# ON THE NONTERMINAL COMPLEXITY OF LEFT RANDOM CONTEXT EOL GRAMMARS 

Petr Zemek<br>Doctoral Degree Programme (1), FIT BUT<br>E-mail: xzemek02@stud.fit.vutbr.cz

Supervised by: Alexander Meduna<br>E-mail: meduna@fit.vutbr.cz


#### Abstract

The present paper studies the nonterminal complexity of left random context EOL grammars. More specifically, it proves that every recursively enumerable language can be generated by a left random context EOL grammar with nine nonterminals. In the conclusion, some open problems related to the achieved result are stated.


Keywords: Formal languages, left random context EOL grammars, nonterminal complexity

## 1 INTRODUCTION

A left random context EOL grammar, introduced in $[7]^{1}$, is a variant of an E0L grammar regulated by context conditions (see Chapter 8 of [2]). Basically, it is an E0L grammar where every rewriting rule is equipped with two finite sets of nonterminals. One set contains permitting symbols while the other has forbidding symbols. A rule like this can rewrite a symbol provided that each of its permitting symbols occurs to the left of the rewritten symbol in the current sentential form while each of its forbidding symbols is absent therein.

As demonstrated by several recent studies, such as $[1,3,5,6,8]$, the study of descriptional complexity of formal models represents a modern trend in formal language theory. One of the studied parameter is the number of nonterminals needed to sustain the power of a grammar. We follow this trend by showing that every recursively enumerable language can be generated by a left random context EOL grammar with nine nonterminals.

To demonstrate that the number of nonterminals of a computationally complete grammar can be bounded, several techniques have been used [1]. In [5] and [8], the authors simulate a phrase-structure grammar in the Geffert normal form (see [4]), where the number of nonterminals is bounded by a very small constant. In [3], the number of nonterminals of graph controlled, programmed, and matrix grammars to generate any recursively enumerable language is reduced by a simulation of a register machine, which is a Turing machine with integers stored in counters rather than with words written on tapes. A similar variant of a Turing machine, called two-counter machine, is used in [1] to prove that every recursively enumerable language can be generated by a scattered context grammar with only two nonterminals. We use the technique based on the Geffert normal form.

The paper is organized as follows. First, Section 2 gives all the necessary terminology. Then, Section 3 proves that every recursively enumerable language can be generated by a left random context EOL grammar with nine nonterminals. In the conclusion of this paper, Section 4 states two open problems related to the achieved result.

[^0]
## 2 PRELIMINARIES

We assume that the reader is familiar with formal language theory (see [9]). For a set, $Q, \operatorname{card}(Q)$ denotes the cardinality of $Q$, and $2^{Q}$ denotes the power set of $Q$. For an alphabet (finite non-empty set), $V, V^{*}$ represents the free monoid generated by $V$ under the operation of concatenation. The unit of $V^{*}$ is denoted by $\varepsilon$. Define $V^{+}=V^{*}-\{\varepsilon\}$. For a string, $x \in V^{*},|x|$ denotes the length of $x$, and $\operatorname{alph}(x)$ denotes the set of symbols occurring in $x$.

A phrase-structure grammar is a quadruple, $G=(N, T, P, S)$, where $N$ is an alphabet of nonterminals, $T$ is an alphabet of terminals, $N \cap T=\emptyset, S \in N$ is the start symbol, and $P \subseteq(N \cup T)^{*} N(N \cup T)^{*} \times$ $(N \cup T)^{*}$ is a finite relation, called the set of rules. Each $(x, y) \in P$ is written as $x \rightarrow y$. Set $V=N \cup T$. The relation of a direct derivation, symbolically denoted by $\Rightarrow$, is defined as follows: if $u, v \in V^{*}$, and $x \rightarrow y \in P$, then $u x v \Rightarrow u y v$ in $G$. Let $\Rightarrow^{n}$ and $\Rightarrow^{*}$ denote the $n$th power of $\Rightarrow$, for some $n \geq 0$, and the reflexive-transitive closure of $\Rightarrow$, respectively. The language of $G$ is denoted by $L(G)$ and defined as $L(G)=\left\{w \in T^{*} \mid S \Rightarrow^{*} w\right\}$.

