Simulation of Scattered Context Grammars and Phrase-Structured Grammars by Symbiotic E0L Grammars

Programming language theory

Tomáš Kopeček

January 2005

Abstract: This paper contains more examples to formerly introduced concept of formal language equivalency. That is, for two models, there is a substitution by which we change each string of every yield sequence in one model so that sequence ofs string resulting from this change represents a yield sequence in the other equvalent model, these two models closely simulates each other; otherwise they do not. In this paper are shown two cases of such simulations.

Contents

1 Introduction

In the [1] was introduced quite new method of compraing two grammatical systems. Before this paper there was almost vague comparations of grammars limited by similarity of generated languages. This new approach comes with comparing not only generated languages but also similarity of generating process.

Because we have many different transformations from one type of grammar to another in the theory of formal languages, we sometimes want to describe similarity of such converted grammars. On the second hand, we need to examine this similarity in the practice. For example we try to find some usable representation of some grammar for use in a compiling system. We can do some transformations but we still want to achieve same result in new grammar with almost same number of derivation steps and so on.

So, the concepts of *m-close simulation* and some others were introduced in [1].

In the section 2 are recalled some well-known notions of the formal language theory. Section 3 introduces new conversion from scattered context grammars to symbiotic E0L grammrs. Similar conversion from phrase-structured grammars is described in Section 4. Next section deals with description of derivation simulations from previous two sections. Here are repeated some needed definitions of concepts of derivation similarity and proved two theorems about previous conversions. Section 6 includes proved results as a whole.

2 Preliminaries

This paper assumes that the reader is familiar with the language theory (see [2], [4], [6]).

Let V be an alphabet. V^* denotes the free monoid generated by V under the operation of concatenation. Let ε be the unit of V^* and $V^+ = V^* - \{\varepsilon\}$. Given a word, $w \in V^*$, $|w|$ represents the length of w and $alph(w)$ denotes the set of all symbols occuring in w. Moreover, $sub(w)$ denotes the set of all subwords of w. Let R be a binary relation on a set W. Instead of $u \in R(v)$, where $u, v \in W$, we write vRu in this paper.

A *scattered context grammar* is an ordered quadruple $G = (V, T, P, S)$, where V, T, and S are the total alphabet of G, the set of terminals $T \subseteq V$, and the axiom $S \in V - T$, respectively. P is a finite set of productions of the form $(A_1, \dots, A_n) \rightarrow (x_1, \dots, x_n)$, for some $n \geq 1$, where $A_i \in V - T$ and $x_i \in V^*$ form $1 \leq i \leq n$. If $p \in P$ is of the form $(A_1, A_2, \dots, A_n) \rightarrow (x_1, x_2, \dots, x_n), u = u_1A_1u_2A_2 \dots u_nA_nu_{n+1}, v = u_1x_1u_2x_2 \dots u_nx_nu_{n+1}$, where $u_i \in V^*$, for $i = 1, 2, \dots, n$, then u directly derives v in G according to p, denoted by $u \Rightarrow_G v[p]$ or, simply $u \Rightarrow v$. In a standard manner, we extend $\Rightarrow_G$ to $\Rightarrow_G^n$, where $n \geq 0$, and based on $\Rightarrow_G^n$, we define $\Rightarrow_G^*$, which is transitive and reflexive closure of $\Rightarrow$. Let $S \Rightarrow_G^* x$ is called a successful derivation. The language of G, $L(G)$, is defined as $L(G) = \{x : S \Rightarrow_G^* x, x \in T^*\}$. For any $p \in P$ of the form $(A_1, A_2, \dots, A_n) \rightarrow (x_1, x_2, \dots, x_n), left(p)$ means string $A_1A_2 \dots A_n$ and $right(p)$ string $x_1x_2 \dots x_n$.

A *phrase-structured grammar* is and ordered quadruple $G = (V, T, P, S)$, where V, T, and S are the total alphabet of G, the set of terminals $T \subseteq V$, and the axiom $S \in V - T$, respectively. P is a finite set of productions of the form $x \rightarrow y$, where $x \in V^+$ and

$y \in V^*$. If $p \in P$ is of the form $x \to y, u = u_1 x u_2, v = u_1 y u_2$, where $u, v \in V^*$, then u directly derives v in G according to p, denoted by $u \Rightarrow_G v[p]$ or, simply $u \Rightarrow v$. In a standard manner, we extend $\Rightarrow_G$ to $\Rightarrow_G^n$, where $n \geq 0$, and based on $\Rightarrow_G^n$, we define $\Rightarrow_G^*$, which is transitive and reflexive closure of $\Rightarrow$. Let $S \Rightarrow_G^* x$ is called a successful derivation. The language of G, $L(G)$, is defined as $L(G) = \{x : S \Rightarrow_G^* x, x \in T^*\}$. For any $p \in P$ of the form $x \to y, left(p)$ means string x and $right(p)$ string y.

A *symbiotic E0L grammar* (see [3]) is a quadruple $G = (W, T, P, S)$, where W, T, and S are the set of generators $W \subseteq (V \cup V^2)$, the set of terminals $T \subseteq V$, and the axiom $S \in V - T$, respectively. P is a finite set of productions of the form $A \to x$, $A \in V$, $x \in V^*$. The direct derivation relation is defined in the following way: let $x, y \in W^*$ such that $x = a_1 a_2 \dots a_n$, $a_i \in V$, $y = y_1 y_2 \dots y_n$, $y_i \in V^*$, and productions $a_i \to y_i \in P$ for all $i = 1, \dots, n$. Then, x directly derives y, $x \Rightarrow_G y$ in symbols. The language of G is $L(G) = \{w \in T^* : S \Rightarrow_G^* w\}$.

3 Simulation of Scattered Context Grammars

Construction 1.

Input: A scattered context grammar, $G = (V, T, P, S)$

Output: A symbiotic E0L grammars, G'

Algorithm: At first, we introduce a new alphabet, $V' = V \cup \{@, \#, S'\} \cup V'' \cup \widetilde{T}, \widetilde{T} = \{\widetilde{a} : a \in T\}, V'' = \{\langle i, j\rangle : 0 < i \leq Card(P), 0 \leq j \leq k\}$. Let τ be a homomorphism from T to $\widetilde{T}$ such that $\tau(a) = \widetilde{a}$ for all $a \in T$. Define a language W, over V' as $W = V \cup \{@, \#, S'\} \cup \widetilde{T} \cup \{\langle i, j\rangle \langle i, j\rangle : 0 < i \leq Card(P), 0 \leq j \leq k\}$. Then, construct a symbiotic E0L grammar $G' = (W, T, P', S')$, where the set of productions is defined in the following way:

1. add $S' \to @S\#$ to P';

2. for every production $n: (A_1, A_2, \dots, A_k) \to (x_1, x_2, \dots, x_k) \in P$ add these rules to P' (where n is a label, $0 <\leq Card(P)$:

$$\begin{array}{lcl} A_1 & \to & \langle n, 0\rangle \, \tau(x_1) \, \langle n, 1\rangle \\ A_2 & \to & \langle n, 1\rangle \, \tau(x_2) \, \langle n, 2\rangle \\ & \vdots & \\ A_k & \to & \langle n, k-1\rangle \, \tau(x_k) \, \langle n, k\rangle \end{array}$$

3. add $@ \to @ \langle i, 0\rangle, 0 < i \leq Card(P)$ to P';

4. add $\# \to \langle i, k\rangle \#$, to P' for each production $i: (A_1, A_2, \dots, A_k) \to (x_1, x_2, \dots, x_k) \in P$;

5. add $@ \to \varepsilon$;

6. add $\# \to \varepsilon$;

7. for each $A \in V \cup \widetilde{T}$ add productions of this form to P': $A \rightarrow \langle i, j \rangle A \langle i, j \rangle, 0 < i \leq Card(P), 0 \leq j \leq k$;

8. add these productions to P': $\langle i, j \rangle \rightarrow \varepsilon, 0 < i \leq Card(P), 0 \leq j \leq k$;

9. add production $\widetilde{a} \rightarrow a$ for each $a \in T$ to P'.

