On the Power of Permitting Semi-conditional Grammars

Zsolt Gazdag[1(✉)] and Krisztián Tichler[2]

[1] Department of Foundations of Computer Science,
University of Szeged, Szeged, Hungary
`gazdag@inf.u-szeged.hu`
[2] Department of Algorithms and Their Applications,
Eötvös Loránd University, Budapest, Hungary
`ktichler@inf.elte.hu`

Abstract. Permitting semi-conditional grammars are such extensions of context-free grammars where each rule is associated with a word v, and such a rule can be applied to a sentential form u only if v is a subword of u. In this paper we show that the class of languages generated by permitting semi-conditional grammars with no erasing rules is strictly included in the class of context-sensitive languages.

Keywords: 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. One of their good properties is that their membership problem is efficiently solvable. These 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-dependencies, but the membership problem for them is already PSPACE-complete. In [11] Joshi gave several properties that a grammar should have in order to be able to model natural languages. These properties are the ability to handle *limited cross-serial dependencies*, the *constant growth* of the associated language, and the *polynomial time solvability of the membership problem*. Grammars satisfying these properties are called *mildly context-sensitive grammars*.

One way to enrich CF grammars with context sensitivity and, in turn, raise their generative power is to control their derivations by context conditions. For example, in *random context grammars* (RCG's) [18] 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 RCG's with erasing rules have equal power to that of Turing machines, 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* (rRCG's) in this paper. It turned out that even with this very limited ability of controlling the derivations these grammars are equivalent to random context grammars [2,15]. Moreover, permitting rRCG's are as powerful as permitting RCG's [2], and this is the case for the forbidding variants too if erasing rules are allowed [8].

In [16] a variant of random context grammars, called *semi-conditional grammars* (SCG's) were introduced. In these grammars every rule r is associated with two words, a permitting word v_1 and a forbidding one v_2, and r is applicable only if v_1 is a subword of the sentential form, but v_2 is not. Moreover, 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 [16] that SCG's without erasing rules and with degree $(1, 2)$ or $(2, 1)$ have equal power to that of CS grammars. This clearly means that these grammars are too powerful to meet all the conditions of mild context-sensitivity. It turned out in [16], on the other hand, that these grammars with degree $(1, 1)$ cannot generate all languages in CS. The invention of SCG's was motivated by the grammars of Kelemen [13], where only a permitting word was associated to each rule (we call these grammars *permitting semi-conditional grammars* in this paper). In [16] it remained an open question whether permitting SCG's can generate all CS languages. In this paper we show that there is a CS language that cannot be generated by any permitting RCG if the use of erasing rules is not allowed. Some results concerning grammars mentioned in this introduction are given in Fig. 1.

The proof of our result is based on a pumping lemma similar to the one presented in [6], where it was shown that the class of languages generated by permitting RCG's with no erasing rules is strictly included in the class of languages generated by RCG's. In the proof of the pumping lemma in [6] the following property was essential: 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
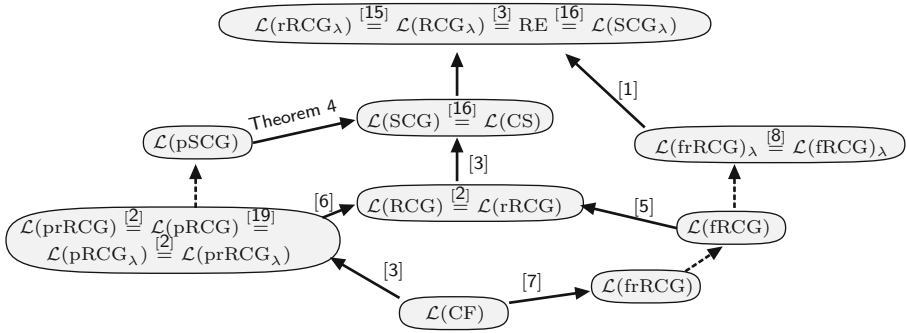
Fig. 1. A comparison of the power of some variants of grammars mentioned in the introduction. Arrows with solid lines represent strict inclusions, while arrows with dashed lines indicate inclusions which are not known to be strict. References to the presented equalities or strict inclusions are also given. Inclusions represented by dashed lines follow from definitions. SCG_λ, RCG_λ, and $rRCG_\lambda$ (resp. SCG, 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.

