1)\%4$, w' is replaced with $v''w''$ satisfying $v'' \in \widehat{F}_{j''}$, $w'' \in O_{j''}$ and the alternation measure of w'' wrt $\widehat{\mathcal{O}}_{j''}$ is at least $\phi_p^{-1}(\|w'\|_{\widehat{\mathcal{O}}_{j'}})$.

Thus, in order to delete w and elements of $\bigcup_{j=0}^4 \widehat{O}_j$ inserted during the derivation, $\mathcal{G}_4$ has to be applied first and then the grammars $\mathcal{G}_j$, $\mathcal{G}_{(j+1)\%4}$, $\mathcal{G}_{(j+2)\%4}$, and $\mathcal{G}_{(j+3)\%4}$ should be applied in a cycle for some $j \in [0,3]$. The construction of $\mathcal{G}_{\phi_{i+1}^{-1}}$ guarantees that this scenario is the only way to eliminate $w \in I^*$ by

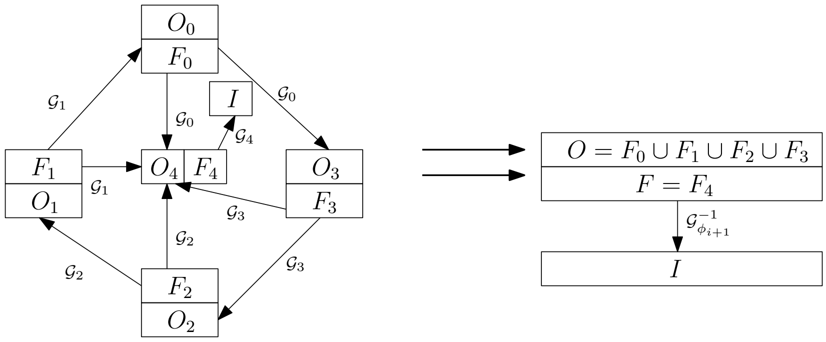

Fig. 5. An illustration for the construction of $\mathcal{G}^{-1}_{\phi_{i+1}}$. An edge from a vertex A to a vertex B labeled with C denotes that the A is the "frontier" alphabet of the grammar C and B is a subset of the "input" alphabet of C, i.e., $A = F(C)$ and $B \subseteq \mathcal{O}(C)$.

g, such that the trace of g contains only the elements of $\widehat{F}_0 \cup \cdots \cup \widehat{F}_4$. The key point here is that each application of $\mathcal{G}_j$ for $j \in [0,3]$ has to leave a symbol from $\widehat{F}_j$, so the number of symbols from $\widehat{F}_0 \cup \cdots \cup \widehat{F}_3$ in the sentential form is a kind of a counter denoting the number of applications of the function ϕ_p^{-1}.

Below, we present a formal proof of the lemma by checking whether all statements of Def. 5 are satisfied for $\mathcal{G}^{-1}_{\phi_{i+1}}$.

First, note that the statement 3 of Def. 5 follows from the assumption that $\mathcal{G}_4$ satisfies $\mathsf{sh}_i(\mathcal{I}, \widehat{\mathcal{O}}_4, \widehat{F}_4, \widehat{g}_4)$, and the fact that only symbols from $\mathcal{G}_4$ are able to eliminate symbols from I.

Next, we show that the statement 1 of Def. 5 holds. W.l.o.g., assume that $w = w[1] \cdots w[n]$, where $w[j] \in I_{j\%4}$ for each $j \in [1,n]$. Consider a derivation described by the following algorithm:

