A Pumping Lemma for Permitting Semi-conditional Languages

Zsolt Gazdag*

Department of Foundations of Computer Science
University of Szeged
gazdag@inf.u-szeged.hu

Krisztián Tichler

Department of Algorithms and their Applications
Eötvös Loránd University
ktichler@inf.elte.hu

Erzsébet Csuhaj-Varjú

Department of Algorithms and their Applications
Eötvös Loránd University
csuhaj@inf.elte.hu

Permitting semi-conditional grammars are such extensions of context-free grammars where each rule is associated with a word w and such a rule can be applied to a sentential form u only if w is a subword of u. We consider a generalization of these grammars where each rule r is associated with a set of words P and r is applicable only if every word in P occurs in u. The paper investigates the generative power of these grammars with no erasing rules. A pumping lemma is proven for their languages, and it is shown that they are strictly weaker than context-sensitive grammars. Moreover, their generating power is compared to that of forbidding random context grammars with no erasing rules.

Keywords: semi-conditional grammars; permitting context; generative power

1. Introduction

Context-free (CF) grammars are extensively studied since they serve as formal models in many areas of computer science. They have many good properties. For example, their membership problem is efficiently solvable. CF grammars were invented by Noam Chomsky to describe the structures of words in sentences of natural languages. However, it turned out that certain natural languages contain phenomena such as cross-serial dependencies that cannot be handled by CF grammars (see e.g. [12]). The more powerful context-sensitive (CS) grammars are able to model cross-serial dependencies, but the membership problem for them is PSPACE-complete, i.e., not efficiently solvable.

2

One way to enrich CF grammars with context sensitivity and raise their generative power is to control their derivations by context conditions. For example, in conditional grammars (CGs) [7, 16] a regular language is added to every context-free rule and a rule is applied only to sentential forms in the associated language. It turned out that these grammars are equivalent to CS grammars when erasing rules are not allowed, and with erasing rules they are Turing-equivalent [19].

Many variations of conditional grammars have already been investigated. In *random context grammars* (RCGs) [20] two sets of nonterminals, a permitting P and a forbidding one Q, are associated to every context-free rule. Then a rule is applicable, if it is applicable in the context-free sense and nonterminals in Q do not occur, while every nonterminal in P does occur in the current sentential form. If in an RCG each rule is associated with an empty forbidding set (resp. permitting set), then the grammar is called a *permitting* (resp. *forbidding*) RCG.

It turned out that RCGs have equal power to that of Turing machines (see e.g. [3]), thus recently a restricted variant of them was introduced and investigated [15]. In these grammars the permitting and forbidding sets are associated to the nonterminals rather than to the rules. Moreover, one of these sets is always a singleton and the other one is empty. We will call these grammars *restricted random context grammars* (rRCGs). It was shown that even with this very limited ability of controlling the derivations these grammars are equivalent to random context grammars [2, 15]. Moreover, permitting rRCGs are as powerful as permitting RCGs [2], and this is the case for the forbidding variants too if erasing rules are allowed [9].

Păun [17], motivated by the grammars of Kelemen [13], introduced another variant of conditional grammars called *semi-conditional grammars* (SCGs). In these grammars every rule r is associated with two words, a permitting word w_1 and a forbidding one w_2, and r is applicable only if w_1 is a subword of the sentential form, but w_2 is not. An SCG G is of degree (i, j) if the length of its permitting words is at most i and that of the forbidding words is at most j. It was shown in [17] that these grammars with degree $(1, 0)$ or $(0, 1)$ can already generate non-semilinear (hence non-context-free) languages. Moreover, with degree $(1, 2)$ or $(2, 1)$ they determine exactly the class of CS languages if erasing rules are not allowed and with erasing rules they are Turing-equivalent.

In [17] it was also shown that SCGs without erasing rules and with degree $(1, 1)$ cannot generate every CS language. However, it remained an open question if this property still holds if we consider SCGs with degree $(i, 0)$ for some $i \geq 2$. In [10] we gave a negative negative answer to this question by showing that there is a CS language that cannot be generated by any permitting SCG (permitting SCGs are SCGs with degree $(i, 0)$ for some $i \geq 0$).

The present paper is an extended version of [10]. Here we consider SCGs in a more general form: each rule is associated with two sets of words P and Q, and such a rule can be applied to a sentential form u only if every word of P is a subword of u, and no word in Q is a subword of u (cf. also Definition 3.2.6 in [18]). We call these grammars *generalized semi-conditional grammars* or gSCGs to be short
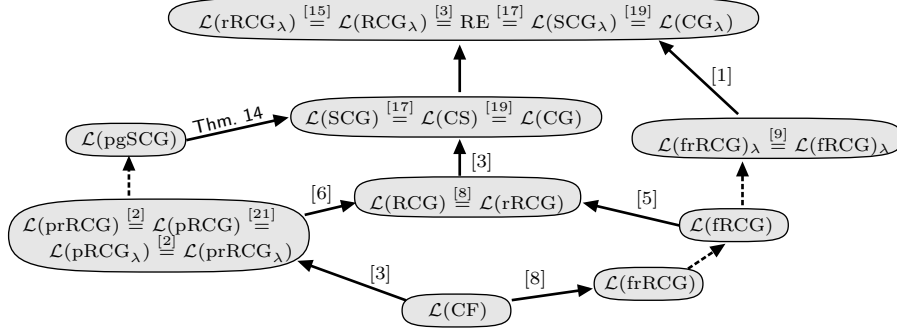
Fig. 1. A comparison of the power of some variants of grammars mentioned in the introduction. Arrows with solid lines represent strict inclusions, while arrows with dashed lines indicate inclusions which are not known to be strict. References to the presented equalities or strict inclusions are also given. Inclusions represented by dashed lines follow from definitions. $gSCG_\lambda$, RCG_λ, and $rRCG_\lambda$ (resp. gSCG, RCG, and rRCG) denote the classes of the corresponding grammars with erasing rules (resp. with no erasing rules). For a class of grammars C, $\mathcal{L}(C)$ denotes the class of languages generated by grammars in C, and pC (resp. fC) denotes that subclass of C, where only permitting (resp. forbidding) context conditions are used.

(notice that gSCGs are generalizations of RCGs too). It is known that gSCGs have equal computational power to that of SCGs (see Theorem 3.2.6 in [18]). Concerning the power of permitting gSCGs (pgSCGs for short), in this paper we show that they are still not able to generate every CS language. A comparison of some language classes generated by grammars discussed so far is given in Figure 1.

The key to prove the results of this paper is a pumping lemma (Lemma 13) which was motivated by a pumping lemma for permitting RCGs proved in [6]. In more details, in [6] it was shown that sufficiently long derivations of a permitting RCG with no erasing rules always contain two sentential forms α and β such that β is derived from α and, for every nonterminal A, $|\alpha|_A \leq |\beta|_A$ (here $|\alpha|_A$ and $|\beta|_A$ denote the number of occurrences of A in α and β, respectively). This property follows from Dickson's lemma [4] which states that any infinite sequence $v_1, v_2, \ldots$ of n-vectors over the natural numbers contains an infinite sub-sequence $v_{i_1} \leq v_{i_2} \leq \cdots$, where $\leq$ is the componentwise ordering of n-vectors. To prove our pumping lemma we had to employ such sentential forms α and β that satisfy a stronger condition: if u is a permitting word of G, then β should contain at least as many occurrences of u as the number of these strings is in α. In our work we used Higman's lemma [11], which ensures that in any infinite sequence $v_1, v_2, \ldots$ of words, there is an infinite subsequence $v_{i_1} \leq_s v_{i_2} \leq_s \cdots$, where $\leq_s$ is the subsequence (or scattered subword) relation. However, to find an appropriate α and β we could not apply directly Higman's lemma to the sentential forms of a derivation, but rather to certain carefully defined words obtained from these sentential forms.

Using Lemma 13 we could compare the generative power of permitting gSCGs

4

and forbidding RCGs as follows. We could show that, for every $i \geq 1$, there is a language L such that L can be generated by an fRCG but cannot by pgSCGs with degree $(i, 0)$.

The paper is organized as follows. First, we introduce the necessary notions and notations. Then, in Section 3 we present the results of the paper. Finally, we give some concluding remarks in Section 4.

2. Preliminaries

We define here the necessary notions, however we assume that the reader is familiar with the basic concepts of the theory of formal languages. For a comprehensive guide we refer to [19]. An *alphabet* Σ is a finite, nonempty set of symbols whose elements are also called *letters*. *Words* over Σ are finite sequences of letters in Σ. As usual, Σ^* denotes the set of all words over Σ including the *empty word* ε. For a word $u \in \Sigma^*$, $|u|$ denotes the *length of* u. $\mathbb{N}$ denotes the set of natural numbers. For $n, m \in \mathbb{N}$, $n < m$, $[n, m]$ denotes the set $\{n, n + 1, \ldots, m\}$. If $n = 1$, then $[n, m]$ is denoted by $[m]$. The *set of positions in* u ($pos(u)$ for short) is $[|u|]$.