Let $G$ be a phrase-structure grammar. $G$ is in the Geffert normal form (see [4]) if it is of the form $G=(\{S, A, B, C\}, T, P \cup\{A B C \rightarrow \varepsilon\}, S)$, where $P$ contains only rules of the form (i) $S \rightarrow u S a$, (ii) $S \rightarrow u S v$, (iii) $S \rightarrow u v$, where $u \in\{A, A B\}^{*}, v \in\{B C, C\}^{*}$, and $a \in T$.

Lemma 1 (see [4]). Let $K$ be a recursively enumerable language. Then, there is a phrase-structure grammar in the Geffert normal form, $G$, such that $L(G)=K$. In addition, every successful derivation in $G$ is of the form $S \Rightarrow^{*} w_{1} w_{2} w_{3}$ by rules from $P$, where $w_{1} \in\{A, A B\}^{*}, w_{2} \in\{B C, C\}^{*}, w_{3} \in T^{*}$, and $w_{1} w_{2} w_{3} \Rightarrow^{*} w_{3}$ is derived by $A B C \rightarrow \varepsilon$.

A left random context EOL grammar (see [7]) is a quadruple, $G=(V, T, P, w)$, where $V$ is the total alphabet, $T \subseteq V$ is an alphabet of terminals, $N=V-T$ is an alphabet of nonterminals, $w \in V^{+}$is the start string, and $P \subseteq V \times V^{*} \times 2^{N} \times 2^{N}$ is finite. By analogy with phrase-structure grammars, elements of $P$ are called rules and instead of $(X, y, U, W) \in P$, we write $\lfloor X \rightarrow y, U, W\rfloor$. The relation of a direct derivation, symbolically denoted by $\Rightarrow$, is defined as follows: if $u=X_{1} X_{2} \cdots X_{k}, v=y_{1} y_{2} \cdots y_{k}$, $\left\lfloor X_{i} \rightarrow y_{i}, U_{i}, W_{i}\right\rfloor \in P, U_{i} \subseteq \operatorname{alph}\left(X_{1} X_{2} \cdots X_{i-1}\right)$, and alph $\left(X_{1} X_{2} \cdots X_{i-1}\right) \cap W_{i}=0$, for all $i, 1 \leq i \leq k$, for some $k \geq 1$, then $u \Rightarrow v$ in $G$. For $\lfloor X \rightarrow y, U, W\rfloor \in P, U$ and $W$ are called the left permitting context and the left forbidding context, respectively. Let $\Rightarrow^{n}$ and $\Rightarrow^{*}$ denote the $n$th power of $\Rightarrow$, for some $n \geq 0$, and the reflexive-transitive closure of $\Rightarrow$, respectively. The language of $G$ is denoted by $L(G)$ and defined as $L(G)=\left\{x \in T^{*} \mid w \Rightarrow^{*} x\right\}$.

## 3 MAIN RESULT

In this section, we prove that every recursively enumerable language can be generated by a left random context EOL grammar with nine nonterminals.

Theorem 1. Let $K$ be a recursively enumerable language. Then, there is a left random context E0L grammar, $H=(V, T, P, w)$, such that $L(H)=K$ and $\operatorname{card}(N)=9$.