1. eliminate w using $\mathcal{G}_4$:
$$wg \Rightarrow^*_{\text{ind.ass.}} v_0 w_0 \widehat{g}_4 g \Rightarrow^*_{(ii)} v_0 w_0 g,$$
such that $v_0 \in \widehat{F}_4^+$, $w_0 \in \widehat{O}_4^*$, and $\|w_0\|_{\widehat{\mathcal{O}}_4} = \phi_i^{-1}(\|w\|_{\mathcal{I}})$.
2. $k := 0$
3. As long as $|w_k| > 0$, repeat:
   (a) eliminate w_k, using $\mathcal{G}_{(k+1)\%4}$ (stage U_{k+1}):
   $$w_k g \Rightarrow_{(i)} w_k \widehat{g}_{k'} g \Rightarrow^*_{\text{ind.ass.}} v_{k+1} w_{k+1} \widehat{g}_{k'} g \Rightarrow_{(ii)} v_{k+1} w_{k+1} g,$$
   such that $k' = (k+1)\%4$, $v_{k+1} \in \widehat{F}_{k'}^+$, $w_{k+1} \in \widehat{O}_{k'}^+$, $\|w_{k+1}\|_{\widehat{\mathcal{O}}_{k'}} = \phi_i^{-1}(\|w_k\|_{\widehat{\mathcal{O}}_{k\%4}})$ for $k > 0$ and $\|w_1\|_{\widehat{\mathcal{O}}_1} = \phi_i^{-1}(\|w_0\|_{\widehat{\mathcal{O}}_4})$.
   (b) $k := k + 1$

The above algorithm describes a derivation
$$wg \Rightarrow^* v_0 w_0 g \Rightarrow^* v_0 v_1 w_1 g \Rightarrow^* v_0 v_1 v_2 \ldots v_r w_r g,$$
where $v_0 \in \widehat{F}_4^+$, $v_k \in \widehat{F}_{k\%4}^+$ for $k > 0$, $w_r = \lambda$ and $r + 1 = \min_s(\phi_i^{(s)}(0)) \geq \|w\|_{\mathcal{I}} = n$. Since $\phi_i^{(s)}(0) = \phi_i^{(s-1)}(1) = \phi_{i+1}(s-1)$, we see that $r = \phi_{i+1}^{-1}(n)$, where $n = \|w\|_{\mathcal{I}}$. This shows that $\mathcal{G}_{\phi^{-1}_{i+1}}$ satisfies the statement 1 of Def. 5.

It remains to show that $\mathcal{G}_{\phi_{i+1}^{-1}}$ satisfies the statement 2 of Def. 5. Let $w \in I^+$ be an alternating word wrt $\mathcal{I}$ and let

$$U \equiv (wg \Rightarrow^* zg) \text{ for } z \in (F \cup O)^* = (\bigcup_{j=0}^{4} \widehat{F}_j)^*. \tag{4.1}$$

be a derivation of $\mathcal{G}_{\phi_{i+1}^{-1}}$, where g eliminates w. Let the production rules (i)-(ii) for $j \in [0,4]$ be considered as the rules of $\mathcal{G}_j$. Then, one can split U into stages $U_0, \ldots, U_m$ such that U_j applies the rules of one of $\mathcal{G}_0, \ldots, \mathcal{G}_4$ for each $j \geq 0$ and each two consecutive stages apply different subgrammars from the set $\{\mathcal{G}_l\}_{l \in [0,4]}$. Observe that

- the productions of $\mathcal{G}_0 \cup \ldots \cup \mathcal{G}_3$ do not allow to delete elements of I;
- the productions of $\mathcal{G}_4$ do not allow to delete elements inserted by the productions of $\mathcal{G}_0 \cup \ldots \cup \mathcal{G}_3$.

Therefore, U_0 should apply productions of $\mathcal{G}_4$ and none of $U_1, \ldots, U_m$ applies $\mathcal{G}_4$. Moreover, the whole word w should be deleted in U_0 (recall that no element of I appears in the final sentential form of U, by the assumption (4.1)).

Let $U_j \equiv (z_{j-1}g \Rightarrow^* z_j g)$ for $j > 0$ and $U_0 \equiv (wg \Rightarrow^* z_0 g)$. We claim that the following properties are satisfied:

P1. $z_0 = v_0 w_0$, where $v_0 \in \widehat{F}_4^+$, $w_0 \in \widehat{O}_4^+$ and $\|w_0\|_{\widehat{\mathcal{O}}_4} \geq \phi_i^{-1}(\|w\|_{\mathcal{I}})$.

P2. $z_j = v_0 v_1 \ldots v_j w_j$ for $j > 0$ such that $v_j \in \widehat{F}_{\gamma(j)}^+$, $w_j \in \widehat{O}_{\gamma(j)}^+$ and $\|w_j\|_{\widehat{\mathcal{O}}_{\gamma(j)}} \geq (\phi_i^{-1})^{(j)}(\|w\|_{\mathcal{I}})$, where $\gamma(j) = (j + k - 1)\%4$ for some fixed $k \in [0,3]$.