Theorem 1. *Let $G = (V, T, P, S)$ be a scattered context grammar. Let G' be a symbiotic E0L grammar constructed by Construction 1 with G as its input. Then, $L(G) = L(G')$.*

Proof. Let ω be a homomorphism from V' to $V' - V''$ defined as $\omega(a) = \varepsilon$ for all $a \in V''$, and $\omega(a) = a$, for all $a \in V' - V''$.

Claim 1. *For every $w \in W^*$ holds,*

1. *$S' \Rightarrow_G^+ w$ if and only if $@S\# \Rightarrow_G^* w$;*

2. *$S' \Rightarrow_G^+ w$ implies $S' \notin sub(w)$.*

Proof. By the definition of P', it is easy to see that the very first derivation step always rewrites S' to $@S\#$. Moreover, no productions generate S'; thus, S' appears in no sentential form derived from S'. □

Claim 2. *For all $u, v \in W^+, S' \notin sub(uv), u \Rightarrow_{G'} v$ if and only if $\omega(u) \Rightarrow_{G'} v$.*

Proof. Examine the definition of P'. Clearly, all occurrences of symbols from V'' are always erased during $u \Rightarrow_{G'} v$, so they play no role in the generation of v. By the definition of W and ω, $\omega(u) \in W^*$; therefore $\omega(u) \Rightarrow_{G'}$ is a valid derivation in G'.

Note that this property of derivations in G' allows us to ignore symbols of the form $\langle i, j \rangle$ occuring in left-hand sides of derivation steps. □

Claim 3. *Let $@y\# \Rightarrow_{G'} @x\#$, where $y = a_1 a_2 \dots a_n$ for some $a_i \in V, x \in W^*, n \geq 0$. Then, $@x\# = @ \langle i, 0 \rangle \langle i, 0 \rangle x_1 \langle i, 1 \rangle \langle i, 1 \rangle \dots \langle i, 1 \rangle \langle i, 1 \rangle x_m \langle i, 2 \rangle \langle i, 2 \rangle x_{m+1} \langle i, 2 \rangle \langle i, 2 \rangle \dots \langle i, k-1 \rangle x_n \langle i, k \rangle \langle i, k \rangle x_{n+1} \langle i, k \rangle \dots x_m \langle i, k \rangle \langle i, k \rangle \#$, where $x_j \in V^*$ for all $j = 1, 2, \dots,$ m and some i.*

Proof. Since x is surrounded by @ and # in $@x\#$, G' surely rewrites $@x\#$ in such way, that @ is rewritten to some $@ \langle i, 0 \rangle$ and # to $\langle j, k \rangle, 0 \leq i, j \leq Card(P)$. Every A_l can be rewritten either to $\langle i, j \rangle x_l \langle i, j \rangle$ or (if such production exists) to $\langle i, j-1 \rangle x_l \langle i, j \rangle$, where $0 < i \leq Card(P), 0 \leq j \leq k, x_i \in V^*$. Thus, $@x\# = @ \langle i, 0 \rangle \alpha_1 z_1 \beta_1 \alpha_2 z_2 \beta_2 \dots \alpha_n z_n \beta_n$ $\langle j, k \rangle \#$ with $\alpha_l = \langle i, j \rangle, z_l = x_l, \beta_l = \langle i, j \rangle$, or $\alpha_l = \langle i, j-1 \rangle, z_l = x_l, \beta_l = \langle i, j \rangle$, for all $l = 1, 2, \dots n$. However, $@x\#$ must be a string over W. Inspect the definition of W to see that $@x\# \in W^*$ if and only if $\alpha_1 = \langle i, 0 \rangle$ and $\beta_n \langle i, k \rangle$. Then, β_1 could be only $\langle i, 0 \rangle$ or $\langle i, 1 \rangle$. In same way α_n could be only $\langle i, k \rangle$ or $\langle i, k-1 \rangle$. We can simply show, that we can get only sentential form $@x\# = @ \langle i, 0 \rangle \langle i, 0 \rangle x_1 \langle i, 1 \rangle \langle i, 1 \rangle \dots \langle i, 1 \rangle \langle i, 1 \rangle x_m \langle i, 2 \rangle$ $\langle i, 2 \rangle x_{m+1} \langle i, 2 \rangle \langle i, 2 \rangle \dots \langle i, k-1 \rangle x_n \langle i, k \rangle \langle i, k \rangle x_{n+1} \langle i, k \rangle \dots x_m \langle i, k \rangle \langle i, k \rangle \#$.

□

Claim 4. *Let* $@y\# \Rightarrow_{G'} x$*, where* $y = a_1 a_2 \dots a_n$ *and* $\{@, \#\} \cap sub(x) = \emptyset$ *for some* $a_i \in V, x \in W^*, n \geq 0$*. Then,* $x = \langle i,0\rangle \langle i,0\rangle \tau(t_1) \langle i,0\rangle \dots \langle i,k\rangle \tau(t_n) \langle i,k\rangle \langle i,k\rangle$*, where* $t_i \in T^*$ *for all* $i = 1, 2, \dots, n$*.*

Proof. Prove this claim by analogy with the proof of Claim 3. □

The following claim shows that Claims 3 and 4 cover all possible ways of rewriting of a string having the form $@y\#, y \in V^*$, in G'.

Claim 5. *Let* $@y\# \Rightarrow_{G'} u, y \in V^*$*. Then, either* $u = @x\#, x \in W^*$*, or* $u \in W^*, \omega(u) \in T^*$*, and* $\{@, \#\} \cap sub(u) = \emptyset$*.*

Proof. Return to the proof of Claim 3. Suppose that @ is rewritten to $@\langle i,0\rangle$ and # is rewritten to ε. Then we can construct only strings of the form $z = @x\langle i,j\rangle y \langle i,k\rangle$, where $x \in W^*, y \in V^*$ and last symbol of y is from $V - V''$. It is clear, that $z \notin W^*$. Analogously, suppose that @ is rewritten to ε and # is rewritten to $\langle i,k\rangle \#$. As before, such a sentential form is out of W^*. □

Claim 6. *Let* $u \Rightarrow_{G'} v, u \in W^*, \{@, \#\} \cap sub(u) = \emptyset$*. Then* $v \in T^*$*.*

Proof. From the Claim 5 we see, that $\omega(u) \in T^*$. Then, we have to consider only productions with its left sides from $\widetilde{T}$, because it is the only possibility. Such productions are of the form $\widetilde{t} \rightarrow \langle i,j\rangle t \langle i,j\rangle$ or $\widetilde{t} \rightarrow t$, where $t \in T, 0 < i \leq Card(P), j \geq 1$. Then, string v could have one of the following forms:

1. $u = \langle i,j\rangle t \langle i,j\rangle y, t \in T, 0 < i \leq Card(P), 0 \leq j, y \in (V'' \cup T)^*$;
2. $u = x\langle i,j\rangle t \langle i,j\rangle y, x \in T^*, t \in T, y \in (V'' \cup T)^*$;
3. $u = t_1 t_2 \dots t_n, t_i \in T$.

It is easy to see, that only third form is the legal one. The others are out of W. □

Claim 7. *Every derivation in* G' *is a prefix of*

$$\begin{array}{rl} S' & \Rightarrow_{G'} @w_0\# \\ & \Rightarrow_{G'} @w_1\# \\ & \vdots \\ & \Rightarrow_{G'} @w_n\# \\ & \Rightarrow_{G'} u \\ & \Rightarrow_{G'} t \end{array}$$

where $w_0 = S$*,* $w_i \in W^*$*,* $\omega(u) = \tau(t)$*,* $t \in T^*$*,* $0 \leq i \leq n$*,* $n \geq 0$*.*

Proof. By the proof of Claim 1, S' is always rewritten to $@w_0\#$, where $w_0 = S$. Then, Claim 5 tells us that there are two possible forms of derivations rewriting $\omega(@w_i\#)$ and, hence, $@w_i\#$. First, G' can generate a sequence of n sentential forms that belong to $\{@\}W^*\{\#\}$, for some $n \geq 0$ (their form is described in Claim 3). Second, G' can rewrite $@w_n\#$ to $u \in W^*$, satisfying $\omega(u) \in \widetilde{T}^*$ (see Claim 4). By the Claim 6 the only form, to which could be rewritten u is t. Therefore, $u \Rightarrow_{G'} t$ such that $t \in T^*$ and $\omega(u) = \tau(t)$. After that, no other derivation step can be made from t because P' contains no production that rewrites terminals. □

Claim 8. *For all* $x, y \in V^*, u \in W^*$ *it holds*

$$y \Rightarrow_G x \text{ if and only if } @y\# \Rightarrow_{G'} @u\#$$

where $x = \omega(u)$.