Proof. Let $K$ be a recursively enumerable language. Then, by Lemma 1, there is a phrase-structure grammar in the Geffert normal form, $G=(\{S, A, B, C\}, T, P \cup\{A B C \rightarrow \varepsilon\}, S)$, such that $L(G)=K$. We next construct a left random context E0L grammar, $H$, such that $L(H)=L(G)$. Set $N=\{S, A$, $B, C\}, V=\{S, A, B, C\} \cup T$, and $N^{\prime}=N \cup\{\bar{A}, \hat{A}, \bar{B}, \hat{B}, \#\}$. Define $H=\left(V^{\prime}, T, P^{\prime}, S \#\right)$, where initially $V^{\prime}=N^{\prime} \cup T$ and $P^{\prime}=\{\lfloor a \rightarrow a, \emptyset, \emptyset\rfloor \mid a \in T\} \cup\{\lfloor X \rightarrow X, \emptyset,\{\#\}\rfloor \mid X \in\{A, B, C\}\}$. To finish the construction, apply the following six steps:
(1) add $\lfloor \# \rightarrow \#, \emptyset,\{\bar{A}, \bar{B}, \hat{A}, \hat{B}\}\rfloor,\lfloor \# \rightarrow \#,\{\hat{A}, \hat{B}, C\},\{S\}\}\rfloor$, and $\left\lfloor \# \rightarrow \varepsilon, \emptyset, N^{\prime}-\{\#\}\right\rfloor$ to $P^{\prime}$;
(2) for each $S \rightarrow u S a \in P$, where $u \in\{A, A B\}^{*}$ and $a \in T$, add $\lfloor S \rightarrow u S \# a, \emptyset,\{\bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor$ to $P^{\prime}$;
(3) for each $S \rightarrow u S v \in P$, where $u \in\{A, A B\}^{*}$ and $v \in\{B C, C\}^{*}$, $\operatorname{add}\lfloor S \rightarrow u S v, \emptyset,\{\bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor$ to $P^{\prime} ;$
(4) for each $S \rightarrow u v \in P$, where $u \in\{A, A B\}^{*}$ and $v \in\{B C, C\}^{*}$, add $\lfloor S \rightarrow u v, \emptyset,\{\bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor$ to $P^{\prime} ;$
(5) add $\lfloor A \rightarrow \bar{A}, \emptyset,\{S, \bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor,\lfloor B \rightarrow \bar{B}, \emptyset,\{S, \bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor,\lfloor A \rightarrow \hat{A}, \emptyset,\{S, \bar{A}, \bar{B}, \hat{A}, \hat{B}$, $\#\}\rfloor$, and $\lfloor B \rightarrow \hat{B}, \emptyset,\{S, \bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor$ to $P^{\prime}$;
(6) add $\lfloor\bar{A} \rightarrow A, \emptyset,\{S, A, B, C, \hat{A}, \hat{B}, \#\}\rfloor,\lfloor\bar{B} \rightarrow B, \emptyset,\{S, A, B, C, \hat{A}, \hat{B}, \#\}\rfloor,\lfloor\hat{A} \rightarrow \varepsilon, \emptyset,\{S, A, B, C$, $\hat{A}, \hat{B}, \#\}\rfloor,\lfloor\hat{B} \rightarrow \varepsilon,\{\hat{A}\},\{S, A, B, C, \hat{B}, \#\}\rfloor,\lfloor C \rightarrow \varepsilon,\{\hat{A}, \hat{B}\},\{S, A, B, C, \#\}\rfloor$ to $P^{\prime}$.
$H$ simulates derivations of the form specified in Lemma 1. Rules in $P$ are simulated by rules from (2) through (4). $A B C \rightarrow \varepsilon$ is simulated in two steps. First, rules introduced in (5) are used to prepare the application of rules from (6). Then, the latter rules perform the actual erasure of $A B C$. For example, $A A B C B C \# a \# \Rightarrow \bar{A} \hat{A} \hat{B} C B C \# a \# \Rightarrow A B C \# a \#$ in $H$.

The role of \# is twofold. First, it ensures that every sentential form of $H$ is of the form $w_{1} w_{2}$, where $w_{1} \in\left(N^{\prime}-\{\#\}\right)^{*}$ and $w_{2} \in(T \cup\{\#\})^{*}$. Since left permitting and left forbidding contexts cannot contain terminals, a mixture of symbols from $T$ and $N$ in $H$ could lead to false sentences. Second, if any of $\bar{A}, \bar{B}, \hat{A}$, or $\hat{B}$ are present, $A B C \rightarrow \varepsilon$ has to be simulated. Therefore, it prevents derivations of the form $A a \Rightarrow \hat{A} a \Rightarrow a$ in $H$ (notice that the start string of $H$ is $S \#$ ). Furthermore, we exploit the fact that in every derivation step of $H$, all symbols have to be rewritten. Consequently, if rules from (5) are used improperly, the derivation is blocked, and so no partial erasures are possible.
Observe that every sentential form of $G$ and $H$ contains at most one occurrence of $S$. In a derivation step of $H$, only a single rule from $P \cup\{A B C \rightarrow \varepsilon\}$ can be simulated at once. $A B C \rightarrow \varepsilon$ can be simulated only if $S$ is not present. \#'s can be eliminated in a single step by an application of rules from (1); however, only if there are no nonterminals present in the current sentential form. Based on these observations and on Lemma 1, we see that every successful derivation in $H$ is of the form $S \# \Rightarrow^{*} w_{1} w_{2} \# a_{1} \# a_{2} \cdots \# a_{n} \# \Rightarrow^{*} \# a_{1} \# a_{2} \cdots \# a_{n} \# \Rightarrow a_{1} a_{2} \cdots a_{n}$, where $w_{1} \in\{A, A B\}^{*}, w_{2} \in\{B C, C\}^{*}$, $a_{i} \in T$, for all $i, 1 \leq i \leq n$, for some $n \geq 0$.