First, we show that P1 holds. From the above consideration and 4.1, we know that w should be eliminated in U_0, using $\mathcal{G}_4$. It remains to show that $\mathsf{tr}(\widehat{g}_4) \in \widehat{F}_4^+ \widehat{O}_4^+$ in U_0. Recall that the grammars $\mathcal{G}_0, \ldots, \mathcal{G}_3$ do not allow to delete symbols from $\Sigma_4 \setminus \widehat{O}_4$ and only the grammars $\mathcal{G}_0, \ldots, \mathcal{G}_3$ are applied in $U_1, \ldots, U_m$. Therefore, if $\mathsf{tr}(\widehat{g}_4)$ contains symbols outside of $\widehat{O}_4 \cup \widehat{F}_4$, these symbols will not be deleted in U. This contradicts the assumption (4.1). So, due to the assumption s(4) on page 15, U_0 finishes with a sentential form $v_0 w_0$ satisfying P1 (by the stat. 2 of Def. 5).

Now, we make the inductive step which proves the statement P2. Assume that the stage U_j is finished by a sentential form $z_j = v_0 v_1 \ldots v_j w_j$, satisfying P2 (if $j > 0$) or P1 (if $j = 0$). If $j = 0$ then U_{j+1} can apply any of $\mathcal{G}_0, \ldots, \mathcal{G}_3$, say $\mathcal{G}_k$. If $j \neq 0$, then U_{j+1} should apply the grammar $\mathcal{G}_{\gamma(j+1)}$, since we consider greedy derivations (i.e., the rightmost symbol of w_j should be deleted in U_{j+1}). Thus, the symbol $\widehat{g}_{\gamma(j+1)}$ is introduced at the beginning of U_{j+1}, this symbol is able to eliminate w_j but it is not able to eliminate any symbol of v_j. The following claim says that the whole w_j has to be eliminated in U_{j+1}.

Claim 6 *Assume that the grammar $\mathcal{G}_l$ is applied in the stage*

$$U_j \equiv (z_{j-1}g \Rightarrow^* z_j g = uyg),$$

where y is a word inserted in U_j. Then, no element of u is deleted in U.

Proof. The assumption $s(l)$ on page 15 and the fact that we consider greedy derivations imply that y contains some elements from $\widehat{F}_l$ or elements from $\Sigma_l \setminus (\widehat{F}_l \cup \widehat{O}_l)$, which follows from the statement 2 of Def. 5. Furthermore, $\mathcal{G}_{(l+1)\%4}$ is the only grammar among $\{\mathcal{G}_k\}_{k\in[0,4]} \setminus \{\mathcal{G}_l\}$, which allows to delete elements inserted by $\mathcal{G}_l$. Since we consider greedy derivations, U_{j+1} applies the grammar $\mathcal{G}_{(l+1)\%4}$. But $\mathcal{G}_{(l+1)\%4}$ is not able to delete elements inserted by $\mathcal{G}_l$, except of the elements of $\widehat{O}_l$. On the other hand, at least one element outside of $\widehat{O}_l$ is inserted in U_j (see stat. 2 of Def. 5). Thus, only symbols introduced in the previous stage, U_j, can be deleted in U_{j+1} for each $j \in [0, m-1]$. □

So, by Claim 6, if an nonempty prefix of $w_j \in \widehat{O}^+_{\gamma(j)}$ is not deleted in U_{j+1}, it will not be deleted in the whole U. This contradicts the assumption (4.1). Thus, by the assumption $(s(j+1))$, we see that

$$U_{j+1} \equiv (v_j w_j g \Rightarrow^* v_j v_{j+1} w_{j+1} g),$$

such that $v_{j+1} \in \widehat{F}_{\gamma(j+1)}$, $w_{j+1} \in \widehat{O}_{\gamma(j+1)}$, and

$$\|w_{j+1}\|_{\widehat{\mathcal{O}}_{\gamma(j+1)}} \geq \phi_i^{-1}(\|w_j\|_{\widehat{\mathcal{O}}_{\gamma(j)}}) \geq (\phi_i^{-1})^{(j+1)}(\|w\|_{\mathcal{I}}).$$

This finishes the proof of P1, P2 and the statement 2 of Def. 5 for $\mathcal{G}^{-1}_{\phi_i}$. □

5 Reduction

In this section, we prove that the variable membership problem for leftist grammars is nonprimitive recursive and that there is no primitive recursive upper bound for the complexity of the ("static") membership problem.