Proof. Let $b = b_1 b_2 \dots b_n, b_i \in V$ and $x \in V''$, then $\gamma(b, x) = x b_1 x x b_2 x \dots x b_n x$.

Only If: Let $y \Rightarrow_G x$. Express y and x as $y = a_1 A_1 a_2 A_2 \dots a_n A_n a_{n+1}$ and $x = a_1 x_1 a_2 x_2 \dots a_n x_n x_{n+1}$ and corresponding production from P: l: $(A_1, A_2, \dots, A_n) \rightarrow (x_1, x_2, \dots, x_n)$, which is applied during $y \Rightarrow_G x$. Then, for such production exist n corresponding productions in P' (see Construction 1). Then, with use of Claim 3, we can construct $@y\# \Rightarrow_{G'} @\, \langle l, 0\rangle\, \gamma(a_1, \langle l, 0\rangle)\, \langle l, 0\rangle\, x_1\, \langle l, 1\rangle\, \gamma(a_2, \langle l, 1\rangle)\, \langle l, 1\rangle\, \langle l, 2\rangle \dots \langle l, n\rangle\, \gamma(x_{n-1}, \langle l, n-1\rangle)\, \langle l, n-1\rangle\, x_n\, \langle l, n\rangle\, \gamma(x_{n+1}, \langle l, n\rangle)\, \langle l, n\rangle\, \#$, where $\gamma(b, x), a \in V^*, x \in V''$ is defined as this: Obviously, $\omega(y) = a_1 x_1 a_2 x_2 \dots a_n x_n x_{n+1} = x$.

If: Let $@y\# \Rightarrow_{G'} @u\#$. Express y as $y = a_1 a_2 \dots a_n, a_i \in V, n \geq 0$. By the proof of Claim 3 we can see, that each a_i could be rewritten to $\langle l, m\rangle\, x_i\, \langle l, m\rangle$ or to $\langle l, m-1\rangle\, x_i\, \langle l, m\rangle$ (by the proof of Claim 2 we ignore $a_i \in V''$). In the first case it corresponds to use no rule in G. In the second case there will be (by the proof of the claim) n such cases corresponding to use productions derived from original production $(A_1, A_2, \dots, A_n) \rightarrow (y_1, y_2, \dots, y_n)$. Then, $y \Rightarrow_G x$ such that $x = x_1 x_2 \dots x_n = \omega(u)$. □

Claim 9. *For all* $t \in T^*, y \in V^*, u \in W^*$, *it hold*

$$y \Rightarrow_G t \text{ if and only if } @y\# \Rightarrow_{G'} u$$

where $\tau(t) = \omega(u)$.

Proof. By the analogy with the proof of Claim 8. □

From the above claims, it is easy to prove that

$$S \Rightarrow_G^* t \text{ if and only if } S' \Rightarrow_{G'}^+ t$$

for all $t \in T^*$.

Only If: Let $S \Rightarrow_G v_1 \Rightarrow_G v_2 \Rightarrow_G \dots \Rightarrow_G v_n \Rightarrow_t$ for some $n \geq 0$. Then, there exists $S' \Rightarrow_{G'} @S\# \Rightarrow_{G'} @w_1\# \Rightarrow_{G'} @w_2\# \Rightarrow_{G'} \dots \Rightarrow_{G'} @w_n\# \Rightarrow_{G'} u \Rightarrow_{G'} t$, where $v_i = \omega(w_i)$ for all $i = 1, \dots, n$ and $\tau(t) = \omega(u)$.

If: By Claim 7, $S' \Rightarrow^+_{G'} t$ has the form $S' \Rightarrow_{G'} @S\# \Rightarrow_{G'} @w_1\# \Rightarrow_{G'} @w_2\# \Rightarrow_{G'} \ldots \Rightarrow_{G'} @w_n\# \Rightarrow_{G'} u \Rightarrow_{G'} t, n \geq 0$. For this derivation we can construct $S \Rightarrow_G v_1 \Rightarrow_G v_2 \Rightarrow_G \ldots \Rightarrow_G v_n \Rightarrow_G t$ so that $v_i = \omega(w_i)$ for all $i = 1, \ldots, n$.

Therefore, $L(G) = L(G)'$, and the theorem holds. ■

4 Simulation of Phrase-Structured Grammars

Construction 2.

Input: A phrase-structured grammar, $G = (V, T, P, S)$

Output: A symbiotic E0L grammars, G'

Algorithm: Introduce a new alphabet, $V' = V \cup \{@, \#, \widetilde{@}, \widetilde{\#}, S'\} \cup V'' \cup \widetilde{T}, V'' = \{\langle i,j \rangle : 0 < i \leq Card(P), 0 \leq j \leq k\}, \widetilde{T} = \{\widetilde{a} : a \in T\}$. Let τ be a homomorphism from T to $\widetilde{T}$ such that $\tau(a) = \widetilde{a}$ for all $a \in T$. Define a language W, over V' as $W = V \cup \{@, \#, \widetilde{@}, \widetilde{\#}, S'\} \cup \widetilde{T} \cup \{\langle i,j \rangle \langle i,j \rangle : 0 < i \leq Card(P), 0 \leq j \leq k\}$. Then, construct a symbiotic E0L grammar $G' = (W, T, P', S')$, where the set of productions is defined in the following way:

1. add $S' \to @S\#$ to P';
2. for every production $n\colon X_1X_2\ldots X_n \to y \in P$ add these rules to P' (where n is a label, $0 <\leq Card(P)$:
$$\begin{aligned} X_1 &\to \langle n,0 \rangle \tau(y) \langle n,1 \rangle \\ X_2 &\to \langle n,1 \rangle \langle n,2 \rangle \\ X_3 &\to \langle n,2 \rangle \langle n,3 \rangle \\ &\vdots \\ X_n &\to \langle n,k-1 \rangle \langle n,k \rangle \end{aligned}$$
3. add $@ \to @ \langle i,0 \rangle, 0 < i \leq Card(P)$ to P';
4. add $\# \to \langle i,k \rangle \#$, to P' for each production $i\colon X_1X_2\ldots X_k \to y \in P$;
5. add $@ \to \varepsilon$;
6. add $\# \to \varepsilon$;
7. for each $A \in V \cup \widetilde{T}$ add productions of this form to P': $A \to \langle i,j \rangle A \langle i,j \rangle, 0 < i \leq Card(P), 0 \leq j \leq k$;
8. add these productions to P': $\langle i,j \rangle \to \varepsilon, 0 < i \leq Card(P), 0 \leq j \leq k$;
9. add production $\widetilde{a} \to a$ for each $a \in T$ to P'.

Theorem 2. *Let $G = (V, T, P, S)$ be a phrase-structured grammar. Let G' be a symbiotic E0L grammar constructed by Construction 2 with G as its input. Then, $L(G) = L(G')$.*

Proof. Let ω be a homomorphism from V' to $V' - V''$ defined as $\omega(a) = \varepsilon$ for all $a \in V''$, and $\omega(a) = a$, for all $a \in V' - V''$.

Claim 10. *For every $w \in W^*$ holds,*

1. $S' \Rightarrow_G^+ w$ *if and only if* $@S\# \Rightarrow_G^* w$;
2. $S' \Rightarrow_G^+ w$ *implies* $S' \notin sub(w)$.