Due to space requirements, we omit some details in what follows. The reader can easily fill them in. To establish $L(H)=L(G)$, we prove two claims. The first claim shows how derivations of $G$ are simulated by $H$. It is then used to prove $L(G) \subseteq L(H)$. Define the homomorphism $\varphi$ from $V^{*}$ to $V^{*}$ as $\varphi(X)=X$, for all $X \in N$, and $\varphi(a)=\# a$, for all $a \in T$.

Claim 1. If $S \Rightarrow^{n} x \Rightarrow^{*} z$ in $G$, for some $n \geq 0$, where $x \in V^{*}, z \in T^{*}$, then $S \# \Rightarrow^{*} \varphi(x) \#$ in $H$.

Proof. This claim is established by induction on $n$, where $n \geq 0$. Basis: For $n=0$, the claim clearly holds. Induction Hypothesis: Suppose that the claim holds for all derivations of length $l$ or less, where $l \leq n$, for some $n \geq 0$. Induction Step: Consider any derivation of the form $S \Rightarrow^{n+1} w \Rightarrow^{*} z$ in $G$, where $w \in V^{*}, z \in T^{*}$. Since $n+1 \geq 1$, this derivation can be expressed as $S \Rightarrow^{n} x \Rightarrow w \Rightarrow^{*} z$, for some $x \in V^{*}$. Without any loss of generality, we may assume that $x=x_{1} x_{2} x_{3} x_{4}$, where $x_{1} \in\{A$, $A B\}^{*}, x_{2} \in\{S, \varepsilon\}, x_{3} \in\{B C, C\}^{*}$, and $x_{4} \in T^{*}$ (see Lemma 1 and [4]). Next, we consider all possible forms of $x \Rightarrow w$ in $G$, covered by the following four cases-(i) through (iv).
(i) (Application of $S \rightarrow u S a \in P$.) Let $x=x_{1} S x_{3} x_{4}, w=x_{1} u S a x_{3} x_{4}$, and $S \rightarrow u S a \in P$, where $u \in\{A, A B\}^{*}$ and $a \in T$. Then, by the induction hypothesis, $S \# \Rightarrow^{*} \varphi\left(x_{1} S x_{3} x_{4}\right) \#$ in $H$. By (2), $\lfloor S \rightarrow u S \# a, \emptyset,\{\bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor \in P^{\prime}$. Since $\varphi\left(x_{1} S x_{3} x_{4}\right) \#=x_{1} S x_{3} \varphi\left(x_{4}\right) \#$ and $\operatorname{alph}\left(x_{1} S x_{3}\right) \cap\{\bar{A}$,
$\bar{B}, \hat{A}, \hat{B}, \#\}=\emptyset, x_{1} S x_{3} \varphi\left(x_{4}\right) \# \Rightarrow x_{1} u S \# a x_{3} \varphi\left(x_{4}\right) \#$ in $H$. As $\varphi\left(x_{1} u S a x_{3} x_{4}\right) \#=x_{1} u S \# a x_{3} \varphi\left(x_{4}\right) \#$, the induction step is completed for (i).
(ii) (Application of $S \rightarrow u S v \in P$.) Let $x=x_{1} S x_{3} x_{4}, w=x_{1} u S v x_{3} x_{4}$, and $S \rightarrow u S v \in P$, where $u \in\{A$, $A B\}^{*}$ and $v \in\{B C, C\}^{*}$. Proceed by analogy with (i) by using a rule from (3).