$v_1, v_2, \ldots$ of n-vectors over the natural numbers contains an infinite sub-sequence $v_{i_1} \leq v_{i_2} \leq \ldots$, where $\leq$ is the componentwise ordering of n-vectors.

In the proof of our pumping lemma (Lemma 2) we need to find such sentential forms α and β in a derivation of a permitting SCG G 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 is in α. To do so we will use Higman's lemma [10], which ensures that in any infinite sequence $v_1, v_2, \ldots$ of words, there is an infinite subsequence $v_{i_1} \leq_s v_{i_2} \leq_s \ldots$, where $\leq_s$ is the subsequence (or scattered subword) relation. However, to find an appropriate α and β we cannot apply directly Higman's lemma to the sentential forms of a derivation, but rather to certain carefully defined words obtained from these sentential forms.

The paper is organized as follows. First, we introduce the necessary notions and notations. Then, in Sect. 3 we present the main result of the paper. Finally, we give some concluding remarks in Sect. 4. Due to space reasons, some proofs are omitted in the paper. The interested reader can find them in [9].

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 [17]. 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 letter $a \in \Sigma$ and a word $u \in \Sigma^*$, $|u|$ denotes the *length of* u and $|u|_a$ is the *number of occurrences of a in 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 $i \in \text{pos}(u)$ and $m \geq 1$ be such that $i + m - 1 \in \text{pos}(u)$. 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 $f : \text{pos}(v) \to \text{pos}(u)$ 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 [10] (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 $u_{i_1} \leq_s u_{i_2} \leq_s \ldots$.*

A *semi-conditional grammar* (SCG 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 $p, q \in (V \cup \Sigma)^*$. For such a rule r, the words p and q are called the *permitting* and *forbidding contexts* of r, respectively. The *right-hand side* of r (denoted by 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) p is a subword of $u_1 A u_2$, and (ii) if $q \neq \varepsilon$, then q is not 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 called a *sentential form of G* (or just a *sentential form* if G is clear from the context). A *derivation der* from α to β is a sequence $\alpha_1 \Rightarrow \alpha_2 \Rightarrow \ldots \Rightarrow \alpha_{n+1}$ of sentential forms, 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 α and β be sentential forms and *der* a derivation from α to β. The *sentential form vector of der* (denoted by sfv(der)) is $(u_1, \ldots, u_k)$ ($k = |\alpha|, u_i \in V_G^*, i \in [k]$),

such that $\beta = u_1 \ldots u_k$ and, for every $i \in [k]$, u_i is derived from $\mathrm{subw}(\alpha, i, 1)$. Let $der : \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n$ $(n \geq 1)$ and $der' : \alpha_n \Rightarrow \ldots \Rightarrow \alpha_k$ $(k \geq n)$. Then $der\, der'$ denotes the derivation $\alpha_1 \Rightarrow \ldots \Rightarrow \alpha_k$.

If, for every rule $(A \to \alpha, p, q)$ in R, $q = \varepsilon$, then G is a *permitting SCG*, or a *pSCG* for short. Let $G = (V, \Sigma, R, S)$ be a pSCG. The *set of permitting contexts of G* is $pw(G) = \{p \mid (r, p, q) \in R\}$ and $\max_{pw(G)} = \max\{|u| \mid u \in pw(G)\}$. We denote by $\mathcal{L}(\mathrm{pSCG})$ and $\mathcal{L}(\mathrm{CS})$ the families of languages generated by pSCG's and context-sensitive grammars, respectively.