One can show by standard methods of computability theory that the problem whether a one-tape Turing machine M halts on an empty input in $Ack(|M|)$ space is nonprimitive recursive. A question whether a one-tape TM halts on empty input in $f(n)$ space can be reduced to the question if a linear-bounded automaton (LBA) M' simulating M step by step accepts a word $\flat^{f(n)}$ where $\flat$ is a blank symbol. In [4], a reduction from the membership problem for LBAs to the membership problem for leftist grammars is presented. We combine a grammar obtained as a result of this reduction with grammars which "compute" the functions ϕ_p and ϕ_p^{-1}. As a result, we reduce (in exponential time) the problem whether a Turing Machine M accepts an empty word in space $Ack(|M|)$ to the membership problem for a leftist grammar.

Corollary 2. *The variable membership problem for leftist grammars is non-primitive recursive.*

Proof. First, we recall some results concerning leftist grammars.

Theorem 1. *[4][1] Let M be a linear-bounded automaton (LBA) with an input alphabet Σ_M, and let $\mathcal{I} = (\bigcup_{a\in\Sigma_M} I(a,0), \bigcup_{a\in\Sigma_M} I(a,1))$ be a partition of the set $I = \bigcup_{a\in\Sigma_M, i\in[0,1]} I(a,i)$. For a word $w \in \Sigma_M^*$ of length n, let*

$$R(w) = \{v_1 \cdots v_n \mid v_i \in I(w[i], i\%2) \text{ for each } i \in [1,n]\}.$$

[1] This result corresponds to a part of the main construction from [4].

Then, there exists a grammar $\mathcal{G}_M$ of size poly$(|M|)$ and a set O with a partition $\mathcal{O} = (O_0, \ldots, O_3)$, such that $I \cap O = \emptyset$, $I(a, j)$ is homogenous in $\mathcal{G}_M$for each $a \in \Sigma_M$, $j \in [0, 1]$. Moreover, $w \in L(M)$ iff the following conditions are satisfied for each $z \in R(w)$:

1. *there exists a derivation $zg \Rightarrow^*_{\mathcal{G}_M} w'g$ such that g eliminates z, $|w'| = |z|$ and w' is an alternating word wrt $\mathcal{O}$;*
2. *if g eliminates z in a derivation $zg \Rightarrow^*_{\mathcal{G}_M} w'g$, then*
   - *$w' \notin O^*$* ***or***
   - *w' is an alternating word wrt $\mathcal{O}$, and $\|w'\|_{\mathcal{O}} \geq |z| = n$.*

(Note that, due to the fact that $I(a, j)$'s are homogeneous, the above statements hold for each $z \in R(w)$ iff they hold for any $z \in R(w)$.)
If $\mathcal{G}_M$ satisfies these conditions, we denote this by $\mathsf{cmp}_M(\mathcal{I}, \mathcal{O}, g)$.

Theorem 2. *[5][2] Let I be a set of symbols and let $\mathcal{I} = (I_0, I_1)$ be a partition of I. Then, there exists a grammar $\mathcal{G}$ and a set O with the partition $\mathcal{O} = (O_0, O_1)$ such that elements of O cannot insert other symbols and $z \in (O_0 \cup O_1)^* O_j$ is able to eliminate $w \in (I_0 \cup I_1)^* I_j$ iff $\|z\|_{\mathcal{O}} \geq \|w\|_{\mathcal{I}}$. This property is denoted by* $\mathsf{eq}(\mathcal{I}, \mathcal{O})$.