Proof. By the definition of P', it is easy to see that the very first derivation step always rewrites S' to $@S\#$. Moreover, no productions generate S'; thus, S' appears in no sentential form derived from S'. □

Claim 11. *For all $u, v \in W^+, S' \notin sub(uv), u \Rightarrow_{G'} v$ if and only if $\omega(u) \Rightarrow_{G'} v$.*

Proof. Examine the definition of P'. Clearly, all occurrences of symbols from V'' are always erased during $u \Rightarrow_{G'} v$, so they play no role in the generation of v. By the definition of W and ω, $\omega(u) \in W^*$; therefore $\omega(u) \Rightarrow_{G'}$ is a valid derivation in G'.

Note that this property of derivations in G' allows us to ignore symbols of forms $\langle i, j \rangle$ occuring in left-hand sides of derivation steps. □

Claim 12. *Let $@y\# \Rightarrow_{G'} @x\#$, where $y = a_1a_2 \dots a_n$ for some $a_i \in V, x \in W^*, n \geq 0$. Then, $@x\# = @ \langle i,0 \rangle \langle i,0 \rangle x_1 \langle i,0 \rangle \langle i,0 \rangle x_2 \langle i,0 \rangle \dots \langle i,0 \rangle \langle i,1 \rangle x_m \langle i,2 \rangle \langle i,2 \rangle \langle i,3 \rangle \langle i,3 \rangle \dots \langle i,k \rangle \langle i,k \rangle x_n \langle i,k \rangle \langle i,k \rangle x_{n+1} \langle i,k \rangle \dots \langle i,k \rangle \#$, where $x_j \in V^*$ for all $j = 1, 2, \dots, m$ and some $0 < i \leq Card(P)$.*

Proof. Since x is surrounded by @ and # in $@x\#$, G' surely rewrites $@x\#$ in such way, that @ is rewritten to some $@ \langle i,0 \rangle$ and # to $\langle j,k \rangle, 0 \leq i,j \leq Card(P)$. Every A_l can be rewritten either to $\langle i,j \rangle x_l \langle i,j \rangle$ or (if such production exists) to $\langle i,j-1 \rangle x_l \langle i,j \rangle$, where $0 < i \leq Card(P), 0 \leq j \leq k, x_i \in V^*$. Thus, $@x\# = @ \langle i,0 \rangle \alpha_1 z_1 \beta_1 \alpha_2 z_2 \beta_2 \dots \alpha_n z_n \beta_n \langle j,k \rangle \#$ with $\alpha_l = \langle i,j \rangle, z_l = x_l, \beta_l = \langle i,j \rangle$, or $\alpha_l = \langle i,j-1 \rangle, z_l = x_l, \beta_l = \langle i,j \rangle$, for all $l = 1, 2, \dots n$. However, $@x\#$ must be a string over W. Inspect the definition of W to see that $@x\# \in W^*$ if and only if $\alpha_1 = \langle i,0 \rangle$ and $\beta_n \langle i,k \rangle$. Then, β_1 could be only $\langle i,0 \rangle$ or $\langle i,1 \rangle$. In same way α_n could be only $\langle i,k \rangle$ or $\langle i,k-1 \rangle$. We can simply show, that we can get only sentential form $@ \langle i,0 \rangle \langle i,0 \rangle x_1 \langle i,0 \rangle \langle i,0 \rangle x_2 \langle i,0 \rangle \dots \langle i,0 \rangle \langle i,1 \rangle x_m \langle i,2 \rangle \langle i,2 \rangle \langle i,3 \rangle \langle i,3 \rangle \dots \langle i,k \rangle \langle i,k \rangle x_n \langle i,k \rangle \dots \langle i,k \rangle x_{n+1} \langle i,k \rangle \dots \langle i,k \rangle \#$ □

Claim 13. *Let $@y\# \Rightarrow_{G'} x$, where $y = a_1a_2 \dots a_n$ and $\{@, \#\} \cap sub(x) = \emptyset$ for some $a_i \in V, x \in W^*, n \geq 0$. Then, $x = \langle i,0 \rangle \langle i,0 \rangle \tau(t_1) \langle i,0 \rangle \dots \langle i,k \rangle \tau(t_k) \langle i,k \rangle \langle i,k \rangle$, where $t_i \in T^*$ for all $i = 1, 2, \dots, n$.*

Proof. Prove this claim by analogy with the proof of Claim 12. □

The following claim shows that Claims 12 and 13 cover all possible ways of rewriting of a string having the form $@y\#, y \in V^*$, in G'.

Claim 14. *Let $@y\# \Rightarrow_{G'} u, y \in V^*$. Then, either $u = @x\#, x \in W^*$, or $u \in W^*, \omega(u) \in T^*$, and $\{@, \#\} \cap sub(u) = \emptyset$.*

Proof. Return to the proof of Claim 12. Suppose that @ is rewritten to $@\langle i,0\rangle$ and # is rewritten to ε. Then we can construct only strings of the form $z = @x\langle i,j\rangle y\langle i,k\rangle$, where $x \in W^*, y \in V^*$ and last symbol of y is from $V - V''$. It is clear, that $z \notin W^*$. Analogously, suppose that @ is rewritten to ε and # is rewritten to $\langle i,k\rangle \#$. As before, such a sentential form is out of W^*. □

Claim 15. *Let* $u \Rightarrow_{G'} v, u \in W^*, \{@,\#\} \cap sub(u) = \emptyset$. *Then* $v \in T^*$.

Proof. From the Claim 14 we see, that $\omega(u) \in T^*$. Then, we have to consider only productions with its left sides from $\widetilde{T}$, because it is the only possibility. Such productions are of the form $\widetilde{t} \to \langle i,j\rangle t \langle i,j\rangle$ or $\widetilde{t} \to t$, where $t \in T, 0 < i \leq Card(P), j \geq 1$. Then, string v could have one of the following forms:

1. $u = \langle i,j\rangle t \langle i,j\rangle y, t \in T, 0 < i \leq Card(P), 0 \leq j, y \in (V'' \cup T)^*$;
2. $u = x\langle i,j\rangle t \langle i,j\rangle y, x \in T^*, t \in T, y \in (V'' \cup T)^*$;
3. $u = t_1 t_2 \ldots t_n, t_i \in T$.

It is easy to see, that only third form is the legal one. The others are out of W. □

Claim 16. *Every derivation in* G' *is a prefix of*

$$\begin{aligned} S' &\Rightarrow_{G'} @w_0\# \\ &\Rightarrow_{G'} @w_1\# \\ &\ \ \vdots \\ &\Rightarrow_{G'} @w_n\# \\ &\Rightarrow_{G'} u \\ &\Rightarrow_{G'} t \end{aligned}$$

where $w_0 = S$, $w_i \in W^*$, $\omega(u) = \tau(t)$, $t \in T^*$, $0 \leq i \leq n$, $n \geq 0$.

Proof. By the proof of Claim 10, S' is always rewritten to $@w_0\#$, where $w_0 = S$. Then, Claim 14 tells us that there are two possible forms of derivations rewriting $\omega(@w_i\#)$ and, hence, $@w_i\#$. First, G' can generate a sequence of n sentential forms that belong to $\{@\}W^*\{\#\}$, for some $n \geq 0$ (their form is described in Claim 12). Second, G' can rewrite $@w_n\#$ to $u \in W^*$, satisfying $\omega(u) \in \widetilde{T}^*$ (see Claim 13). By the Claim 15 the only form, to which could be rewritten u is t. Therefore, $u \Rightarrow_{G'} t$ such that $t \in T^*$ and $\omega(u) = \tau(t)$. After that, no other derivation step can be made from t because P' contains no production that rewrites terminals. □

Claim 17. *For all* $x, y \in V^*, u \in W^*$ *it holds*

$$y \Rightarrow_G x \text{ if and only if } @y\# \Rightarrow_{G'} @u\#$$

where $x = \omega(u)$.

Proof. Let $b = b_1 b_2 \dots b_n, b_i \in V$ and $x \in V''$, then $\gamma(b, x) = x b_1 x x b_2 x \dots x b_n x$.