(iii) (Application of $S \rightarrow u v \in P$.) Let $x=x_{1} S x_{3} x_{4}, w=x_{1} u v x_{3} x_{4}$, and $S \rightarrow u v \in P$, where $u \in\{A$, $A B\}^{*}$ and $v \in\{B C, C\}^{*}$. Proceed by analogy with (i) by using a rule from (4).
(iv) (Application of $A B C \rightarrow \varepsilon$.) Let $x=x_{1}^{\prime} A B C x_{3}^{\prime} x_{4}, w=x_{1}^{\prime} x_{3}^{\prime} x_{4}$, where $x_{1} x_{2}=x_{1}^{\prime} A B C x_{2}^{\prime}$, so $x \Rightarrow w$ in $G$ by $A B C \rightarrow \varepsilon$. Then, by the induction hypothesis, $S \# \Rightarrow^{*} \varphi\left(x_{1}^{\prime} A B C x_{3}^{\prime} x_{4}\right) \#$ in $H$. Let $x_{1}^{\prime}=$ $X_{1} X_{2} \cdots X_{k}$, where $k=\left|x_{1}^{\prime}\right|$. Since $\varphi\left(x_{1}^{\prime} A B C x_{3}^{\prime} x_{4}\right) \#=x_{1}^{\prime} A B C x_{3}^{\prime} \varphi\left(x_{4}\right) \#$ and alph $\left(x_{1}^{\prime} A B C x_{3}^{\prime}\right) \cap\{\bar{A}$, $\bar{B}, \hat{A}, \hat{B}, \#\}=\emptyset, x_{1}^{\prime} A B C x_{3}^{\prime} \varphi\left(x_{4}\right) \# \Rightarrow \bar{x}_{1}^{\prime} \hat{A} \hat{B} C x_{3}^{\prime} \varphi\left(x_{4}\right) \#$ in $H$ by rules from (5), where $\bar{x}_{1}^{\prime}=$ $\bar{X}_{1} \bar{X}_{2} \cdots \bar{X}_{k}$. Since $\operatorname{alph}\left(\bar{x}_{1}^{\prime}\right) \cap\{S, A, B, C, \hat{A}, \hat{B}, \#\}=\emptyset, \hat{A} \in \operatorname{alph}\left(\bar{x}_{1}^{\prime} \hat{A}\right), \hat{A}, \hat{B} \in \operatorname{alph}\left(\bar{x}_{1}^{\prime} \hat{A} \hat{B}\right)$, and $\hat{A}, \hat{B}, C \in \operatorname{alph}\left(\bar{x}_{1}^{\prime} \hat{A} \hat{B} C x_{2}\right), \bar{x}_{1}^{\prime} \hat{A} \hat{B} C x_{3}^{\prime} \varphi\left(x_{4}\right) \# \Rightarrow x_{1}^{\prime} x_{3}^{\prime} \varphi\left(x_{4}\right) \#$ in $H$ by rules from (6). As $\varphi\left(x_{1}^{\prime} x_{3}^{\prime} x_{4}\right) \#=$ $x_{1}^{\prime} x_{3}^{\prime} \varphi\left(x_{4}\right) \#$, the induction step is completed for (iv).

Observe that these four cases cover all possible forms of $x \Rightarrow w$ in $G$, so the claim holds.
The second claim demonstrates how $G$ simulates derivations of $H$. It is then used to prove $L(H) \subseteq$ $L(G)$. Define the homomorphism $\tau$ from $V^{* *}$ to $V^{*}$ as $\tau(X)=X$, for all $X \in N, \tau(\bar{A})=\tau(\hat{A})=A$, $\tau(\bar{B})=\tau(\hat{B})=B, \tau(a)=a$, for all $a \in T$, and $\tau(\#)=\varepsilon$.
Claim 2. If $S \# \Rightarrow^{n} x \Rightarrow^{*} z$ in $H$, for some $n \geq 0$, where $x \in V^{\prime *}, z \in T^{*}$, then $S \Rightarrow^{*} \tau(x)$ in $G$.