3   The Main Result

Here we show that pSCG's 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 pSCG (Theorem 4). The proof, roughly, consists of the following main steps. First we define the notion of *m-embedding* (Definition 1). 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. Then we show that if a pSCG $G = (V, \Sigma, R, S)$ with $m = \max_{pw(G)}$ has a derivation *der* from α to β $(\alpha, \beta \in V_G^*)$ such that $|\alpha| < |\beta|$ and α can be *m-embedded* to β, then the *der* can be "pumped" so that the obtained derivation is a valid derivation of G (cf. Lemma 2, which we will often refer to as our *pumping lemma*). Finally, we show that sufficiently long derivations of G always contain sentential forms α and β such that α can be *m-embedded* to β (Lemma 3). To this end we will use the fact that $\leq_s$ is a wqo on V_G^*.

Definition 1. 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 1. *Here we give two examples to demonstrate the notion of m-embedding.*

(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 2. *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 $f = \operatorname{inv}_m(g)$. Then the following statements hold.*

(i) *For every $i \in \operatorname{pos}(\alpha)$, $|\{t \in \operatorname{pos}(\beta) \mid f(t) = i\}| \leq m$.*
(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 $g(j) - g(i) \geq m$.*

Proof. See [9]. □

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

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 $\operatorname{subw}(\alpha, i, 1) = \operatorname{subw}(\alpha, j, 1)$, then take such a j and let $\hat{f}(i) = f(j)$. Notice that f is always an extension of itself.

Example 2. *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 $f : [5] \rightarrow [3]$ be the following partial function. $f(2) = f(3) = 3$, $f(5) = 1$. Then $\operatorname{subst}(\mathbf{u}, \alpha, f) = Au_3 u_3 B u_1 = ACCCCBAA$ 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 $\operatorname{subst}(\mathbf{u}, \alpha, \hat{f})$ equal to $AC^4 A^4$ or $AC^6 A^2$, respectively.*

The following lemma will be crucial in the proof of our pumping lemma.

Lemma 1. *Let $G = (V, \Sigma, R, S)$ be a pSCG, $m = \max_{pw(G)}$, and $\alpha, \alpha', \beta \in V_G^*$. Assume that $\alpha \Rightarrow_G^* \alpha'$ and $\alpha \leadsto_m \beta$. Let der be a derivation from α to α', g an m-embedding of α to β, and $f = \operatorname{inv}_m(g)$. Then, for every extension $\hat{f}$ of f, $\beta \Rightarrow_G^* \beta_{\hat{f}}$, where $\beta_{\hat{f}} = \operatorname{subst}(\operatorname{sfv}(der), \beta, \hat{f})$.*

Proof. Let $\beta' = \operatorname{subst}(\operatorname{sfv}(der), \beta, f)$. We first show that $\beta \Rightarrow_G^* \beta'$ 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 $\alpha_0 \Rightarrow_G \ldots \Rightarrow_G \alpha_n$, der_2 is $\alpha_n \Rightarrow_G \alpha_{n+1}$, $\alpha_0 = \alpha$, and $\alpha_{n+1} = \alpha'$. Let $\beta_n = \operatorname{subst}(\operatorname{sfv}(der_1), \beta, f)$. By the induction hypothesis, there is a derivation der_1' from β to β_n. Let $(u_1, \ldots, u_k) = \operatorname{sfv}(der_1)$ ($k = |\alpha|$). Assume that G rewrites a nonterminal A during $\alpha_n \Rightarrow_G \alpha_{n+1}$ using a rule $r = (A \rightarrow \gamma, p, \varepsilon)$ (see Fig. 2 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 $i_1 < i_2 \ldots < i_\xi$ 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 $\beta' = \beta_{n+1} = \mathrm{subst}((\gamma), \beta_n, h)$ where $h : \mathrm{pos}(\beta) \to \{1\}$ is defined as follows: $h(j) = 1$ if $j \in \{\kappa_1, \ldots, \kappa_\xi\}$, and it is undefined otherwise. Therefore, to prove $\beta_n \Rightarrow_G^* \beta'$ 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.