Only If: Let $y \Rightarrow_G x$. Express y and x as $y = a_1 A_1 a_2 A_2 \dots a_n A_n a_{n+1}$ and $x = a_1 x_1 a_2 x_2 \dots a_n x_n x_{n+1}$ and corresponding production from P: l: $(A_1, A_2, \dots, A_n) \to (x_1, x_2, \dots, x_n)$, which is applied during $y \Rightarrow_G x$. Then, for such production exist n corresponding productions in P' (see Construction 2). Then, with use of Claim 12, we can construct $@y\# \Rightarrow_{G'} @ \langle l, 0\rangle \gamma(a_1, \langle l, 0\rangle) \langle l, 0\rangle x_1 \langle l, 1\rangle \gamma(a_2, \langle l, 1\rangle) \langle l, 1\rangle \langle l, 2\rangle \dots \langle l, n\rangle \gamma(x_{n-1}, \langle l, n-1\rangle) \langle l, n-1\rangle x_n \langle l, n\rangle \gamma(x_{n+1}, \langle l, n\rangle) \langle l, n\rangle \#$, where $\gamma(b, x), a \in V^*, x \in V''$ is defined as this: Obviously, $\omega(y) = a_1 x_1 a_2 x_2 \dots a_n x_n x_{n+1} = x$.

If: Let $@y\# \Rightarrow_{G'} @u\#$. Express y as $y = a_1 a_2 \dots a_n, a_i \in V, n \geq 0$. By the proof of Claim 12 we can see, that each a_i could be rewritten to $\langle l, m\rangle x_i \langle l, m\rangle$ or to $\langle l, m-1\rangle x_i \langle l, m\rangle$ (by the proof of Claim 11 we ignore $a_i \in V''$). In the first case it corresponds to use no rule in G. In the second case there will be (by the proof of the claim) n such cases corresponding to use productions derived from original production $(A_1, A_2, \dots, A_n) \to (y_1, y_2, \dots, y_n)$. Then, $y \Rightarrow_G x$ such that $x = x_1 x_2 \dots x_n = \omega(u)$. □

Claim 18. *For all $t \in T^*, y \in V^*, u \in W^*$, it hold*

$$y \Rightarrow_G t \text{ if and only if } @y\# \Rightarrow_{G'} u$$

where $\tau(t) = \omega(u)$.

Proof. By the analogy with the proof of Claim 17. □

From the above claims, it is easy to prove that

$$S \Rightarrow^*_G t \text{ if and only if } S' \Rightarrow^+_{G'} t$$

for all $t \in T^*$.

Only If: Let $S \Rightarrow_G v_1 \Rightarrow_G v_2 \Rightarrow_G \dots \Rightarrow_G v_n \Rightarrow_t$ for some $n \geq 0$. Then, there exists $S' \Rightarrow_{G'} @S\# \Rightarrow_{G'} @w_1\# \Rightarrow_{G'} @w_2\# \Rightarrow_{G'} \dots \Rightarrow_{G'} @w_n\# \Rightarrow_{G'} u \Rightarrow_{G'} t$, where $v_i = \omega(w_i)$ for all $i = 1, \dots, n$ and $\tau(t) = \omega(u)$.

If: By Claim 16, $S' \Rightarrow^+_{G'} t$ has the form $S' \Rightarrow_{G'} @S\# \Rightarrow_{G'} @w_1\# \Rightarrow_{G'} @w_2\# \Rightarrow_{G'} \dots \Rightarrow_{G'} @w_n\# \Rightarrow_{G'} u \Rightarrow_{G'} t, n \geq 0$. For this derivation we can construct $S \Rightarrow_G v_1 \Rightarrow_G v_2 \Rightarrow_G \dots \Rightarrow_G v_n \Rightarrow_G t$ so that $v_i = \omega(w_i)$ for all $i = 1, \dots, n$.

Therefore, $L(G) = L(G)'$, and the theorem holds. ■

5 Derivation simulations

5.1 Definitions

Now we have to repeat some needed definitions. Definitions as a whole were introduced in [1] and there can be found reasons of their existence and so on. Here we only repeat their readings because they will be used in the following subsections.

Definition 1. A *string-relation system* is a quadruple $\Psi = (W, \Rightarrow, W_0, W_F)$, where W is a language, $\Rightarrow$ is a binary relation on W, $W_0 \subseteq W$ is a set of *start strings*, and $W_F \subseteq W$ is a set of *final strings*.

Every string, $w \in W$, represents a 0-step string-relation sequence in Ψ. For every $n \geq 1$, a sequence $w_0, w_1, \dots w_n$, $w_i \in W$, $0 \leq i \leq n$, is an *n-step string-relation sequence*, symbolically written as $w_0 \Rightarrow w_1 \Rightarrow \dots \Rightarrow w_n$, if for each $0 \leq i \leq n-1$, $w_i \Rightarrow w_{i+1}$.

If there is a string-relation sequence $w_0 \Rightarrow w_1 \Rightarrow \dots \Rightarrow w_n$, where $n \geq 0$, we write $w_0 \Rightarrow^n w_n$. Furthermore, $w_0 \Rightarrow^* w_n$ means that $w_0 \Rightarrow^n w_n$ for some $n \geq 0$, and $w_0 \Rightarrow^+ w_n$ means that $w_0 \Rightarrow^n w_n$ for some $n \geq 1$. Obviously, from the mathematical point of view, $\Rightarrow^+$ and $\Rightarrow^*$ are the transitive closure of $\Rightarrow$ and the transitive and reflexive closure of $\Rightarrow$, respectively.

Let $\Psi = (W, \Rightarrow, W_0, W_F)$ be a string-relation system. A string-relation sequence in Ψ, $u \Rightarrow^* v$, where $u, v \in W$, is called a *yield sequence*, if $u \in W_0$. If $u \Rightarrow^* v$ is a yield sequence and $v \in W_F$, $u \Rightarrow^* v$ is *successful*.

Let $D(\Psi)$ and $SD(\Psi)$ denote the set of all yield sequences and all successful yield sequences in Ψ, respectively.

Definition 2. Let $\Psi = (W, \Rightarrow_\Psi, W_0, W_F)$ and $\Omega = (W', \Rightarrow_\Omega, W'_0, W'_F)$ be two string-relation systems, and let σ be a substitution from W' to W. Furthermore, let d be a yield sequence in Ψ of the form $w_0 \Rightarrow_\Psi w_1 \Rightarrow_\Psi \dots \Rightarrow_\Psi w_{n-1} \Rightarrow_\Psi w_n$, where $w_i \in W$, $0 \leq i \leq n$, for some $n \geq 0$. A yield sequence, h, in Ω *simulates d with respect to σ*, symbolically written as $h \rhd_\sigma d$, if h is of the form $y_0 \Rightarrow_\Omega^{m_1} y_1 \Rightarrow_\Omega^{m_2} \dots \Rightarrow_\Omega^{m_{n-1}} y_{n-1} \Rightarrow_\Omega^{m_n} y_n$, where $y_j \in W'$, $0 \leq j \leq n$, $m_k \geq 1$, $1 \leq k \leq n$, and $w_i \in \sigma(y_i)$ for all $0 \leq i \leq n$. If, in addition, there exists $m \geq 1$ such that $m_k \leq m$ for each $1 \leq k \leq n$, then h *m-closely simulates d with respect to σ*, symbolically written as $h \rhd_\sigma^m d$.

Definition 3. Let $\Psi = (W, \Rightarrow_\Psi, W_0, W_F)$ and $\Omega = (W', \Rightarrow_\Omega, W'_0, W'_F)$ be two string-relation systems, and let σ be a substitution from W' to W. Let $X \subseteq D(\Psi)$ and $Y \subseteq D(\Omega)$. *Y simulates X with respect to σ*, written as $Y \rhd_\sigma X$, if the following two conditions hold:

1. for every $d \in X$, there is $h \in Y$ such that $h \rhd_\sigma d$;
2. for every $h \in Y$, there is $d \in X$ such that $h \rhd_\sigma d$.