Let $u \in \Sigma^*$. A word v is a *scattered subword* of u, if v can be obtained from u by erasing some (possibly zero) letters. Moreover, v is a *subword* of u if there are words $u_1, u_2 \in \Sigma^*$ such that $u = u_1 v u_2$. Let $i \in pos(u)$ and $m \geq 1$ be such that $i + m - 1 \in pos(u)$. Then $subw(u, i, m)$ denotes that subword of u which starts on the ith position and has length m. It will always be clear from the context whether we consider an arbitrary subword of u or that one which starts on a certain position. Those subwords of u that have length m are also called m-*subwords*. The *subsequence relation* $\leq_s$ over Σ^* is a binary relation defined as follows. For $u, v \in \Sigma^*$, $u \leq_s v$, if u is a scattered subword of v. Let $f : [k] \to [l]$ ($k, l \geq 1$) be a (partial) function. The *domain* and *range* of f, denoted by $dom(f)$ and $ran(f)$, respectively, are defined as follows: $dom(f) = \{i \in [k] \mid \exists j \in [l] : f(i) = j\}$ and $ran(f) = \{i \in [l] \mid \exists j \in [k] : f(j) = i\}$. If $I \subsetneq [k]$, then $f|_I$ denotes the restriction of f to I. Let $u, v \in \Sigma^*$ and $f : pos(v) \to pos(u)$ be a (partial) function. If, for every $i \in dom(f)$, $subw(v, i, 1) = subw(u, f(i), 1)$, then we call f *letter-preserving*. A *well-quasi-ordering* (wqo for short) on a set S is a reflexive, transitive binary relation $\leq$ such that any infinite sequence $a_1, a_2, \ldots$ ($a_i \in S, i \geq 1$) contains a pair $a_j \leq a_k$ with $j < k$. The following result is due to [11] (see also [14]).

Proposition 1. *Let Σ be an alphabet. Then $\leq_s$ is a wqo on Σ^*. Consequently, for every infinite sequence $u_1, u_2, \ldots$ ($u_i \in \Sigma^*, i \geq 1$), there is an infinite subsequence $u_{i_1} \leq_s u_{i_2} \leq_s \cdots$.*

A *generalized semi-conditional grammar* (*gSCG* for short) is a 4-tuple $G = (V, \Sigma, R, S)$, where V and Σ are alphabets of the *nonterminal* and *terminal* symbols, respectively (it is assumed that $V \cap \Sigma = \emptyset$), $S \in V$ is the *start symbol*, and R is a finite set of *production rules* of the form $(A \to \alpha, P, Q)$, where $A \in V, \alpha \in (V \cup \Sigma)^+$ (that is $A \to \alpha$ is a usual non-erasing context-free rule), and P and Q are finite

disjoint sets of words in $(V \cup \Sigma)^+$. For a rule $r = (A \to \alpha, P, Q)$, P and Q are called the *permitting* and *forbidding context* of r, respectively. The *right-hand side* of r (denoted by $\mathrm{rhs}(r)$) is α. We will often denote $V \cup \Sigma$ by V_G. The *derivation relation* $\Rightarrow_G$ of G is defined as follows. For every word $u_1, u_2, \alpha \in V_G^*$ and $A \in V$, $u_1 A u_2 \Rightarrow_G u_1 \alpha u_2$ if and only if there is a rule $(A \to \alpha, P, Q) \in R$ such that (i) every word in P is a subword of $u_1 A u_2$, and (ii) no word in Q is a subword of $u_1 A u_2$. We will often write $\Rightarrow$ instead of $\Rightarrow_G$ when G is clear from the context. As usual, the *reflexive, transitive closure of* $\Rightarrow$ is denoted by $\Rightarrow^*$ and the *language generated by* G is $L(G) = \{u \in \Sigma^* \mid S \Rightarrow^* u\}$. A word $\alpha \in V_G^*$ is often called a *sentential form of* G (or just a *sentential form*) if $S \Rightarrow^* \alpha$.

Example 2. (Cf. Example 1.1.7 in [3]) *Consider the gSCG* $G = (\{S, A, B, D\}, \{a\}, \{r_1, r_2, \ldots, r_5\}, S)$, *where* $r_1 = (S \to AA, \emptyset, \{B, D\})$, $r_2 = (A \to B, \emptyset, \{S, D\})$, $r_3 = (B \to S, \emptyset, \{A, D\})$, $r_4 = (A \to D, \emptyset, \{S, B\})$, *and* $r_5 = (D \to a, \emptyset, \{S, A, B\})$.

Notice that every rule in R *has an empty permitting context and that the forbidding contexts are subsets of the set of nonterminal symbols. It can be seen moreover that* $L(G) = \{a^{2^n} \mid n \geq 1\}$. *Indeed, consider the word* S^i *for some* $i \geq 1$. *To this word we have to apply* r_1 *as long as* S *occurs in the sentential form (rules with* A *on the left-hand side are forbidden to use by* S). *Thus we get* A^{2i}. *Now we can apply only* r_2 *or* r_4. *If we apply* r_2, *then we should apply it as long as we get* B^{2i}. *To this word we can apply only* r_3 *and we should apply this rule until we get* S^{2i}. *On the other hand, if we apply* r_4 *to* A^{2i}, *then we should apply this rule until we get* D^{2i}. *Now we can apply only* r_5 *until we get* a^{2i}.

Let $G = (V, \Sigma, R, S)$ be a gSCG. A *derivation der* from α to β in G is a sequence $\alpha_1 \Rightarrow_G \alpha_2 \Rightarrow_G \cdots \Rightarrow_G \alpha_{n+1}$ of words in V_G^* for some $n \geq 0$ such that $\alpha_1 = \alpha$ and $\alpha_{n+1} = \beta$. The *length of der* (denoted by $|der|$) is n. Let $\alpha, \beta \in V_G^*$ and *der* be a derivation from α to β. The *sentential form vector of der* (denoted by $\mathrm{sfv}(der)$) is $(u_1, \ldots, u_k)$ $(k = |\alpha|, u_i \in V_G^*, i \in [k])$, such that $\beta = u_1 \cdots u_k$ and, for every $i \in [k]$, u_i is derived from $\mathrm{subw}(\alpha, i, 1)$. Let $der : \alpha_1 \Rightarrow \cdots \Rightarrow \alpha_n$ $(n \geq 1)$ and $der' : \alpha_n \Rightarrow \cdots \Rightarrow \alpha_k$ $(k \geq n)$. Then $der\, der'$ denotes the derivation $\alpha_1 \Rightarrow \cdots \Rightarrow \alpha_k$.

Let $i, j \geq 0$. G is of *degree* (i, j) if, for every rule $(p, P, Q) \in R$, P (resp. Q) contains only words with length at most i (resp. j). For $i \geq 0$, gSCGs with degree $(i, 0)$ are also called *permitting gSCGs*, or *pgSCGs* for short. Let $G = (V, \Sigma, R, S)$ be a pgSCG. The *set of permitting contexts of* G is $pw(G) = \bigcup_{(p,P,Q) \in R} P$ and $\max_{pw(G)} = \max\{|u| \mid u \in pw(G)\}$.

A *random context grammar* (RCG) is a semi-conditional grammar, where the sets of permitting and forbidding contexts are subsets of the nonterminal alphabet. An RCG $G = (V, \Sigma, R, S)$ is a *forbidding* RCG (fRCG) if, for every rule $(p, P, Q) \in R$, P is empty. Notice that the grammar G in Example 2 is an fRCG.

We denote by $\mathcal{L}(\mathrm{pgSCG})$, $\mathcal{L}(\mathrm{fRCG})$ and $\mathcal{L}(\mathrm{CS})$ the families of languages generated by pgSCGs, fRCGs and context-sensitive grammars, respectively. Moreover, for a number $i \geq 0$, $\mathcal{L}(\mathrm{pgSCG}_i)$ denotes the class of languages generated by pgSCGs

6

of degree $(i, 0)$.

3. The Main Results

Here we show first that pgSCGs are strictly weaker than context sensitive grammars by proving that the language $L = \{a^{2^{2^n}} \mid n \geq 0\}$ cannot be generated by any pgSCG (Theorem 14). We show this by using a pumping lemma (Lemma 13). The proof of this lemma consists of the following main steps.