Since G can apply r at the step $\alpha_n \Rightarrow_G \alpha_{n+1}$, α_n should contain the permitting context p. Then there are $\mu \in [k]$ and $\nu \in [0, m-1]$ such that p occurs in the subword $u_\mu \ldots u_{\mu+\nu}$ of α_n (notice that G has no erasing rules). Since g is an m-embedding of α to β, it is clear that $v_{g(\mu)} \ldots v_{g(\mu)+\nu} = u_\mu \ldots u_{\mu+\nu}$. Thus, there is at most one index $j \in [\xi]$ such that $v_{g(\mu)} \ldots v_{g(\mu)+\nu}$ contains that A which occurs on the κ_jth position in β_n. If no such j exists, then G can rewrite all A's occurring on positions κ_j $(j \in [\xi])$ in β_n, since the permitting context p is always present as a subword of $v_{g(\mu)} \ldots v_{g(\mu)+\nu}$. Otherwise let $j \in [\xi]$ be such that the subword $v_{g(\mu)} \ldots v_{g(\mu)+\nu}$ includes that A which occurs on the κ_jth position of β_n. Then G should rewrite first those A's in β_n that occur on positions other that κ_j and, at the last step, that A which occurs on the κ_jth position. Therefore $\beta_n \Rightarrow_G^* \beta_{n+1} = \beta'$ which implies that $\beta \Rightarrow_G^* \beta'$.

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, for each derivation step in der', there is a $j \in dom(f)$ such that $[j, j+m-1] \subseteq dom(f)$, and the necessary permitting context is in a sentential form derived from $\mathrm{subw}(\beta, j, m)$. In other words, those 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 $i \in \mathrm{pos}(\beta) - dom(f)$ such that $\hat{f}(i) = f(j)$, for some $j \in dom(f)$. Let u be the $f(j)$th word in $\mathrm{sfv}(der)$ and $X = \mathrm{subw}(\beta, j, 1)$. Then G derives u during der' from this X. On the other hand, by the definition of $\hat{f}$, $\mathrm{subw}(\beta, i, 1) = X$. 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.

To prove out pumping lemma we need the following preparation. Let $g : \alpha \leadsto_m \beta$ be an m-embedding. The mapping

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

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 3. *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 $|\beta| = l$. Suppose that $g : \alpha \leadsto_m \beta$ with $f = \mathrm{inv}_m(g)$ and $\bar{g} = \mathrm{cmp}_g$.*

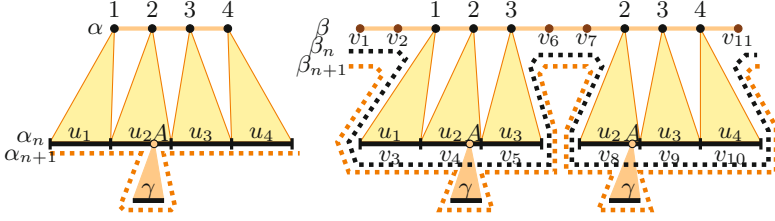**Fig. 2.** The inductive proof of Lemma 1 assuming $m = 3$

Let us introduce the notations $x_i = \begin{cases} z_{f(i)} & \text{if } i \in dom(f) \\ \text{subw}(\beta, i, 1) & \text{if } i \notin dom(f) \end{cases}$,

$\zeta(i, r) = \sum_{j=1}^{i-1} |z_j| + r$ $(i \in [k], r \in [|z_i|])$ *and* $\xi(i, r) = \sum_{j=1}^{i-1} |x_j| + r$ $(i \in [l], r \in [|x_i|])$. *Then for the mapping* $g'(\zeta(i, r)) := \xi(\bar{g}(i), r)$ $(\zeta(i, r) \in [l - m + 1])$,

$$g' : \beta \leadsto_m \text{subst}(\mathbf{u}, \beta, f)$$

holds, where $\mathbf{u} = (z_1, \ldots, z_k)$.

Proof. See [9].