Let m be a positive integer. *Y m-closely simulates X with respect to σ*, $Y \rhd_\sigma^m X$, provided that:

1. for every $d \in X$, there is $h \in Y$ such that $h \rhd_\sigma^m d$;
2. for every $h \in Y$, there is $d \in X$ such that $h \rhd_\sigma^m d$.

Definition 4. Let $\Psi = (W, \Rightarrow_\Psi, W_0, W_F)$ and $\Omega = (W', \Rightarrow_\Omega, W'_0, W'_F)$ be two string-relation systems. If there exists a substitution σ from W' to W such that $D(\Omega) \rhd_\sigma D(\Psi)$ and $SD(\Omega) \rhd_\sigma SD(\Psi)$, then Ω is said to be Ψ's *derivation simulator* and *successful-derivation simulator*, respectively. Furthermore, if there is an integer, $m \geq 1$, such that

$D(\Omega) \rhd_\sigma^m D(\Psi)$ and $SD(\Omega) \rhd_\sigma^m SD(\Psi)$, Ω is called an *m-close derivation simulator* and *m-close successful-derivation simulator* of Ψ, respectively. If there exists a homomorphism ρ from W' to W such that $D(\Omega) \rhd_\rho D(\Psi)$, $SD(\Omega) \rhd_\rho SD(\Psi)$, $D(\Omega) \rhd_\rho^m D(\Psi)$, and $SD(\Omega) \rhd_\rho^m SD(\Psi)$, then Ω is Ψ's *homomorphic derivation simulator*, *homomorphic successful-derivation simulator*, *m-close homomorphic derivation simulator* and *m-close homomorphic successful-derivation simulator*, respectively.

5.2 Derivation simulation of Scattered Context Grammars

Definition 5. Let $G = (V, T, P, S)$ be a scattered context grammar. Let $\Rightarrow_G$ be the direct derivation relation in G. For $\Rightarrow_G$ and every $l \geq 0$, set

$$\Delta(\Rightarrow_G, l) = \{x \Rightarrow_G y : x \Rightarrow_G y \Rightarrow_G^i w, x, y \in V^*, w \in T^*, i + 1 = l, i \geq 0\}.$$

Next, let $G_1 = (V_1, T_1, P_1, S_1)$ and $G_2 = (V_2, T_2, P_2, S_2)$ be scattered context grammars. Let $\Rightarrow_{G_1}$ and $\Rightarrow_{G_2}$ be the derivation relations of G_1 and G_2, respectively. Let σ be a substitution from V_2 to V_1. *G_2 simulates G_1 with respect to σ*, $D(G_2) \rhd_D (G_1)$ in symbols, if there exists two natural numbers $k, l \geq 0$ so that the following conditions hold:

1. $\Psi_1 = (V_1^*, \Rightarrow_{G_1}, \{S_1\}, T_1^*)$ and $\Psi_2 = (V_2^*, \Rightarrow_{\Psi_2}, W_0, W_F)$ are string-relation systems corresponding to G_1 and G_2, respectively, where $W_0 = \{x \in V_2^* : S_2 \Rightarrow_{G_2}^k x\}$ and $W_F = \{x \in V_2^* : x \Rightarrow_{G_2}^l w, w \in T_2^*, \sigma(w) \subseteq T_1^*\}$;
2. relation $\Rightarrow_{\Psi_2}$ coincides with $\Rightarrow_{G_2} - \Delta(\Rightarrow_{G_2}, l)$;
3. $D(\Psi_2) \rhd_\sigma D(\Psi_1)$.

In case that $SD(\Psi_2) \rhd_\sigma SD(\Psi_1)$, *$G_2$ simulates successful derivations of G_1 with respect to σ*; in symbols, $SD(G_2) \rhd_\sigma SD(G_1)$.

Definition 6. Let G_1 and G_2 be scattered context grammars with total alphabets V_1 and V_2, terminal alphabets T_1 and T_2, and axioms S_1 and S_2, respectively. Let σ be a substitution from V_2 to V_1. *G_2 m-closely simulates G_1 with respect to σ* if $D(G_2) \rhd_\sigma D(G_1)$ and there exists $m \geq 1$ such that the corresponding string-relation systems Ψ_1 and Ψ_2 satisfy $D(\Psi_2) \rhd_\sigma^m D(\Psi_1)$. In symbols, $D(G_2) \rhd_\sigma^m D(G_1)$.

Analogously, *G_2 m-closely simulates successful derivations of G_1 with respect to σ*, denoted by $SD(G_2) \rhd_\sigma^m SD(G_1)$, if $SD(\Psi_2) \rhd_\sigma^m SD(\Psi_1)$ and there exists $m \geq 1$ such that $SD(G_2) \rhd_\sigma^m SD(G_1)$.

Definition 7. Let G_1 and G_2 be two scattered context grammars. If there exists a substitution σ such that $D(G_2) \rhd_\sigma D(G_1)$, then G_2 is said to be G_1's *derivation simulator*.

By analogy with Definition 7, the reader can also define *homomorphic*, *m-close*, and *successful-derivation simulators* of scattered context grammars.

Theorem 3. *Let $G = (V, T, P, S)$ be a scattered context grammar and $G' = (W, T, P', S')$ be a symbiotic E0L grammar constructed by using Construction 1 with G as its input. Then, there exists a homomorphism $\widetilde{\omega}$ such that $D(G') \rhd^1_{\widetilde{\omega}} D(G)$ and $SD(G') \rhd^1_{\widetilde{\omega}} SD(G)$.*

Proof. Let $\Psi = (V^*, \Rightarrow_G, \{S\}, T^*)$ be a string-relation system corresponding to G. Let $\widetilde{\omega}$ be the homomorphism defined in the proof of Theorem 1. Let $\Psi' = ((V')^*, \Rightarrow_{\Psi'}, W_0, W_F)$ be a string-relation system corresponding to G', where

$$\begin{aligned}
\Rightarrow_{\Psi'} &= \Rightarrow_{G'} - \{\langle i,0\rangle \langle i,0\rangle \tau(t_1) \langle i,0\rangle \dots \langle i,k\rangle \tau(t_n) \langle i,k\rangle \langle i,k\rangle\} \Rightarrow_{G'} t_1 t_2 \dots t_n : \\
&\qquad 0 < i \le Card(P), t_j \in T^*, 1 \le j \le n, n \ge 0\} \\
W_0 &= \{@S\#\} \\
W_F &= \{\langle i,0\rangle \langle i,0\rangle \tau(t_1) \langle i,0\rangle \dots \langle i,k\rangle \tau(t_n) \langle i,k\rangle \langle i,k\rangle : \\
&\qquad 0 < i \le Card(P), t_j \in T^*, 1 \le j \le n, n \ge 0\}
\end{aligned}$$

It is easy to verify, that Ψ and Ψ' satisfy (1) through (3) of Definition 5; of course $S' \Rightarrow^1_{G'} @S\#$ and for every $u \in W_F, u \Rightarrow_1 G't$, where $t \in T^*$ (see Claim 7 in the proof of Theorem 1). Next, we show that $D(\Psi') \rhd^1_{\widetilde{\omega}} D(\Psi)$. By Definition 3, we have to establish that

1. for every $d \in D(\Psi)$, there exists $h \in D(\Psi')$ such that $h \rhd^1_{\widetilde{\omega}} d$;
2. for every $h \in D(\Psi')$, there exists $d \in D(\Psi)$ so that $h \rhd^1_{\widetilde{\omega}} d$.

(Note that most of this proof is based on substitutions and claims introduced in the proof of Theorem 1).