(1) First we define the notion of m-embedding (Definition 3). Intuitively, a word α can be m-embedded to a word β, if there is an injective mapping of the m-subwords of α to the m-subwords of β such that this mapping preserves the order of these words and satisfies certain additional conditions.
(2) Then we show that if a pgSCG $G = (V, \Sigma, R, S)$ has a derivation of the form $\alpha \Rightarrow^* \beta \Rightarrow^* \gamma$ where $\alpha, \beta \in V_G^*, \gamma \in \Sigma^*, |\alpha| < |\beta|$ and α can be m-embedded to β, then this derivation can be "extended" into a derivation $\alpha \Rightarrow^* \gamma'$ for some $\gamma' \in \Sigma^*$ with $|\gamma| < |\gamma'| \leq (m+1)|\gamma|$ (Lemma 9).
(3) Finally, we show that, for any $m \geq 1$, sufficiently long derivations of a pgSCG G always contain sentential forms α and β such that α can be m-embedded to β (Lemma 12). To this end we will use the fact that $\leq_s$ is a wqo on V_G^*.

Definition 3. *Let Σ be an alphabet, $\alpha, \beta \in \Sigma^*$, $k = |\alpha|$, $l = |\beta|$, and $m \geq 1$. An m-embedding of α to β is a strictly increasing function $g : [k - m + 1] \to [l]$ such that the following (partial) mapping $f : \mathrm{pos}(\beta) \to \mathrm{pos}(\alpha)$ is letter-preserving and well defined: for every $i \in [k - m + 1]$ and $\kappa \in [0, m - 1]$, $f(g(i) + \kappa) = i + \kappa$. If g is an m-embedding, then the above f is denoted by $\mathrm{inv}_m(g)$. Moreover, if an m-embedding of α to β exists, then we denote this by $\alpha \rightsquigarrow_m \beta$.*

Example 4. *Here we give two examples to demonstrate the notion of m-embedding (see also Fig. 2).*

(1) Let $\alpha = BAAB$ and $\beta = BAAAB$. Any 3-subword of α is a subword of β, too. Due to the letter-preserving property, only the following g can be a 3-embedding of α to β: $g(1) = 1$ and $g(2) = 3$. The mapping $f = \mathrm{inv}_m(g)$ is letter-preserving, but not well defined. Indeed, with $i = 1$ and $\kappa = 2$ we get $f(g(i) + \kappa) = f(3) = 3$, while with $i = 2$ and $\kappa = 0$, $f(g(i) + \kappa) = f(3) = 2$. This implies that there is no 3-embedding of α to β.
(2) Let $\alpha = ABBAC$, $\beta = AABBAABAC$ and g be the following strictly increasing function: $g(1) = 2, g(2) = 3$, and $g(3) = 7$. Then the mapping $f = \mathrm{inv}_m(g)$ is letter-preserving and well defined: $f(i) = i - 1$ $(i \in [2, 5])$ and $f(i) = i - 4$ $(i \in [7, 9])$. Thus g is a 3-embedding of α to β.

The following properties of m-embeddings will be useful in what follows.

Proposition 5. *Let Σ be an alphabet, $m \geq 1$, and $\alpha, \beta \in \Sigma^*$. Assume that g is an m-embedding of α to β and $f = \mathrm{inv}_m(g)$. Then the following statements hold.*
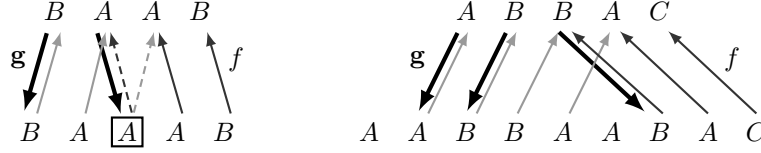
Fig. 2. A visualization of the mappings used in Example 4.

(i) *For every $i \in \mathrm{pos}(\alpha)$, $|\{t \in \mathrm{pos}(\beta) \mid f(t) = i\}| \leq m$.*
(ii) *If $|\alpha| = |\beta|$, then $\alpha = \beta$.*
(iii) *If $i, j \in pos(\alpha)$ with $j = i + 1$, then $g(j) - g(i) = 1$ or $g(j) - g(i) \geq m$.*
(iv) *If $s, t \in pos(\beta)$ with $s < t$ and $f(s) = f(t)$, then $t - s \geq m - 1$.*

Proof. We show (i) first. Let $i \in pos(\alpha)$ and $J = \{j \in \mathrm{pos}(\alpha) \mid i \in [j, j + m - 1]\}$. Clearly, $|J| \leq m$. Let $S = ran(g|_J)$. Since g is strictly increasing, $|S| = |J|$. Moreover, using that for every $s \in S$, $f|_{[s,s+m-1]}$ is a strictly increasing function to the set $[f(s), f(s) + m - 1]$, we get that there is exactly one $t \in [s, s + m - 1]$ with $f(t) = i$. It is also clear that for every $t \in \mathrm{pos}(\beta)$ with $f(t) = i$ there is an $s \in S$ with $t \in [s, s + m - 1]$. Thus, $|\{t \in \mathrm{pos}(\beta) \mid f(t) = i\}| \leq |S| = |J| \leq m$.

To see (ii) it is enough to observe that since g is a strictly increasing function, it should map every $i \in [k - m + 1]$ to i ($k = |\alpha|$). Thus $f(i) = i$ for every $i \in [k - m + 1]$, and the statement follows using that f is letter preserving.

To prove (iii) let $d = g(i + 1) - g(i)$. If $d \leq m - 1$ then $f(g(i) + d) = f(g(i + 1)) = i + 1$ and $f(g(i) + d) = i + d$ (by the definition of f for $i + 1$ and $\kappa = 0$, and for i and $\kappa = d$, respectively). Since f must be a (partial) function, we get that in this case $d = 1$.

To prove (iv) assume that $f(s) = f(t)$ for some $s, t \in \mathrm{pos}(\beta)$ and $s < t$. Let $i, j \in [k - m + 1]$ ($k = |\alpha|$) and $\mu, \nu \in [0, m - 1]$ be such that $s = g(i) + \mu$ and $t = g(j) + \nu$. Then $f(s) = i + \mu$ and $f(t) = j + \nu$ by the definition of f. Now, $s \notin [g(j), t]$, otherwise $f(s) = f(t)$ contradicts the definition of an m-embedding (notice that $f|_{[g(j),g(j)+m-1]}$ is an injective function and $[g(j), t] \subseteq [g(j), g(j) + m - 1]$). This proves $g(i) < g(j)$, and thus $i < j$ by the strictly increasing property of g. According to (iii) $g(j) - g(i) = j - i + d$ where either $d = 0$ or $d \geq m - 1$. Thus $(t - \nu) - (s - \mu) = g(j) - g(i) = j - i + d = (f(t) - \nu) - (f(s) - \mu) + d = -\nu + \mu + d$. Consequently, $t - s = d$. Since $s < t$, only $d \geq m - 1$ is possible, which finishes the proof of (iv). $\square$

We will also need the following operation which inserts words into certain positions of a word. Let Σ be an alphabet and $\alpha = X_1 \cdots X_k$ ($k \geq 1, X_i \in \Sigma, i \in [k]$). Let moreover $u_1, \ldots, u_l \in \Sigma^*$ and $f : [k] \to [l]$ be a (partial) function. The *substitution of* $\mathbf{u} = (u_1, \ldots, u_l)$ *into* α *by* f (denoted by $\mathrm{subst}(\mathbf{u}, \alpha, f)$) is the word $\beta = v_1 \cdots v_k$, where v_i ($i \in [k]$) is defined as follows. If $f(i)$ is defined, then let $v_i = u_{f(i)}$, and let $v_i = X_i$ otherwise.

8
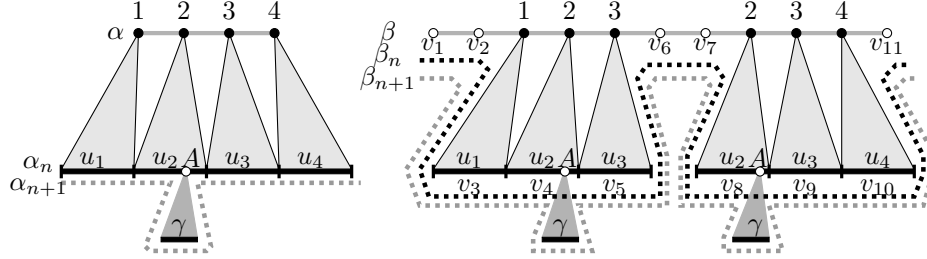
Fig. 3. The inductive proof of Lemma 7 assuming $m = 3$ and $|P| = 1$.

Sometimes we will need to extend a function f used in a substitution. An *extension of f* (with respect to α) is a function $\hat{f}$ defined as follows. For every $i \in dom(f)$, $\hat{f}(i) = f(i)$, and for every $i \in [k] - dom(f)$, $\hat{f}$ is either undefined or defined as follows: if there is a $j \in dom(f)$ such that $\mathrm{subw}(\alpha, i, 1) = \mathrm{subw}(\alpha, j, 1)$, then take such a j and let $\hat{f}(i) = f(j)$. Notice that f is always an extension of itself.