Lemma 2. *Let* $G = (V, \Sigma, R, S)$ *be a pSCG and* $m = \max_{pw(G)}$. *Suppose that* $\alpha \Rightarrow^* \beta$, $\beta \Rightarrow^* \gamma$, $\alpha \leadsto_m \beta$, *and* $|\alpha| < |\beta|$ *hold for some* $\alpha, \beta, \gamma \in V_G^*$. *Then there is 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. Let moreover $f = \text{inv}_m(g)$, and der', der'' be any derivations from α to β and from β to γ, respectively. Applying Proposition 3 with these parameters and $\Sigma = V_G$, $\mathbf{u} = \text{sfv}(der')$ we get $g' : \beta \leadsto_m \beta'$, where $\beta' = \text{subst}(\mathbf{u}, \beta, f)$. Let $f' = \text{inv}_m(g')$. By Lemma 1 $\beta \Rightarrow^* \beta'$.

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 $r \in [|x_i|]$, 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 $x_i = \text{subw}(\beta, i, 1)$, 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 $\text{subw}(\beta', \tau, 1) = \text{subw}(\beta, \hat{f}'(\tau), 1) = \text{subw}(\beta, f'(\tau'), 1) = \text{subw}(\beta', \tau', 1)$. Thus $\hat{f}'$ is an extension of f'.

Now let $\gamma' = \text{subst}((v_1, \ldots, v_l), \beta', \hat{f}')$, where $(v_1, \ldots, v_l) = \text{sfv}(der'')$. By Lemma 1, $\beta' \Rightarrow_G^* \gamma'$. This, together with $\alpha \Rightarrow_G^* \beta$ and $\beta \Rightarrow_G^* \beta'$ implies $\alpha \Rightarrow_G^* \gamma'$, 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 $f'(\tau) = i$. Thus, each v_i $(i \in [l])$ is substituted for a position in β' by f'. Therefore $|\gamma| = |v_1 \ldots v_l| \leq |\text{subst}((v_1, \ldots, v_l), \beta', f')| \leq |\text{subst}((v_1, \ldots, v_l), \beta', \hat{f}')| = |\gamma'|$. 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 2, 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).

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

Example 3. *Let* $G = (V, \Sigma, R, S)$ *be a pSCG,* $\alpha = ABBA$, $\beta = EABBFBACA$ $(A, B, C, E, F \in V \cup \Sigma)$, *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* $\alpha \rightsquigarrow_m \beta$ *with the following* m-*embedding* g: $g(1) = 2$, $g(2) = 3$, *and* $g(3) = 6$. *Then, according to Lemma 2, we can give a* $\gamma' \in \Sigma^*$ *with the following properties:* (i) $\alpha \Rightarrow^* \gamma'$ *and* (ii) $|\gamma| < |\gamma'| \leq (m+1)|\gamma|$.

Assume, for instance, that $\mathrm{sfv}(der') = (z_1, z_2, z_3, z_4)$, *where* $z_1 = E$, $z_2 = AB$, $z_3 = BF$, *and* $z_4 = BACA$. *Let* $f = \mathrm{inv}_m(g)$. *Then* $\beta' = \mathrm{subst}(\mathrm{sfv}(der'), \beta, f) = x_1 \ldots x_9$, *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 Lemma 2, 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* $f' = \mathrm{inv}_m(g')$. *Then* $dom(f') = [2, 5] \cup [8, 13]$. *Let us define now* $\hat{f}'$ *according to the proof of Lemma 2, 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* $\tau \in [1, 15] - dom(f')$, $\mathrm{subw}(\beta', \tau, 1) = \mathrm{subw}(\beta, \hat{f}'(\tau), 1)$. *Assume that* $\mathrm{sfv}(der'') = v_1 \ldots v_9$, *where* $v_i \in \Sigma^*$ $(i \in \mathrm{pos}(\beta))$. *Then* $\gamma' = \mathrm{subst}(\mathrm{sfv}(der''), \beta', \hat{f}') = v_1 v_1 \ldots v_5 v_5 \ldots v_9 v_8 v_9$. *By* (i) *of Lemma 2,* $\alpha \Rightarrow^* \gamma'$, *and it is easy to check that* γ' *satisfies Statement* (ii) *too.*

The following proposition together with Lemma 3 will be used to show that sufficiently long derivations of pSCG's always contain sentential forms α and β satisfying the conditions of Lemma 2. The statement can be seen using Proposition 1 (see also the proof of Lemma 2 in [6]).