(1) Let $d \in D(\Psi)$. Express d as $d = v_0 \Rightarrow_G v_1 \Rightarrow_G v_2 \Rightarrow_G \dots \Rightarrow_G v_n, v_0 = S$, for some $n \ge 0$. For $n = 0$, there is $@S\# \in \Psi'$ such that zero-length derivations S and $@S\#$ satisfy $S \rhd_1 \widetilde{\omega}@S\#$. Assume that $n > 0$. Then, according to Claims 2 and 8, $v_i \Rightarrow_G v_{i+1}$ if and only if $@w_i\# \Rightarrow_{G'} @w_{i+1}\#$, where $v_{i+1} = \omega(w_{i+1}) = \widetilde{\omega}(@w_{i+1}\#), w_i, w_{i+1} \in W^*, 0 \le i \le n-1$. Moreover, by the definition of $\Psi', @w_i\# \Rightarrow_{\Psi'} @w_{i+1}\#$ for all $i = 0, \dots, n-1$. Hence, by induction on the length of derivations in G, the reader can easily establish that for every $d \in D(\Psi)$, there exists $h \in D(\Psi')$ such that $h \rhd^1_{\widetilde{\omega}} d$.

(2) Let $h \in D(\Psi)$. By the definition of $\Rightarrow_{\Psi'}$ and Claim 7, every yield sequence in Ψ' is a prefix of $@w_0\# \Rightarrow_{\Psi'} @w_1\# \Rightarrow_{\Psi'} \dots \Rightarrow_{\Psi'} @w_n\# \Rightarrow_{\Psi'} u$, where $w_0 = s$, $w_i \in W^*$, $u \in W_F$, $0 \le i \le n$, $n \le 0$. The zero-length derivation $@s\#$ is a 1-close simulation of s from G. Claims 2 and 8 imply that for every $@w_i\# \Rightarrow_{\Psi'} @w_{i+1}\#$, there exists $v_i \Rightarrow_G v_{i+1}$ for some $v_i, v_{i+1} \in V^*$, $v_{i+1} = \omega(w_{i+1}) = \widetilde{\omega}(@w_{i+1}\#)$, $0 \le i \le n-1$. Furthermore, according to Claims 4 and 9, for $@w_n\# \Rightarrow_{\Psi'} u$, there exists $v_n \Rightarrow_G t$ such that $t \in T^*$, $\tau(t) = \omega(u)$; that is, $\widetilde{\omega}(u) = t$. Clearly, every derivation step in h is a simulation of a corresponding derivation step in d; as a result, $h \rhd^1_{\widetilde{\omega}} d$.

Next, we prove that $SD(G') \rhd^1_{\widetilde{\omega}} SD(G)$. From (2), it follows that every successful yield sequence $h \in SD(\Psi')$ is a 1-close simulation of a derivation $s \Rightarrow^*_G t$ with $t \in T^*$. To prove that for every $d \in SD(\Psi)$, there exists $h \in SD(\Psi')$ such that $h \rhd^1_{\widetilde{\omega}} d$, return to case (1) in this proof. Assume that $v_0 \Rightarrow_G^n v_n$, $v_n \in T^*$, $n \geq 1$. Then, there exists a derivation $@w_{n-1}\# \Rightarrow_{\Psi'} u$, $u \in W_F$ (see Claim 9), such that $\tau(v_n) = \omega(u)$ which implies $\widetilde{\omega}(u) = v_n$. Therefore, we get $h \rhd_{\widetilde{\omega}}^1 d$, so $SD(G') \rhd_{\widetilde{\omega}}^1 SD(G)$. $\square$

Theorems 1 and 3 show that for every scattered context grammar $G = (V, T, P, S)$, there exists a symbiotic E0L grammar $G' = (W', T, P', S')$ such that

1. $L(G) = L(G')$;
2. G' is a 1-close homomorphic derivation simulator of G';
3. G' is a 1-close homomorphic successful-derivation simulator of G;
4. To simulate G, G' uses one initial derivation step ($S' \Rightarrow_{G'} @S\#$) and one derivation step that removes auxiliary symbols ($\langle i,0\rangle\,\langle i,0\rangle\,\tau(t_1)\,\langle i,0\rangle \ldots \langle i,k\rangle\,\tau(t_n)\,\langle i,k\rangle\,\langle i,k\rangle \Rightarrow_{G'} t_1t_2\ldots t_n : 0 < i \leq Card(P), t_j \in T^*, 1 \leq j \leq n, n \geq 0$).

5.3 Derivation simulation of Phrase-Structured Grammars

Definition 8. Let $G = (V, T, P, S)$ be a phrase-structured grammar. Let $\Rightarrow_G$ be the direct derivation relation in G. For $\Rightarrow_G$ and every $l \geq 0$, set

$$\Delta(\Rightarrow_G, l) = \{x \Rightarrow_G y : x \Rightarrow_G y \Rightarrow_G^i w, x, y \in V^*, w \in T^*, i + 1 = l, i \geq 0\}.$$

Next, let $G_1 = (V_1, T_1, P_1, S_1)$ and $G_2 = (V_2, T_2, P_2, S_2)$ be phrase-structured grammars. Let $\Rightarrow_{G_1}$ and $\Rightarrow_{G_2}$ be the derivation relations of G_1 and G_2, respectively. Let σ be a substitution from V_2 to V_1. *G_2 simulates G_1 with respect to σ*, $D(G_2) \rhd_D (G_1)$ in symbols, if there exists two natural numbers $k, l \geq 0$ so that the following conditions hold:

1. $\Psi_1 = (V_1^*, \Rightarrow_{G_1}, \{S_1\}, T_1^*)$ and $\Psi_2 = (V_2^*, \Rightarrow_{\Psi_2}, W_0, W_F)$ are string-relation systems corresponding to G_1 and G_2, respectively, where $W_0 = \{x \in V_2^* : S_2 \Rightarrow_{G_2}^k x\}$ and $W_F = \{x \in V_2^* : x \Rightarrow_{G_2}^l w, w \in T_2^*, \sigma(w) \subseteq T_1^*\}$;
2. relation $\Rightarrow_{\Psi_2}$ coincides with $\Rightarrow_{G_2} - \Delta(\Rightarrow_{G_2}, l)$;
3. $D(\Psi_2) \rhd_\sigma D(\Psi_1)$.

In case that $SD(\Psi_2) \rhd_\sigma SD(\Psi_1)$, *$G_2$ simulates successful derivations of G_1 with respect to σ*; in symbols, $SD(G_2) \rhd_\sigma SD(G_1)$.

Definition 9. Let G_1 and G_2 be phrase-structured grammars with total alphabets V_1 and V_2, terminal alphabets T_1 and T_2, and axioms S_1 and S_2, respectively. Let σ be a substitution from V_2 to V_1. *G_2 m-closely simulates G_1 with respect to σ* if $D(G_2) \rhd_\sigma D(G_1)$ and there exists $m \geq 1$ such that the corresponding string-relation systems Ψ_1 and Ψ_2 satisfy $D(\Psi_2) \rhd_\sigma^m D(\Psi_1)$. In symbols, $D(G_2) \rhd_\sigma^m D(G_1)$.

Analogously, *G_2 m-closely simulates successful derivations of G_1 with respect to σ*, denoted by $SD(G_2) \rhd_\sigma^m SD(G_1)$, if $SD(\Psi_2) \rhd_\sigma^m SD(\Psi_1)$ and there exists $m \geq 1$ such that $SD(G_2) \rhd_\sigma^m SD(G_1)$.

Definition 10. Let G_1 and G_2 be two phrase-structured grammars. If there exists a substitution σ such that $D(G_2) \rhd_\sigma D(G_1)$, then G_2 is said to be G_1's *derivation simulator.*

By analogy with Definition 7, the reader can also define *homomorphic*, *m-close*, and *successful-derivation simulators* of phrase-structured grammars.