Proof. This claim is established by induction on $n$, where $n \geq 0$. Basis: For $n=0$, the claim clearly holds. Induction Hypothesis: Suppose that the claim holds for all derivations of length $l$ or less, where $l \leq n$, for some $n \geq 0$. Induction Step: Consider any derivation of the form $S \# \Rightarrow^{n+1} w \Rightarrow^{*} z$ in $H$, where $w \in V^{*}, z \in T^{*}$. Since $n+1 \geq 1$, this derivation can be expressed as $S \# \Rightarrow^{n} x \Rightarrow w \Rightarrow^{*} z$, for some $x \in V^{*}$. By the induction hypothesis, $S \Rightarrow^{*} \tau(x)$ in $G$. Next, we consider all possible forms of $x \Rightarrow w$ in $H$, covered by the following five cases-(i) through (v).
(i) Let $x=x_{1} S x_{2}$ and $w=x_{1}^{\prime} u S \# a x_{2}^{\prime}$, where $x_{1}, x_{1}^{\prime}, x_{2}, x_{2}^{\prime} \in V^{\prime *}$, such that $x_{1} S x_{2} \Rightarrow x_{1}^{\prime} u S \# a x_{2}^{\prime}$ in $H$ by $\lfloor S \rightarrow u S \# a, \emptyset,\{\bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor$-introduced in (2) from $S \rightarrow u S a \in P$, where $u \in\{A, A B\}^{*}$, $a \in T$-and by the rules introduced in the initialization part, in (1), and in (5). Since $\tau\left(x_{1} S x_{2}\right)=$ $\tau\left(x_{1}\right) S \tau\left(x_{2}\right), \tau\left(x_{1}\right) S \tau\left(x_{2}\right) \Rightarrow \tau\left(x_{1}\right) u \operatorname{Sa} \tau\left(x_{2}\right)$ in $G$. As $\tau\left(x_{1}\right) u \operatorname{Sa} \tau\left(x_{2}\right)=\tau\left(x_{1}^{\prime} u S \# a x_{2}^{\prime}\right)$, the induction step is completed for (i).
(ii) Let $x=x_{1} S x_{2}$ and $w=x_{1}^{\prime} u S v x_{2}^{\prime}$, where $x_{1}, x_{1}^{\prime}, x_{2}, x_{2}^{\prime} \in V^{\prime *}$, such that $x_{1} S x_{2} \Rightarrow x_{1}^{\prime} u S v x_{2}^{\prime}$ in $H$ by $\lfloor S \rightarrow u S v, \emptyset,\{\bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor$ —introduced in (3) from $S \rightarrow u S v \in P$, where $u \in\{A, A B\}^{*}$, $v \in\{B C, C\}^{*}$-and by the rules introduced in the initialization part and in (1) and (5). Proceed by analogy with (i).
(iii) Let $x=x_{1} S x_{2}$ and $w=x_{1}^{\prime} u v x_{2}^{\prime}$, where $x_{1}, x_{1}^{\prime}, x_{2}, x_{2}^{\prime} \in V^{\prime *}$, such that $x_{1} S x_{2} \Rightarrow x_{1}^{\prime} u v x_{2}^{\prime}$ in $H$ by $\lfloor S \rightarrow u v, \emptyset,\{\bar{A}, \bar{B}, \hat{A}, \hat{B}, \#\}\rfloor$-introduced in (4) from $S \rightarrow u v \in P$, where $u \in\{A, A B\}^{*}, v \in\{B C$, $C\}^{*}$-and by the rules introduced in the initialization part and in (1) and (5). Proceed by analogy with (i).
(iv) Let $x=x_{1} \hat{A} x_{2} \hat{B} x_{3} C x_{4}$ and $w=x_{1}^{\prime} x_{2}^{\prime} x_{3}^{\prime} x_{4}^{\prime}$, where $x_{1}, x_{1}^{\prime}, x_{2}, x_{2}^{\prime}, x_{3}, x_{3}^{\prime}, x_{4}, x_{4}^{\prime} \in V^{\prime *}$, such that $x_{1} \hat{A} x_{2} \hat{B} x_{3} C x_{4} \Rightarrow x_{1}^{\prime} x_{2}^{\prime} x_{3}^{\prime} x_{4}^{\prime}$ in $H$ by rules introduced in (6), in the initialization part, and in (1) and (5). Observe that $x_{2}=x_{2}^{\prime}=x_{3}=x_{3}^{\prime}=\varepsilon$, $\operatorname{alph}\left(x_{1} x_{3}\right) \cap\{\hat{A}, \hat{B}\}=\emptyset$, and the only occurrence of $C$ that is erased is the one right next to $\bar{B}$; otherwise, this derivation in $H$ is not possible.