Example 6. *Let $\Sigma = \{A, B, C\}$, $\alpha = ABCBB$ and $\mathbf{u} = (u_1, u_2, u_3)$, where $u_1 = AA, u_2 = ABC, u_3 = CC$. Let furthermore $f : [5] \to [3]$ be the following partial function. $f(2) = f(3) = 3$, $f(5) = 1$. Then $\mathrm{subst}(\mathbf{u}, \alpha, f) = Au_3u_3Bu_1 = ACCCCBAA$ and f has two possible extensions other than f. $\hat{f}(1)$ is undefined. $\hat{f}(2) = \hat{f}(3) = 3$, $\hat{f}(5) = 1$ and $\hat{f}(4)$ is either 1 or 3 resulting in $\mathrm{subst}(\mathbf{u}, \alpha, \hat{f})$ equal to AC^4A^4 or AC^6A^2, respectively.*

The following result will be crucial in the proof of Lemma 9.

Lemma 7. *Let $G = (V, \Sigma, R, S)$ be a pgSCG, $m = 2 \cdot \max_{pw(G)}$, and $\alpha, \alpha', \beta \in V_G^*$. Assume that $\alpha \Rightarrow^* \alpha'$ and $\alpha \rightsquigarrow_m \beta$. Let der be a derivation from α to α', g an m-embedding of α to β, and $f = \mathrm{inv}_m(g)$. Then, for every extension $\hat{f}$ of f, $\beta \Rightarrow^* \beta_{\hat{f}}$, where $\beta_{\hat{f}} = \mathrm{subst}(\mathrm{sfv}(der), \beta, \hat{f})$.*

Proof. Let $\beta' = \mathrm{subst}(\mathrm{sfv}(der), \beta, f)$. We first show that $\beta \Rightarrow^* \beta'$ by induction on $n = |der|$. If $n = 0$, then one can see that $\beta' = \beta$, and thus the statement trivially holds. Assume that it holds for n. We prove it for $n+1$. In this case der can be written as $der = der_1 der_2$, where der_1 is $\alpha_0 \Rightarrow \cdots \Rightarrow \alpha_n$, der_2 is $\alpha_n \Rightarrow \alpha_{n+1}$, $\alpha_0 = \alpha$, and $\alpha_{n+1} = \alpha'$. Let $\beta_n = \mathrm{subst}(\mathrm{sfv}(der_1), \beta, f)$. By the induction hypothesis, there is a derivation der_1' from β to β_n. Let $(u_1, \ldots, u_k) = \mathrm{sfv}(der_1)$ $(k = |\alpha|)$. Assume that G rewrites a nonterminal A during $\alpha_n \Rightarrow \alpha_{n+1}$ using a rule $r = (A \to \gamma, P, \emptyset)$ (see Fig. 3 for an example).

Let $i \in [k]$ and $\kappa \in \mathrm{pos}(u_i)$ be such that the rewritten A occurs on the κth position of u_i. Let $i_1 < i_2 \cdots < i_\xi$ be all the positions in $\mathrm{pos}(\beta)$ with $f(i_j) = i$ $(j \in [\xi])$. Let $(v_1, \ldots, v_l) = \mathrm{sfv}(der_1')$ $(l = |\beta|)$. Then, for every $j \in [\xi]$, $v_{i_j} = u_i$ and thus, for every such j, there is a position $\kappa_j \in \mathrm{pos}(\beta_n)$ satisfying that κ_j corresponds to the κth position in v_{i_j}. Clearly $\mathrm{subw}(\beta_n, \kappa_j, 1) = A$ and $\beta' =$

$\beta_{n+1} = \text{subst}((\gamma), \beta_n, h)$ where $h : \text{pos}(\beta_n) \to \{1\}$ is defined as follows: $h(j) = 1$ if $j \in \{\kappa_1, \ldots, \kappa_\xi\}$, and it is undefined otherwise. Therefore, to prove $\beta_n \Rightarrow^* \beta'$ it is enough to show that G can use r to rewrite each nonterminal A that occurs on a position κ_j ($j \in [\xi]$) in β_n.

Let $m' = \max_{pw(G)}$, then $m = 2m'$. Since G can apply r at the step $\alpha_n \Rightarrow \alpha_{n+1}$, α_n should contain all permitting contexts of P. Then, for every $p \in P$, let $\mu(p) \in [k]$ and $\nu(p) \in [0, m' - 1]$ such that p occurs in the subword $u_{\mu(p)} \cdots u_{\mu(p)+\nu(p)}$ of α_n (notice that G has no erasing rules). Let moreover $P_0 = \{p \in P \mid i \in [\mu(p), \mu(p) + \nu(p)]\}$ and $P_1 = P \setminus P_0$.

First, consider the permitting contexts of P_1. For every $p \in P_1$, p is a subword of $u_{\mu(p)} \cdots u_{\mu(p)+\nu(p)}$ and this word avoids u_i. Since g is an m-embedding of α to β, $v_{g(\mu(p))} \cdots v_{g(\mu(p))+\nu(p)} = u_{\mu(p)} \cdots u_{\mu(p)+\nu(p)}$. Furthermore it contains p and does not contain the positions κ_j ($j \in [\xi]$). So the contexts of P_1 are subwords of the sentential form after rewriting even all the A's at the positions κ_j ($j \in [\xi]$).

Now, consider the permitting contexts of P_0. It can be seen that, for all $p \in P_0$, $u_{\mu(p)} \cdots u_{\mu(p)+\nu(p)}$ is a subword of $u_{max(1,i-m'+1)} \cdots u_{min(k,i+m'-1)}$. Since $|[max(1, i - m' + 1), min(k, i + m' - 1)]| \leq m - 1$, it is a subword of $u_x \cdots u_{x+m-2}$ for some $x \in [k]$. Since g is an m-embedding of α to β, it is clear that $w = v_{g(x)} \cdots v_{g(x)+m-2} = u_x \cdots u_{x+m-2}$. Moreover, w contains all contexts of P_0 and by (iv) of Proposition 5 there is exactly one index $j \in [\xi]$ such that w contains A on the κ_jth position in β_n. Then G should rewrite first those A's in β_n that occur on positions other than κ_j and, at the last step, that A which occurs on the κ_jth position. Therefore $\beta_n \Rightarrow^* \beta_{n+1} = \beta'$ which implies that $\beta \Rightarrow^* \beta'$.

To finish the proof of the lemma consider a derivation der' from β to β'. Looking at the inductive proof of $\beta \Rightarrow^* \beta'$, one can see that, the letters in β that are on such positions which are not included in $dom(f)$ do not occur in the permitting contexts used during der'. Assume that $i \in \text{pos}(\beta) - dom(f)$ such that $\hat{f}(i) = f(j)$, for some $j \in dom(f)$. Let u be the $f(j)$th word in $\text{sfv}(der)$ and $X = \text{subw}(\beta, j, 1)$. Then G derives u during der' from this X. On the other hand, by the definition of $\hat{f}$, $\text{subw}(\beta, i, 1) = X$. Thus, der' can be extended to such a derivation where G, using the appropriate rules simultaneously, derives u also from that X which occurs on the ith position of β. Following this way of thinking one can see that der' can be extended to a derivation of $\beta_{\hat{f}}$ from β which completes the proof of the lemma. $\qquad \square$

Let us consider now the pgSCG G and β, β_f from the previous lemma. We have seen that $\beta \Rightarrow^*_G \beta_f$. In the next proposition we will see that $\beta \rightsquigarrow_m \beta_f$ also holds. This result will be important in the proof of Lemma 9. However, first we need to introduce some more concepts.

Let $g : \alpha \rightsquigarrow_m \beta$ be an m-embedding. The mapping

$$\text{cmp}_g(i) = \begin{cases} g(i) & \text{for } i \in [k - m + 1] \\ g(k - m + 1) + i - (k - m + 1) & \text{for } i \in [k - m + 2, k] \end{cases}$$

10

is called the *completion* of g. Note that $dom(\mathrm{cmp}_g) = pos(\alpha)$ and by the definition of an m-embedding cmp_g is letter-preserving. If $f = \mathrm{inv}_m(g)$ then $f(\mathrm{cmp}_g(i)) = i$ holds for $i \in [k]$.

Proposition 8. *Let $\alpha \in \Sigma^*, |\alpha| = k, z_i \in \Sigma^*, z_i \neq \varepsilon \ (i \in [k])$ and $\beta = z_1 \cdots z_k$ with $|\beta| = l$. Let $m \geq 1$ and suppose that $g : \alpha \rightsquigarrow_m \beta$ with $f = \mathrm{inv}_m(g)$ and $\bar{g} = \mathrm{cmp}_g$. Then*

$$g' : \beta \rightsquigarrow_m \mathrm{subst}(\mathbf{u}, \beta, f),$$

where $\mathbf{u} = (z_1, \ldots, z_k)$ and g' is defined as follows. Let

- $x_i = \begin{cases} z_{f(i)} & \text{if } i \in dom(f) \\ \mathrm{subw}(\beta, i, 1) & \text{if } i \notin dom(f) \end{cases}$,
- $\zeta(i, r) = \sum_{j=1}^{i-1} |z_j| + r$, *where $i \in [k]$ and $r \in \mathrm{pos}(z_i)$, and*
- $\xi(i, r) = \sum_{j=1}^{i-1} |x_j| + r$, *where $i \in [l]$ and $r \in \mathrm{pos}(x_i)$.*

Then $g'(\zeta(i, r)) := \xi(\bar{g}(i), r) \ (\zeta(i, r) \in [l - m + 1])$.