For a Turing machine M and natural numbers p, n, we build the grammar $\mathcal{H}_M$ verifying if M halts on an empty input word in space $\phi_p(n)$. The grammar $\mathcal{H}_M$ consists of:

$$\begin{array}{lll} \mathcal{G}_1 & \text{which satisfies} & \mathsf{ex}_p\,(\mathcal{I}_1, B_1, \mathcal{O}_1, F_1, g_1) \\ \mathcal{G}_2 & \text{which satisfies} & \mathsf{cmp}_{M'}\,(\mathcal{I}_2, \mathcal{O}_2, g_2) \\ \mathcal{G}_3 & \text{which satisfies} & \mathsf{sh}_p\,(\mathcal{I}_3, \mathcal{O}_3, F_3, g_3) \\ \mathcal{G}_4 & \text{which satisfies} & \mathsf{eq}\,(\mathcal{I}_4, \mathcal{O}_4). \end{array}$$

where M' is an LBA simulating M. These grammars are connected by the following relationships:

$$\begin{array}{l} \mathcal{I}_2 := (I_2(\flat, 0), I_2(\flat, 1)), \text{ where } I_2(\flat, i) := O_{1,i} \\ \mathcal{I}_3 := \mathcal{O}_2; \\ \mathcal{I}_4 := (O_{3,0} \cup O_{3,2}, O_{3,1} \cup O_{3,3}) \end{array}$$

where $\mathcal{O}_i = (O_{i,0}, \ldots, O_{i,j_i})$, $O_i = O_{i,0} \cup \cdots \cup O_{i,j_i}$, $j_1 = j_4 = 1$, and $j_2 = j_3 = 3$. The grammar $\mathcal{H}_M$ is obtained by combining alphabets and productions of $\mathcal{G}_1, \ldots, \mathcal{G}_4$, adding symbols $\imath_0$, $\imath_1$, b such that $\mathcal{I}_1 = (\{\imath_0\}, \{\imath_1\})$, $B_1 = \{b\}$, a new final symbol x and new production rules

$$\begin{array}{rll} (i) & x \to^{\mathsf{del}} a & \text{for each } a \in O_{4,0} \cup O_{4,1} \cup F_1 \cup F_3 \cup \{g_3\} \\ (ii) & g_i \to^{\mathsf{del}} g_{i-1} & \text{for } i \in [2, 3] \\ (iii) & a \to^{\mathsf{del}} g_3 & \text{for each } a \in O_{4,0} \cup O_{4,1}. \end{array}$$

Now, the question whether M accepts with empty input in space $\phi_p(n)$ is reduced to the question whether the word

$$R(M, p, n) \equiv \psi_1(n) g_1 g_2 g_3 \psi_4(n) x$$

[2] This theorem describes a simple generalization of a grammar presented in the proof of Theorem 2 in [5]

belongs to the language $L(\mathcal{H}_M)$, where

$$\psi_1(n) = b\imath_0^{n\%2}(\imath_1\imath_0)^{\lfloor n/2\rfloor}, \psi_4(n) = a_0^{n\%2}(a_1a_0)^{\lfloor n/2\rfloor}$$

and $a_j \in O_{4,j}$ for $j \in [0, 1]$.

If M halts on the empty input in space $\phi_p(n)$, then there exists a derivation $R(M, p, n) \Rightarrow^* x$ consisting of the following stages (see Figure 6):

1. Expansion: $\psi_1(n)g_1 \Rightarrow^*_{\mathcal{G}_1} v_1w_1g_1$, where $v_1 \in F_1^+$, $w_1 \in [O_{1,0}](O_{1,1}O_{1,0})^*$ and $\|w_1\|_{\mathcal{O}_1} = \phi_p(n)$.
2. Computation: $v_1w_1g_1g_2 \Rightarrow^* v_1w_2g_2$, where $w_2 \in O_{2,0}(O_{2,1}O_{2,0})^*$ and $\|w_2\|_{\mathcal{O}_2} = \|w_1\|_{\mathcal{O}_1}$. Here, w_1 is considered as an element of $R(w)$, where $w = \flat^{\phi_p(n)}$.
3. Shrinking: $v_1w_2g_2g_3 \Rightarrow^* v_1v_3w_3g_3$, where $\|w_3\|_{\mathcal{O}_3} = \phi_p^{-1}(\|w_2\|_{\mathcal{O}_2}) = \phi_p^{-1}(\phi_p(n)) = n$.
4. Verification: $v_1v_3w_3g_3\psi_4(n) \Rightarrow^* v_1v_3w_4$, where $w_4 \in (O_{4,0} \cup O_{4,1})^*$.
5. Final steps: $v_1v_3w_4x \Rightarrow^* x$.