Therefore, $x=x_{1} \hat{A} \hat{B} C x_{4}$ and $w=x_{1}^{\prime} x_{4}^{\prime}$. Since $\tau\left(x_{1} \hat{A} \hat{B} C x_{4}\right)=\tau\left(x_{1}\right) A B C \tau\left(x_{4}\right), \tau\left(x_{1}\right) A B C \tau\left(x_{4}\right) \Rightarrow$ $\tau\left(x_{1}\right) \tau\left(x_{4}\right)$ by $A B C \rightarrow \varepsilon$ in $G$. As $\tau\left(x_{1}\right) \tau\left(x_{4}\right)=\tau\left(x_{1}^{\prime} x_{4}^{\prime}\right)$, the induction step is completed for (iv).
(v) Let $x \Rightarrow w$ in $H$ only by rules from the initialization part, from (1), from (5), and by the first two rules from (6). As $\tau(x)=\tau(w)$, the induction step is completed for (v).

Observe that these five cases cover all possible forms of $x \Rightarrow w$ in $H$, so the claim holds.

We now prove that $L(H)=L(G)$. Consider a special case of Claim 1 when $x \in T^{*}$. Then, $S \# \Rightarrow^{*} \varphi(x) \#$ in $H$. By (1), $\left\lfloor \# \rightarrow \varepsilon, \emptyset, N^{\prime}-\{\#\}\right\rfloor \in P^{\prime}$. Since $\operatorname{alph}(\varphi(x) \#) \cap\left(N^{\prime}-\{\#\}\right)=\emptyset, \varphi(x) \# \Rightarrow x$ in $H$. Hence, $L(G) \subseteq L(H)$. Consider a special case of Claim 2 when $x \in T^{*}$. Then, $S \Rightarrow^{*} x$ in $G$. Hence, $L(H) \subseteq L(G)$. As $\operatorname{card}\left(N^{\prime}\right)=9$, the theorem holds.

## 4 CONCLUSION

The present paper demonstrated that every recursively enumerable language can be generated by a left random context E0L grammar with nine nonterminals. A question to be investigated is whether this bound is, in fact, optimal. Furthermore, in [7], it is proved that left random context E0L grammars without erasing rules-that is, without any rules of the form $X \rightarrow \varepsilon$-generate precisely the family of context-sensitive languages. Can we also bound the number of nonterminals in terms of this nonerasing version? We suggest these two open problems for further research.

## ACKNOWLEDGEMENT

This work was supported by the research plan MSM0021630528.

## REFERENCES

[1] E. Csuhaj-Varjú and G. Vaszil. Scattered context grammars generate any recursively enumerable language with two nonterminals. Information Processing Letters, 110:902-907, 2010.
[2] J. Dassow and G. Păun. Regulated Rewriting in Formal Language Theory. Springer, 1989.
[3] H. Fernau, R. Freund, M. Oswald, and K. Reinhardt. Refining the nonterminal complexity of graph-controlled, programmed, and matrix grammars. Journal of Automata, Languages and Combinatorics, 12(1-2):117-138, 2007.
[4] V. Geffert. Normal forms for phrase-structure grammars. Theoretical Informatics and Applications, 25(5):473-496, 1991.
[5] T. Masopust. On the descriptional complexity of scattered context grammars. Theoretical Computer Science, 410(1):108-112, 2009.
[6] T. Masopust and A. Meduna. On descriptional complexity of partially parallel grammars. Fundamenta Informaticae, 87(3):407-415, 2008.
[7] A. Meduna and P. Zemek. Left random context ET0L grammars. manuscript.
[8] F. Okubo. A note on the descriptional complexity of semi-conditional grammars. Information Processing Letters, 110(1):36-40, 2009.
[9] G. Rozenberg and A. Salomaa, editors. Handbook of Formal Languages, Volumes 1 through 3. Springer, 1997.


[^0]:    ${ }^{1}$ Let us note that [7] represents a not yet published work containing ongoing research.