Proof. First, observe that $x_1 \cdots x_l = \mathrm{subst}(\mathbf{u}, \beta, f)$ holds by the definition of x_i's and of a substitution. Let us denote this word by β'. Let $f' : \mathrm{pos}(\beta') \to \mathrm{pos}(\beta)$ be the following (partial) function: for every $i \in [l - m + 1]$ and $\kappa \in [0, m - 1]$, $f'(g'(i) + \kappa) := i + \kappa$. g' is strictly increasing, so the following are left to prove: (†) f' is letter-preserving and (‡) f' is well-defined.

To prove (†) it is enough to show that $\mathrm{subw}(\beta, \tau, m) = \mathrm{subw}(\beta', g'(\tau), m)$ holds for $\tau = \zeta(i, r) \in [l - m + 1]$, $i \in [k]$, $r \in \mathrm{pos}(z_i)$. Suppose, that $\tau + m - 1 = \zeta(j, r')$ for some $j \in [k]$ and $r' \in \mathrm{pos}(z_j)$, then $\mathrm{subw}(\beta, \tau, m) = \mathrm{subw}(z_i \cdots z_j, r, m)$. By the assumptions, $|z_p| \geq 1$ holds for all $p \in [k]$, which implies $j - i \leq m - 1$.

Let $\Delta = \Sigma \cup Z$ be an alphabet, where $Z = \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$ is a set of k symbols, satisfying $\Sigma \cap Z = \emptyset$. Let $\mathbf{z} = \mathbf{z}_1 \cdots \mathbf{z}_k$ and $\mathbf{x} = \mathbf{x}_1 \cdots \mathbf{x}_l$, where $\mathbf{x}_i = \mathbf{z}_{f(i)}$, if $f(i)$ is defined, and $\mathbf{x}_i = x_i$, otherwise (that is, $\mathbf{z}$ and $\mathbf{x}$ are words of length k and l, respectively, over Δ). Then $\mathbf{z} \rightsquigarrow_m \mathbf{x}$ by the same m-embedding g, which implies $\mathbf{z}_i \cdots \mathbf{z}_j = \mathbf{x}_{\bar{g}(i)} \cdots \mathbf{x}_{\bar{g}(i)+j-i}$ for $1 \leq i \leq j \leq k$ and $j - i \leq m - 1$ over the alphabet Δ.

It follows that the analogous equation $z_i \cdots z_j = x_{\bar{g}(i)} \cdots x_{\bar{g}(i)+j-i}$ over the alphabet Σ holds, too. So we have $\mathrm{subw}(\beta, \tau, m) = \mathrm{subw}(z_i \cdots z_j, r, m) = \mathrm{subw}(x_{\bar{g}(i)} \cdots x_{\bar{g}(i)+j-i}, r, m) = \mathrm{subw}(\beta', \xi(\bar{g}(i), r), m) = \mathrm{subw}(\beta', g'(\tau), m)$ by the definitions proving (†).

It remains to show (‡). To this end, for every $\tau < \tau' \in \mathrm{pos}(\beta)$, let $c(\tau, \tau') = (g'(\tau') - g'(\tau)) - (\tau' - \tau)$. Notice that $c(\tau, \tau') \geq 0$ due to the fact that g' is strictly increasing. Moreover, $c(\tau, \tau')$ has the following property (⋆) $c(\tau, \tau') = \sum_{j=\tau}^{\tau'-1} c(j, j+1)$. Indeed, $c(\tau, \tau') = (g'(\tau') - g'(\tau)) - (\tau' - \tau) = \sum_{j=\tau}^{\tau'-1} (g'(j+1) - g'(j)) - (\tau' - \tau) = \sum_{j=\tau}^{\tau'-1} c(j, j+1)$. We claim, that (⋆⋆) $c(\tau, \tau') \notin [1, m - 2]$, for any $\tau < \tau' \in \mathrm{pos}(\beta)$. Due to (⋆) it is enough to prove (⋆⋆) for $\tau' = \tau + 1$, since if one of the $c(j, j+1)$'s

is at least $m + 1$, then $c(\tau, \tau')$ should be at least $m + 1$ as well. Thus, we need to show that either $g'(\tau + 1) - g'(\tau) = 1$ or $g'(\tau + 1) - g'(\tau) \geq m$. Let $\tau = \zeta(i, r)$ and $\tau + 1 = \zeta(i', r')$ for some $i, i' \in [k], r \in \mathrm{pos}(z_i), r' \in \mathrm{pos}(z_{i'})$. Then we have the following two cases.

<u>Case 1</u>: $i' = i$ and $r' = r + 1$. Then $g'(\tau + 1) - g'(\tau) = \xi(\bar{g}(i), r + 1) - \xi(\bar{g}(i), r) = 1$.

<u>Case 2</u>: $i' = i + 1$, $r = |z_i|$ and $r' = 1$. By (iii) of Proposition 5, $\bar{g}(i + 1) - \bar{g}(i) \notin [2, m - 1]$ holds, since $\bar{g}|_{[k-m]} = g$ and $\bar{g}(i + 1) - \bar{g}(i) = 1$ for $i \in [k - m + 1, k - 1]$. Therefore, we need to discuss the following two sub-cases.

    <u>Case 2a:</u> $\bar{g}(i+1) = \bar{g}(i) + 1$. Then $g'(\tau+1) - g'(\tau) = \xi(\bar{g}(i) + 1, 1) - \xi(\bar{g}(i), |z_i|) = 1$.

    <u>Case 2b:</u> $\bar{g}(i+1) - \bar{g}(i) \geq m$. Then $g'(\tau+1) - g'(\tau) = \xi(\bar{g}(i+1), 1) - \xi(\bar{g}(i), |z_i|) \geq \xi(\bar{g}(i) + m, 1) - \xi(\bar{g}(i), |z_i|) = \sum_{j=1}^{m-1} \left(\xi(\bar{g}(i) + j + 1, 1) - \xi(\bar{g}(i) + j, 1) \right) + \xi(\bar{g}(i) + 1, 1) - \xi(\bar{g}(i), |z_i|) \geq (m - 1)1 + 1 = m$.

To finish the proof of (‡), assume that $\mu = g'(\tau) + \kappa = g'(\tau') + \kappa'$ holds for some $\tau, \tau' \in [l], \tau < \tau'$, and $\kappa, \kappa' \in [0, m - 1]$. Then $f'(\mu) = \tau + \kappa = \tau' + \kappa'$ should hold for f' being well-defined. $1 \leq \tau' - \tau = g'(\tau') - g'(\tau) - c(\tau, \tau') = \kappa - \kappa' - c(\tau, \tau') \leq m - 1 - c(\tau, \tau')$. (⋆⋆) implies $c(\tau, \tau') = 0$, i.e., $\tau' - \tau = \kappa - \kappa'$ should hold. This proves (‡) and thus we have shown that g' is an m-embedding of β to β' with $f' = \mathrm{inv}_m(g')$. $\qquad\square$

Lemma 9. *Let $G = (V, \Sigma, R, S)$ be a pgSCG and $m = 2 \cdot \max_{pw(G)}$. Suppose that $\alpha \Rightarrow^* \beta$, $\beta \Rightarrow^* \gamma$, $\alpha \rightsquigarrow_m \beta$, and $|\alpha| < |\beta|$ hold for some $\alpha, \beta, \gamma \in V_G^*$. Then there exists a $\gamma' \in \Sigma^*$ such that (i) $\alpha \Rightarrow^* \gamma'$ and (ii) $|\gamma| < |\gamma'| \leq (m + 1)|\gamma|$.*

Proof. Let $k = |\alpha|$, $l = |\beta|$, and g be an m-embedding of α to β with $f = \mathrm{inv}_m(g)$. Let moreover der' and der'' be any derivations from α to β and from β to γ, respectively. Let $\mathbf{u} = \mathrm{sfv}(der')$ and $\beta' = subst(\mathbf{u}, \beta, f)$. By Lemma 7 it holds that $\beta \Rightarrow^* \beta'$. Moreover, applying Proposition 8 with the above parameters and $\Sigma = V_G$, we get $g' : \beta \rightsquigarrow_m \beta'$. Let $f' = \mathrm{inv}_m(g')$ and x_i $(i \in [l])$, $\zeta(i, r)$ $(i \in [k], r \in pos(z_i))$, and $\xi(i, r)$ $(i \in [l], r \in pos(x_i))$ be as defined in Proposition 8 (recall that $\mathbf{u} = (z_1, \ldots, z_k)$).