Theorem 4. *Let $G = (V, T, P, S)$ be a phrase-structured grammar, $G' = (W, T, P', S')$ be a symbiotic E0L grammar constructed by using Construction 2 with G as its input. Then, there exists a homomorphism $\widetilde{\omega}$ such that $D(G') \rhd^1_{\widetilde{\omega}} D(G)$ and $SD(G') \rhd^1_{\widetilde{\omega}} SD(G)$.*

Proof. Let $\Psi = (V^*, \Rightarrow_G, \{S\}, T^*)$ be a string-relation system corresponding to G. Let $\widetilde{\omega}$ be the homomorphism defined in the proof of Theorem 2. Let $\Psi' = ((V')^*, \Rightarrow_{\Psi'}, W_0, W_F)$ be a string-relation system corresponding to G', where

$$\begin{aligned} \Rightarrow_{\Psi'} &= \Rightarrow_{G'} - \{\langle i,0\rangle \langle i,0\rangle \tau(t_1) \langle i,0\rangle \ldots \langle i,k\rangle \tau(t_n) \langle i,k\rangle \langle i,k\rangle\} \Rightarrow_{G'} t_1 t_2 \ldots t_n : \\ &\quad\; 0 < i \leq Card(P), t_j \in T^*, 1 \leq j \leq n, n \geq 0\} \\ W_0 &= \{@S\#\} \\ W_F &= \{\langle i,0\rangle \langle i,0\rangle \tau(t_1) \langle i,0\rangle \ldots \langle i,k\rangle \tau(t_n) \langle i,k\rangle \langle i,k\rangle : \\ &\quad\; 0 < i \leq Card(P), t_j \in T^*, 1 \leq j \leq n, n \geq 0\} \end{aligned}$$

It is easy to verify, that Ψ and Ψ' satisfy (1) through (3) of Definition 8; of course $S' \Rightarrow^1_{G'} @S\#$ and for every $u \in W_F, u \Rightarrow_1 G't$, where $t \in T^*$ (see Claim 16 in the proof of Theorem 2). Next, we show that $D(\Psi') \rhd^1_{\widetilde{\omega}} D(\Psi)$. By Definition 3, we have to establish that

1. for every $d \in D(\Psi)$, there exists $h \in D(\Psi')$ such that $h \rhd^1_{\widetilde{\omega}} d$;
2. for every $h \in D(\Psi')$, there exists $d \in D(\Psi)$ so that $h \rhd^1_{\widetilde{\omega}} d$.

(Note that most of this proof is based on substitutions and claims introduced in the proof of Theorem 2).

(1) Let $d \in D(\Psi)$. Express d as $d = v_0 \Rightarrow_G v_1 \Rightarrow_G v_2 \Rightarrow_G \ldots \Rightarrow_G v_n, v_0 = S$, for some $n \geq 0$. For $n = 0$, there is $@S\# \in \Psi'$ such that zero-length derivations S and $@S\#$ satisfy $S \rhd_1 \widetilde{\omega}@S\#$. Assume that $n > 0$. Then, according to Claims 11 and 17, $v_i \Rightarrow_G v_{i+1}$ if and only if $@w_i\# \Rightarrow_{G'} @w_{i+1}\#$, where $v_{i+1} = \omega(w_{i+1}) = \widetilde{\omega}(@w_{i+1}\#), w_i, w_{i+1} \in W^*, 0 \leq i \leq n-1$. Moreover, by the definition of $\Psi', @w_i\# \Rightarrow_{\Psi'} @w_{i+1}\#$ for all $i = 0, \ldots, n-1$. Hence, by induction on the length of derivations in G, the reader can easily establish that for every $d \in D(\Psi)$, there exists $h \in D(\Psi')$ such that $h \rhd^1_{\widetilde{\omega}} d$.

(2) Let $h \in D(\Psi)$. By the definition of $\Rightarrow_{\Psi'}$ and Claim 16, every yield sequence in Ψ' is a prefix of $@w_0\# \Rightarrow_{\Psi'} @w_1\# \Rightarrow_{\Psi'} \ldots \Rightarrow_{\Psi'} @w_n\# \Rightarrow_{\Psi'} u$, where $w_0 = s$, $w_i \in W^*$, $u \in W_F$, $0 \leq i \leq n$, $n \leq 0$. The zero-length derivation $@s\#$ is a 1-close simulation of s from G. Claims 11 and 17 imply that for every $@w_i\# \Rightarrow_{\Psi'} @w_{i+1}\#$, there exists $v_i \Rightarrow_G v_{i+1}$ for some $v_i, v_{i+1} \in V^*$, $v_{i+1} = \omega(w_{i+1}) = \widetilde{\omega}(@w_{i+1}\#)$, $0 \leq i \leq n-1$.

Furthermore, according to Claims 13 and 18, for $@w_n\# \Rightarrow_{\Psi'} u$, there exists $v_n \Rightarrow_G t$ such that $t \in T^*$, $\tau(t) = \omega(u)$; that is, $\widetilde{\omega}(u) = t$. Clearly, every derivation step in h is a simulation of a corresponding derivation step in d; as a result, $h \rhd^1_{\widetilde{\omega}} d$.

Next, we prove that $SD(G') \rhd^1_{\widetilde{\omega}} SD(G)$. From (2), it follows that every successful yield sequence $h \in SD(\Psi')$ is a 1-close simulation of a derivation $s \Rightarrow^*_G t$ with $t \in T^*$. To prove that for every $d \in SD(\Psi)$, there exists $h \in SD(\Psi')$ such that $h \rhd^1_{\widetilde{\omega}} d$, return to case (1) in this proof. Assume that $v_0 \Rightarrow^n_G v_n$, $v_n \in T^*$, $n \geq 1$. Then, there exists a derivation $@w_{n-1}\# \Rightarrow_{\Psi'} u$, $u \in W_F$ (see Claim 18), such that $\tau(v_n) = \omega(u)$ which implies $\widetilde{\omega}(u) = v_n$. Therefore, we get $h \rhd^1_{\widetilde{\omega}} d$, so $SD(G') \rhd^1_{\widetilde{\omega}} SD(G)$. □

Theorems 2 and 4 show that for every phrase-structured grammar $G = (V, T, P, S)$, there exists a symbiotic E0L grammar $G' = (W', T, P', S')$ such that

1. $L(G) = L(G')$;
2. G' is a 1-close homomorphic derivation simulator of G';
3. G' is a 1-close homomorphic successful-derivation simulator of G;
4. To simulate G, G' uses one initial derivation step ($S' \Rightarrow_{G'} @S\#$) and one derivation step that removes auxiliary symbols ($\langle i,0\rangle \langle i,0\rangle \tau(t_1) \langle i,0\rangle \dots \langle i,k\rangle \tau(t_n) \langle i,k\rangle \langle i,k\rangle \Rightarrow_{G'} t_1 t_2 \dots t_n : 0 < i \leq Card(P), t_j \in T^*, 1 \leq j \leq n, n \geq 0$).

6 Conclusion

In this paper we have gained following results:

1. Every scattered context grammar G can be simulated by a symbiotic E0L grammar G', while these claims hold:
   a) $L(G) = L(G')$;
   b) G' is a 1-close homomorphic derivation simulator of G;
   c) G' is a 1-close homomorphic successful-derivation simulator of G;
2. Every phrase-structured grammar G can be simulated by a symbiotic E0L grammar G', while these claims hold:
   a) $L(G) = L(G')$;
   b) G' is a 1-close homomorphic derivation simulator of G;
   c) G' is a 1-close homomorphic successful-derivation simulator of G;

References

[1] A. Meduna and M. Švec, A Formalization of Derivation Similarity in the Formal Language Theory and its Illustration in Terms of Lindenmayer systems, *Fundamenta Informaticae* **20** (2004)

[2] H. Hamburger and D. Richards, *Logic and Language Models for Computer Science*, Prentice Hall, London, 2002.

[3] A. Meduna, Symbiotic E0L Systems, *Acta Cybernetica* **10** (1992), 165–172.

[4] A. Meduna, *Automata and Languages: Theory and Applications*, Springer, London, 2000.

[5] A. Meduna and M. Švec, Forbidding ET0L grammars, *Theoretical Computer Science* **306** (2003), 449–469.

[6] A. Parkes, *Introduction to Languages, Machines, and Logic: Computable Languages, Abstract Machines, and Formal Logic*, Springer, London, 2002.

[7] G. Rozenberg and A. Salomaa, *The Mathematical Theory of L-Systems*, Academic Press, New York, 1980.

[8] G. Rozenberg and A. Salomaa. (editors), *The Book of L*, Springer, Berlin, 1986.