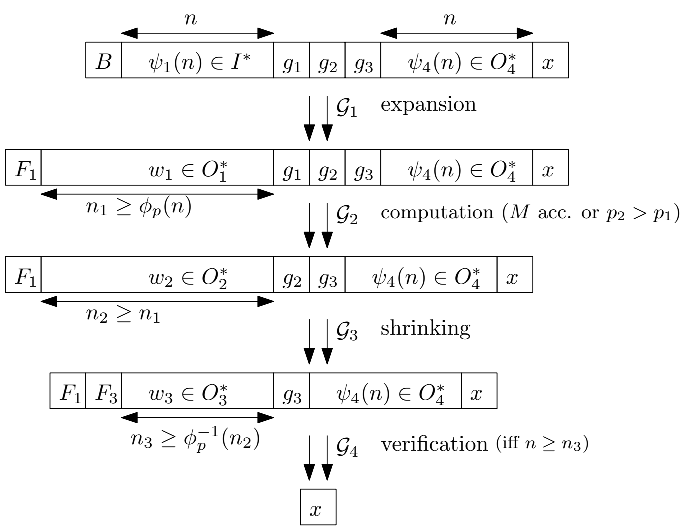

Fig. 6. How the reduction works.

Now, the goal is to show that, if there exists a derivation

$$U \equiv (R(M, p, n) \Rightarrow^* x), \text{ then } M \text{ accepts in space } \phi_p(n). \tag{5.1}$$

Note that, according to the construction of $\mathcal{H}_M$, g_i has to eliminate all symbols from O_{i-1} introduced by g_{i-1} for $i = 2, 3$. Thus, the derivation U can be split into five stages applying $\mathcal{G}_1, \ldots, \mathcal{G}_4$ and the productions $(i) - (iii)$, respectively. Then, (5.1) follows from the assumptions that $\mathcal{G}_1$ and $\mathcal{G}_3$ satisfy p-expansion and p-shrinking, respectively and g_2 satisfies $\mathsf{cmp}_{M'}(\mathcal{I}_2, \mathcal{O}_2, g_2)$. Indeed, the key property is that the functions ϕ_p and ϕ_p^{-1} are "approximated" in such a way that the result may be larger than the actual value of the function, but not smaller than this value. Similarly, if g_2 eliminates a word over O_1, the alternation measure of the eliminated word wrt $\mathcal{O}_1$ is not larger than the alternation measure of the inserted word wrt $\mathcal{O}_2$. Finally, $\psi_4(n)$ is able to eliminate a word over $w_3 \in O_3^*$ iff $\|w_3\|_{\mathcal{I}_4} \leq n$. Therefore, if any of the first three stages works differently as in the above scenario (described by the stages expansion, computation, shrinking, verification and final steps), some elements of $O_1 \cup O_2 \cup O_3$ will not be deleted.

By applying the reduction $R(M, p, n)$ for $p := |M|$ and $n := 3$, we see that the variable membership problem for leftist grammars is nonpritive recursive, since $Ack(m) = \phi_m(3)$.

□

Using a bit more complicated reduction than the above one, we can prove the following result.

Corollary 3. *There is no primitive recursive upper bound for the ("static") membership problem for leftist grammars.*

Proof. It is known that, for each primitive recursive function f, there exists $p \in \mathbb{N}$ such that $f(n) = O(\phi_p(n))$.

For a fixed Turing machine M working in space $\phi_p(n)$ where n is the size of an input word, let the membership problem be to decide, given an input word w of lengths n, if M accepts w. Here, we assume that M is a one-tape Turing machine and that the cells containing an input word are read-only. We describe a polynomial time reduction of the membership problem for a Turing machine M to the membership problem for a leftist grammar.