    Let $\hat{f}'$ be the following function. For every $\tau \in pos(\beta')$, if $\tau \in dom(f')$, then let $\hat{f}'(\tau) = f'(\tau)$. Otherwise let $\tau = \xi(i, r)$, for some $i \in [l]$ and $r \in \mathrm{pos}(x_i)$, and we define $\hat{f}'(\tau)$ as follows. If $i \in dom(f)$, then let $\hat{f}'(\tau) = \zeta(f(i), r)$, and let $\hat{f}'(\tau) = i$, otherwise. Notice that $\hat{f}'$ is a letter preserving function form β' to β. Indeed, if $i \in dom(f)$, then $x_i = z_{f(i)}$, and $x_i = \mathrm{subw}(\beta, i, 1)$, otherwise. Let $\tau \in pos(\beta') - dom(f')$. Since g' is an m-embedding of β to β', there is a $\tau' \in pos(\beta')$ with $f'(\tau') = \hat{f}'(\tau)$. Then $\mathrm{subw}(\beta', \tau, 1) = \mathrm{subw}(\beta, \hat{f}'(\tau), 1) = \mathrm{subw}(\beta, f'(\tau'), 1) = \mathrm{subw}(\beta', \tau', 1)$. Thus $\hat{f}'$ is an extension of f'.

    Now let $\gamma' = subst((v_1, \ldots, v_l), \beta', \hat{f}')$, where $(v_1, \ldots, v_l) = \mathrm{sfv}(der'')$. By Lemma 7, $\beta' \Rightarrow^* \gamma'$. This, together with $\alpha \Rightarrow^* \beta$ and $\beta \Rightarrow^* \beta'$ implies $\alpha \Rightarrow^* \gamma'$, i.e., Statement (i) of the lemma holds. Statement (ii) can be seen as follows. Since g' is an m-embedding of β to β', for every $i \in [l]$, there is a $\tau \in pos(\beta')$ with

12

$f'(\tau) = i$. Thus, each v_i $(i \in [l])$ is substituted for a position in β' by f'. Therefore $|\gamma| = |v_1 \cdots v_l| \leq |\text{subst}((v_1, \ldots, v_l), \beta', f')| \leq |\text{subst}((v_1, \ldots, v_l), \beta', \hat{f}')| = |\gamma'|$. Moreover, since $|\alpha| < |\beta|$, there is an $i \in [k]$ such that $|z_i| \geq 2$. Let $j \in [l]$ with $f(j) = i$. Then $|x_j| \geq 2$, so $|\beta| = l < l + 1 \leq \sum_{s=1}^{l} |x_s| = |\beta'|$. This implies that $|\gamma| = |\gamma'|$ cannot hold, consequently $|\gamma| < |\gamma'|$.

On the other hand, by (i) of Proposition 5, for every $i \in [k]$, z_i is substituted for at most m different positions in β by f. Moreover, one can see that, for every $i \in dom(f)$, $\hat{f}'$ is an injective function from $[\xi(i, 1), \xi(i, |x_i|)]$ to $[l]$. Furthermore, $\hat{f}'$ is injective on the set $\{\tau \mid \tau = \xi(i, 1), i \in [l] - dom(f)\}$, too. Consequently, for every $i \in [l]$, v_i is substituted for at most $m + 1$ different positions in β' by $\hat{f}'$. Therefore, $|\gamma'| \leq (m + 1)|\gamma|$ should hold finishing the proof of Statement (ii). $\square$

Next we demonstrate some of the constructions used in the previous proof.

Example 10. *Let $G = (V, \Sigma, R, S)$ be a pgSCG, $\alpha = ABBA$, $\beta = EABBFBACA$ $(A, B, C, E, F \in V \cup \Sigma)$, and $m = 2$. Let $\gamma \in \Sigma^*$, and assume that G has two derivations der$'$ and der$''$ from α to β and from β to γ, respectively. Clearly $|\alpha| < |\beta|$ and $\alpha \leadsto_m \beta$ with the following m-embedding g: $g(1) = 2$, $g(2) = 3$, and $g(3) = 6$. Then, according to Lemma 9, we can give a $\gamma' \in \Sigma^*$ with the following properties: (i) $\alpha \Rightarrow^* \gamma'$ and (ii) $|\gamma| < |\gamma'| \leq (m + 1)|\gamma|$ (see also Fig. 4).*
Assume, for instance, that sfv(der$'$) $= (z_1, z_2, z_3, z_4)$, where $z_1 = E$, $z_2 = AB$, $z_3 = BF$, and $z_4 = BACA$. Let $f = \text{inv}_m(g)$. Then $\beta' = subst(\text{sfv}(der'), \beta, f) = x_1 \cdots x_9$, where $x_1 = E$, $x_2 = E$, $x_3 = AB$, $x_4 = BF$, $x_5 = F$, $x_6 = BF$, $x_7 = BACA$, $x_8 = C$, and $x_9 = A$. Now, if we define g' according to the proof of Proposition 8, then we get that $dom(g') = [1, 8]$, and $g'(i) = i + 1$ if $i \in [1, 3]$, and $g'(i) = i + 4$, otherwise. It is easy to verify that g' is an m-embedding of β to β'. Let $f' = \text{inv}_m(g')$. Then $dom(f') = [2, 5] \cup [8, 13]$. Let us define now $\hat{f}'$ according to the proof of Lemma 9, that is, $\hat{f}'(1) = 1$, $\hat{f}'(6) = \hat{f}'(7) = 5$, $\hat{f}'(14) = 8$, and $\hat{f}'(15) = 9$. One can check that, for every $\tau \in [1, 15] - dom(f')$, subw$(\beta', \tau, 1) = $ subw$(\beta, \hat{f}'(\tau), 1)$. Assume that sfv(der$''$) $= v_1 \cdots v_9$, where $v_i \in \Sigma^*$ $(i \in \text{pos}(\beta))$. Then $\gamma' = subst(\text{sfv}(der''), \beta', \hat{f}') = v_1 v_1 v_2 \cdots v_5 v_5 v_4 v_5 \cdots v_9 v_8 v_9$. By (i) of Lemma 9, $\alpha \Rightarrow^* \gamma'$, and it is easy to check that γ' satisfies Statement (ii) too.

The following proposition together with Lemma 12 will be used to show that sufficiently long derivations of pgSCGs always contain sentential forms α and β satisfying the conditions of Lemma 9.

Proposition 11. *Let Σ be an alphabet and let $n_1, n_2, \ldots$ be an infinite sequence of numbers in $\mathbb{N}$. Then there is $M \in \mathbb{N}$ such that, for every sequence $v_1, v_2, \ldots, v_n$ where $n \geq M$ and $v_i \in \Sigma^*$ with $|v_i| \leq n_i$ $(i \in [n])$, there are numbers $i < j$ in $[M]$ satisfying $v_i \leq_s v_j$.*

Proof. The proof is based on similar ideas as those used in the proof of Lemma 2 in [6]. Assume that there is no such M. It means that, for every $k \geq 1$, there is a
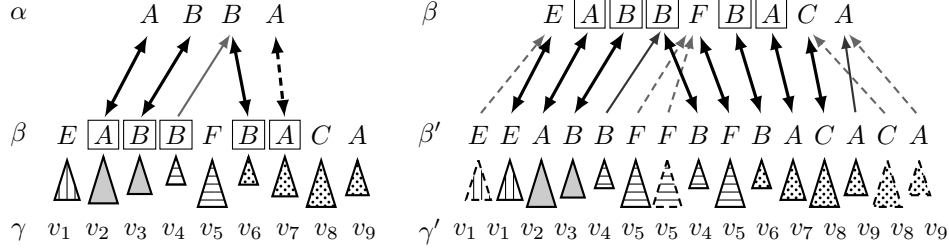
Fig. 4. A visualization of the mappings defined in Example 10. On the left solid arrows pointing down denote g and these arrows with the dashed one denote $\bar{g}$. Arrows pointing up denote f. On the right arrows pointing down denote g', solid arrows pointing up denote f', and (solid and dashed) arrows pointing up denote $\hat{f}'$. Letters in boxes denote the domain of f.

counterexample, a sequence $v_{k1}, \ldots, v_{kk}$ such that $|v_{ki}| \leq n_i$ $(i \in [k])$ and $v_{ki} \not\leq_s v_{kj}$ $(1 \leq i < j \leq k)$.

Since for every v_{k1} $(k \geq 1)$ $|v_{k1}| \leq n_1$ holds, there are only a finite number of possibilities for v_{k1}, and there is thus an infinite index set $\alpha_1 = \{k_1, k_2, \ldots\}$ such that $v_{k1} = v_{k'1}$ $(k, k' \in \alpha_1)$. Choose $u_1 = v_{k_1 1}$. Similarly, of the v_{k2}'s $(k \in \alpha_1)$ there is also an infinite set of indices α_2 such that $v_{k2} = v_{k'2}$ $(k,' k \in \alpha_2)$, choose $u_2 = v_{k2}$ $(k \in \alpha_2)$. Continuing in this manner we get an infinite sequence $u_1, u_2, \ldots$ satisfying $u_i \not\leq_s u_j$ for all $i < j$ which contradicts Proposition 1. $\qquad\square$

Let $G = (V, \Sigma, R, S)$ be a pgSCG and $m \geq 1$. We will apply the above result to appropriate derivations of G in order to find sentential forms α and β satisfying $\alpha \rightsquigarrow_m \beta$. However, Proposition 11 ensures only that we can find such α and β which satisfy $\alpha \leq_s \beta$. Clearly, this does not imply $\alpha \rightsquigarrow_m \beta$. Thus we will apply Proposition 11 not directly to the derivations of G but to sequences of words constructed from these derivations. To this end we will use two functions wdo and p defined below.

Let Σ be an alphabet and $m \geq 1$. We denote by $\Sigma^{\leq m}$ the set of all words in Σ^* with length at most m. Since $\Sigma^{\leq m}$ is a finite set, we will treat it as an alphabet. Now let wdo : $\Sigma^* \rightarrow (\Sigma^{\leq m})^*$ be defined as follows. Let $u \in \Sigma^*$. If $|u| < m$, then let wdo$(u) = u$ (that is, u on the right-hand side is considered as a letter in $\Sigma^{\leq m}$). If $|u| \geq m$, then let wdo$(u) = $ subw$(u, 1, m) \cdots$ subw$(u, |u| - m + 1, m)$ (again, subw(u, i, m) $(i \in [1, |u| - m + 1])$ is considered as a letter in $\Sigma^{\leq m}$). The name wdo comes from the word *window*, since for a word u, wdo(u) is that word whose letters are determined by moving a window of length m on u from left to right. The intuition behind the definition of wdo is the following: if wdo$(\alpha) \leq_s$ wdo(β), then every m-subword of α has to be an m-subword of β too. On the other hand wdo$(\alpha) \leq_s$ wdo(β) still does not imply $\alpha \rightsquigarrow_m \beta$ (see, for example, the first item in Example 4). Thus we will use the following function p before applying wdo on the sentential forms of G. Let Σ be an alphabet and denote by $\hat{\Sigma}$ the alphabet $\{a^{(i)} \mid a \in \Sigma, i \in [m]\}$. Now let $p : \Sigma^* \rightarrow \hat{\Sigma}^*$ be defined as follows. For a word

14

$u = a_1 \cdots a_k \in \Sigma^*$ $(a_i \in \Sigma, i \in [k])$, let $p(u) = a_1^{(1 \mod m')} \cdots a_k^{(k \mod m')}$, where $m' = m - 1$. Intuitively, p associates the number $i \mod m'$ to the ith letter of u (we put this number in parentheses in order not to confuse it with the usual notation of the iteration of a letter). We will see in the proof of the next lemma that for two sentential forms α and β of G, $\mathrm{wdo}(p(\alpha)) \leq_s \mathrm{wdo}(p(\beta))$ implies $\alpha \rightsquigarrow_m \beta$.

Lemma 12. *Let $G = (V, \Sigma, R, S)$ be a pgSCG and $m \geq 1$. Then there is $M \in \mathbb{N}$ such that the following holds. For every derivation $\alpha_0 \Rightarrow \alpha_1 \Rightarrow \cdots \Rightarrow \alpha_n$ of G with $n \geq M$, there are $i < j$ in $[M]$ such that $\alpha_i \rightsquigarrow_m \alpha_j$.*

Proof. Let $\rho = \max\{|\mathrm{rhs}(r)| \mid r \in R\}$ and consider the sequence $n_1, n_2, \ldots$ where $n_i = i\rho$ $(i \geq 1)$. Let moreover M be the number given in Proposition 11 and $\alpha_0 \Rightarrow \alpha_1 \Rightarrow \cdots \Rightarrow \alpha_n$ be a derivation of G with $n \geq M$. Clearly, $|\mathrm{wdo}(p(\alpha_i))| \leq n_i$, for every $i \in [n]$. Then, by Proposition 11, there are numbers $i < j$ in $[M]$ such that $\mathrm{wdo}(p(\alpha_i)) \leq_s \mathrm{wdo}(p(\alpha_j))$. We show that $\alpha_i \rightsquigarrow_m \alpha_j$. To simplify the notation, let us denote $\mathrm{wdo}(p(\alpha_i))$ and $\mathrm{wdo}(p(\alpha_j))$ by u and v, respectively. If $|u| = 1$, then $|\alpha_i| \leq m$ and α_i is a subword of α_j. In this case $\alpha_i \rightsquigarrow_m \alpha_j$ trivially holds. Assume now that $|u| \geq 2$, and let $k = |u|$ and $l = |v|$. Since u is a scattered subword of v, there are $i_1 < \cdots < i_k$ in $\mathrm{pos}(v)$ such that $u = \mathrm{subw}(v, i_1, 1) \cdots \mathrm{subw}(v, i_k, 1)$. Then let $g : [k] \to [l]$ be a strictly increasing function defined as $g(\nu) = i_\nu$ $(\nu \in [k])$. Notice that $k = |\alpha_i| - m + 1$. Let moreover $f : \mathrm{pos}(\alpha_j) \to \mathrm{pos}(\alpha_i)$ be a (partial) function defined as $f(g(\nu) + \kappa) = \nu + \kappa$ $(\nu \in [k], \kappa \in [0, m - 1])$. To see that g is an m-embedding of α_i to α_j it is enough to show that f is letter preserving and well-defined.

Let $\nu \in [k]$. Using the definition of wdo we get that $\mathrm{subw}(p(\alpha_i), \nu, m) = \mathrm{subw}(p(\alpha_j), g(\nu), m)$ and in turn $\mathrm{subw}(\alpha_i, \nu, m) = \mathrm{subw}(\alpha_j, g(\nu), m)$. Thus f is letter preserving. Now, let $\nu \in [k - 1]$. Using again the definition of wdo we get that $\mathrm{subw}(p(\alpha_i), \nu, m) = \mathrm{subw}(p(\alpha_j), g(\nu), m)$ and $\mathrm{subw}(p(\alpha_i), \nu + 1, m) = \mathrm{subw}(p(\alpha_j), g(\nu + 1), m)$. Thus, the upper index added by p to the first letter of $\mathrm{subw}(\alpha_i, \nu, m)$ should match that of $\mathrm{subw}(\alpha_j, g(\nu), m)$. Similar observation holds for the words $\mathrm{subw}(\alpha_i, \nu + 1, m)$ and $\mathrm{subw}(\alpha_j, g(\nu + 1), m)$. This implies that either $g(\nu + 1) - g(\nu) = 1$ or $g(\nu + 1) - g(\nu) \geq m$ should hold. It is easy to see that in both cases the definition of f is consistent. Therefore g is an m-embedding of α_i to α_j. $\qquad\square$

Now we are ready to prove our pumping lemma.

Lemma 13. *Let $G = (V, \Sigma, R, S)$ be a pgSCG, $m = 2 \cdot \max_{pw(G)}$ and assume that $m \geq 1$. Then there is a number K such that for every word $w \in L(G)$ with $|w| \geq K$, there is a word $w' \in L(G)$ with $|w| < |w'| \leq (m + 1)|w|$.*

Proof. Let $N = \max\{|\mathrm{rhs}(r)| \mid r \in R\}$ and M be the number given by Lemma 12. Let $K = MN$ and consider a word $w \in L(G)$ with $|w| \geq K$. We give a word w' satisfying the conditions of the lemma. Let $der : S = \alpha_0 \Rightarrow \alpha_1 \Rightarrow \cdots \Rightarrow \alpha_n = w$ be

one of the shortest derivations of G from S to w. Clearly $n \geq M$. Thus, by Lemma 12, there are $i < j$ in $[M]$ such that $\alpha_i \leadsto_m \alpha_j$ (remember, $m = 2 \cdot \max_{pw(G)}$). We can assume that $|\alpha_i| < |\alpha_j|$. Indeed, assume on the contrary that this is not the case. Then, since G has no erasing rules, $|\alpha_i| = |\alpha_j|$. This, using (ii) of Proposition 5, implies that $\alpha_i = \alpha_j$. This yields that $der' : \alpha_0 \Rightarrow \cdots \Rightarrow \alpha_i \Rightarrow \alpha_{j+1} \Rightarrow \cdots \Rightarrow \alpha_n$ is also a derivation of G from S to w with $|der'| < n$. However this contradicts the assumption that der is a shortest derivation from S to w. Applying Lemma 9 with the parameters $\alpha = \alpha_i, \beta = \alpha_j, \gamma = w$ and $\gamma' = w'$, we get that there is a word $w' \in \Sigma^*$ such that $\alpha_i \Rightarrow^* w'$ and $|w| < |w'| \leq (m+1)|w|$. Since $S \Rightarrow^* \alpha_i$, also $S \Rightarrow^* w'$ holds. Consequently, $w' \in L(G)$. $\qquad\square$

Theorem 14. $\mathcal{L}(\text{pgSCG}) \subsetneq \mathcal{L}(\text{CS})$.