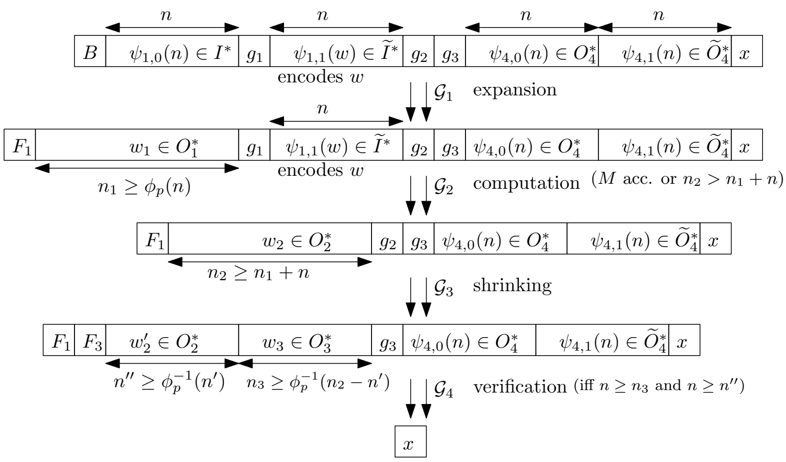

Fig. 7. How the second reduction works (here, we assume that $\widetilde{I} \subset O_1$).

The reduction is

$$R'(M, p, n) \equiv \psi_1(w, n) g_1 g_2 g_3 \psi_4(n) x,$$

where $\psi_1, \psi_4, g_1, g_2, g_3$ are defined similarly as in $R(M, p, n)$, with the following modifications (see Figure 7):

- $\psi_1(n)$ consists of two blocks: $\psi_{1,0}(n) \in I_1^*$ such that $\|\psi_{1,0}(n)\|_{\mathcal{I}} = n$ and $\psi_{1,1}(w) \in \tilde{I}^*$ such that $\|\psi_{1,1}(w)\|_{\mathcal{O}_1} = n$, $\tilde{I} \subseteq O_1$ and $\psi_{1,1}(w)$ "encodes" w (the encoding is similar to the function $R(w)$ in Theorem 1);
- we expect that g_3 eliminates the suffix of the word $w_2 \in O_2^*$ of length $\phi_p(n)$, where $\|w_2\|_{\mathcal{O}_2} \geq \phi_p(n) + n$;
- $\psi_4(n)$ consists of two blocks: $\psi_{4,0}(n)$ which is able to eliminate each $z \in \mathcal{O}_3^*$ such that $\|z\|_{\mathcal{O}_3} \leq n$ and $\psi_{4,1}(n)$ which is able to eliminate each $z \in O_2^*$ such that $\|z\|_{\mathcal{O}_2} \leq n$.

The result holds by the arguments described in the proof of Corollary 2. □

6 Conclusions and Open Problems

We have shown that the membership problem for leftist grammars is nonprimitive recursive. This implies that the accessibility problem in the Saraswat's model is nonprimitive recursive as well. An interesting research direction is to find a restricted variant of the Saraswat's model with feasible complexity and large expressive power.

Interestingly, large complexity of the membership problem for leftist grammars can be obtained by a technique which is similar to the method applied in the proof that verifying lossy channels has nonprimitive recursive complexity [9]. In both models, one can "compute" recursive functions (only) approximately. However, the models allow verifying whether the result of the computation is correct.

References

1. S. Bandyopadhyay, M. Mahajan, K. Narayan Kumar. *A non-regular leftist language.* Manuscript. 2005.
2. O. Cheiner, V. Saraswat. *Security Analysis of Matrix.* Technical report, AT&T Shannon Laboratory, 1999.
3. M. Harrison, W. Ruzzo, J. Ullman. *Protection in operating systems.* Communications of the ACM, 19(8):461–470, August 1976.
4. T. Jurdziński. *On Complexity of Grammars Related to the Safety Problem.* ICALP 2006, LNCS 4052, 432–443. Full version to appear in Theoretical Computer Science.
5. T. Jurdziński, K. Loryś. *Leftist Grammars and the Chomsky Hierarchy.* Theory of Computing Systems, 41(2):233-256, 2007.
6. R. Motwani, R. Panigrahy, V. A. Saraswat, S. Venkatasubramanian. *On the decidability of accessibility problems (extended abstract).* STOC 2000, 306–315.
7. V. Saraswat. *The Matrix Design.* Technical report, AT&T Laboratory, April 1997.
8. V. Saraswat, R. Jagadeesan. *Static support for capability-based programming in Java.* Manuscript.
9. Ph. Schnoebelen. *Verifying Lossy Channel Systems has Nonprimitive Recursive Complexity.* Information Processing Letters 83(5), 251-261, 2002.