Proof. By [17] $\mathcal{L}(\text{pgSCG}) \subseteq \mathcal{L}(\text{CS})$. Thus, since $L = \{a^{2^{2^n}} \mid n \geq 0\}$ is clearly included in $\mathcal{L}(\text{CS})$, it is enough to show that $L \notin \mathcal{L}(\text{pgSCG})$. Assume on the contrary that $L \in \mathcal{L}(\text{pgSCG})$ and let G be a pgSCG with $L(G) = L$. Let moreover $m = 2 \cdot \max_{pw(G)}$. Since L is not a context-free language, we can assume that $m \geq 1$. Then let K be the number of Lemma 13 and let $k \geq K$ be such that $2^{2^k} > m + 1$. Now we put $w = a^{2^{2^k}}$. Clearly $w \in L(G)$ and it is easy to see that $|w| \geq K$. Thus, by Lemma 13, there is a word $w' \in L$ such that $|w| < |w'| \leq (m+1)|w|$.

Clearly, the shortest word $v \in L$ with $|w| < |v|$ is $a^{2^{2^{k+1}}}$. On the other hand, $|w'| \leq (m+1)2^{2^k} < 2^{2^k} \cdot 2^{2^k} = 2^{2^{k+1}} = |v|$. Hence $|w'| < |v|$ which implies that $w' \notin L$. This is a contradiction proving that $L \notin \mathcal{L}(\text{pgSCG})$. $\qquad\square$

Using Lemma 13 we can also show that for every $i \geq 1$, there is a language L in $\mathcal{L}(\text{fRCG})$ such that $L \notin \mathcal{L}(\text{pgSCG}_i)$.

Theorem 15. *For every $i \geq 1$, $\mathcal{L}(\text{fRCG}) \setminus \mathcal{L}(\text{pgSCG}_i) \neq \emptyset$.*

Proof. Let $i \geq 1$ and consider the language $L = \{a^{(2i+2)^n} \mid n \geq 1\}$. It is easy to see that replacing r_1 in Example 2 by $S \to A^{2i+2}$ we get an fRCG generating L. We show that $L \notin \mathcal{L}(\text{pgSCG}_i)$. Assume on the contrary that $L \in \mathcal{L}(\text{pgSCG}_i)$ and let $G = (V, \Sigma, R, S)$ be a pgSCG and $m = 2 \cdot \max_{pw(G)}$. Let moreover K be the number of Lemma 13 and consider $w = a^{(2i+2)^K}$. Then $w \in L$ and thus, by Lemma 13, there is a word $w' \in L$ with $|w| < |w'| \leq (m+1)|w| \leq (2i+1)(2i+2)^K$. On the other hand, the shortest word $v \in L$ with $|v| > |w|$ is $a^{(2i+2)^{K+1}}$. Now $|w'| \leq (2i+1) \cdot (2i+2)^K < (2i+2) \cdot (2i+2)^K = (2i+2)^{K+1} = |v|$. This is a contradiction, thus $L \notin \mathcal{L}(\text{pgSCG}_i)$. Hence $\mathcal{L}(\text{fRCG}) \setminus \mathcal{L}(\text{pgSCG}_i) \neq \emptyset$. $\qquad\square$

4. Conclusions

In this paper we investigated permitting semi-conditional grammars introduced by Kelemen [13], however we considered them in a more general form: every rule of

16

these grammars is associated with a set of words rather than a word. Then such a rule is applicable to a sentential form only if every word in the associated set is a subword of the sentential form. We solved a long lasting open question about the computational power of these grammars by showing that they are strictly weaker than context-sensitive grammars when erasing rules are not allowed. However, there are some interesting questions concerning these grammars that remained open. For example:

(i) Are permitting semi-conditional grammars strictly weaker than CS grammars if erasing rules are allowed?
(ii) Are the inclusions $\mathcal{L}(\text{pgSCG}_i) \subseteq \mathcal{L}(\text{pgSCG}_{i+1})$ $(i \geq 1)$ proper?
(iii) Is the inclusion $\mathcal{L}(\text{pRCG}) \subseteq \mathcal{L}(\text{pgSCG})$ proper?
(iv) Are the classes $\mathcal{L}(\text{pgSCG})$ and $\mathcal{L}(\text{fRCG})$ comparable?

Concerning (i), in [21] it was shown that allowing erasing rules does not increase the generative power of permitting random context grammars. We suspect that this might be the case for pgSCGs too. Concerning (iv), we think that the mentioned classes are incomparable. If it was possible, for example, to simulate pgSCGs with fRCGs, then fRCGs, on the one hand, could employ forbidding contexts, on the other hand they would have the ability of simulating the presence of permitting contexts. We think that this would imply that fRCGs can simulate RCGs. But we know that fRCGs are strictly weaker than RCGs [5]. A similar argumentation applies if we assume that pgSCGs can simulate fRCGs.

References

[1] H. Bordihn and H. Fernau, Accepting grammars and systems: An overview, *Developments in Language Theory, Magdeburg, Germany*, (1995).
[2] J. Dassow and T. Masopust, On restricted context-free grammars, *Journal of Computer and System Sciences* **78**(1) (2012) 293–304.
[3] J. Dassow and G. Păun, *Regulated Rewriting in Formal Language Theory* (Springer-Verlag Berlin Heidelberg, 1989).
[4] L. E. Dickson, Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors, *American Journal of Mathematics* **35**(4) (1913) 413–422.
[5] S. Ewert and A. van der Walt, A shrinking lemma for random forbidding context languages, *Theoretical Computer Science* **237**(1-2) (2000) 149–158.
[6] S. Ewert and A. van der Walt, A pumping lemma for random permitting context languages, *Theoretical Computer Science* **270**(1-2) (2002) 959–967.
[7] I. Friš, Grammars with partial ordering of the rules, *Information and Control* **12**(5) (1968) 415–425.
[8] Z. Gazdag, A note on context-free grammars with rewriting restrictions, *Proceedings of the 2010 Mini-Conference on Applied Theoretical Computer Science*, eds. A. Brodnik and G. Galambos (University of Primorska Press, Koper, 2011).
[9] Z. Gazdag, Remarks on some simple variants of random context grammars, *Journal of Automata, Languages and Combinatorics* **19**(1-4) (2014) 81–92.
[10] Z. Gazdag and K. Tichler, On the power of permitting semi-conditional grammars, *Developments in Language Theory: 21st International Conference, DLT 2017, LNCS 10396*, eds. É. Charlier, J. Leroy and M. Rigo (Springer, Cham, 2017), pp. 173–184.

[11] G. Higman, Ordering by divisibility in abstract algebras, *Proceedings of the London Mathematical Society* **3**(1) (1952) 326–336.

[12] D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (Prentice Hall PTR, Upper Saddle River, NJ, USA, 2000).

[13] J. Kelemen, Conditional grammars: Motivations, definitions, and some properties, *Proc. Conf. Automata, Languages and Mathematical Sciences, Salgótarján*, (1984), pp. 110–123.

[14] J. B. Kruskal, Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture, *Transactions of the American Mathematical Society* **95**(2) (1960) 210–225.

[15] T. Masopust, Simple restriction in context-free rewriting, *Journal of Computer and System Sciences* **76**(8) (2010) 837–846.

[16] G. Păun, On the generative capacity of conditional grammars, *Information and Control* **43**(2) (1979) 178 – 186.

[17] G. Păun, A variant of random context grammars: Semi-conditional grammars, *Theoretical Computer Science* **41** (1985) 1–17.

[18] G. Rozenberg and A. Salomaa, *Handbook of Formal Languages: Volume 2. Linear Modeling: Background and Application* (Springer-Verlag Berlin Heidelberg, 1997).

[19] A. Salomaa, *Formal Languages* (Academic Press, New York, London, 1973).

[20] A. van der Walt, Random context languages, *Inform. Process.* **71** (1972) 66–68.

[21] G. Zetzsche, On erasing productions in random context grammars, *Automata, Languages and Programming, 37th International Colloquium, ICALP (2). LNCS 6199*, eds. S. Abramsky, C. Gavoille, C. Kirchner, F. auf der Heide and P. G. Spirakis (Springer, 2010), pp. 175–186.