Proposition 4. *Let* Σ *be an alphabet and* $n_1, n_2, \ldots$ *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* $v_i \leq_s v_j$.

Let $G = (V, \Sigma, R, S)$ be a pSCG 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 4 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 4 not directly to the derivations of G but to sequences of words

derived 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 : $\Sigma^* \to (\Sigma^{\leq m})^*$ 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$(u) = \text{subw}(u, 1, m) \ldots \text{subw}(u, |u| - m + 1, m)$ (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 $\alpha \leadsto_m \beta$ (see, for example, the first item in Example 1). Thus we will use the following function p before applying wdo on the sentential forms of G. Let Σ be an alphabet. Then $\hat{\Sigma}$ denotes the alphabet $\{a^{(i)} \mid a \in \Sigma, i \in [m]\}$. Now let $p : \Sigma^* \to \hat{\Sigma}^*$ be defined as follows. For a word $u = a_1 \ldots a_k \in \Sigma^*$ ($a_i \in \Sigma, i \in [k]$), $p(u) = a_1^{(1 \mod m)} \ldots a_k^{(k \mod m)}$. 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, wdo$(p(\alpha)) \leq_s$ wdo$(p(\beta))$ implies $\alpha \leadsto_m \beta$.

Lemma 3. *Let $G = (V, \Sigma, R, S)$ be a pSCG and $m \geq 1$. Then there is $M \in \mathbb{N}$ such that the following holds. For every derivation $\alpha_0 \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n$ of G with $n \geq M$, there are $i < j$ in $[M]$ such that $\alpha_i \leadsto_m \alpha_j$.*

Proof. Let $m = \max_{pw(G)}$, $\rho = \max\{|\text{rhs}(r)| \mid r \in R\}$ 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 4 and $\alpha_0 \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n$ be a derivation of G with $n \geq M$. Clearly, $|\text{wdo}(p(\alpha_i))| \leq n_i$, for every $i \in [n]$. Then, by Proposition 4, there are numbers $i < j$ in $[M]$ such that wdo$(p(\alpha_i)) \leq_s$ wdo$(p(\alpha_j))$. We show that $\alpha_i \leadsto_m \alpha_j$. To simplify the notation, let us denote wdo$(p(\alpha_i))$ and 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 $\alpha_i \leadsto_m \alpha_j$ trivially holds. Assume now that $|u| \geq 2$, and let $k = |u|$ and $l = |\alpha_j|$. Since u is a scattered subword of v, there are $i_1 < \ldots < i_k$ in pos(v) such that $u = \text{subw}(v, i_1, 1) \ldots \text{subw}(v, i_k, 1)$. 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 $f : \text{pos}(\alpha_j) \to \text{pos}(\alpha_i)$ 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 subw$(p(\alpha_i), \nu, m) = $ subw$(p(\alpha_j), g(\nu), m)$ and in turn subw$(\alpha_i, \nu, m) = \text{subw}(\alpha_j, g(\nu), m)$. Thus f is letter preserving. Now, let $\nu \in [k - 1]$. Using again the definition of wdo we get that subw$(p(\alpha_i), \nu, m) = \text{subw}(p(\alpha_j), g(\nu), m)$ and subw$(p(\alpha_i), \nu + 1, m) = $ subw$(p(\alpha_j), g(\nu + 1), m)$. Thus, the upper index added by p to the first letter of

subw(α_i, ν, m) should match that of subw$(\alpha_j, g(\nu), m)$. Similar observation holds for the words subw$(\alpha_i, \nu + 1, m)$ and 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.

Theorem 4. $\mathcal{L}(\mathrm{pSCG}) \subsetneq \mathcal{L}(\mathrm{CS})$.

Proof. By [16] $\mathcal{L}(\mathrm{pSCG}) \subseteq \mathcal{L}(\mathrm{CS})$. Thus, since $L = \{a^{2^{2^n}} \mid n \geq 0\}$ is clearly included in $\mathcal{L}(\mathrm{CS})$, it is enough to show that $L \notin \mathcal{L}(\mathrm{pSCG})$. Assume on the contrary that $L \in \mathcal{L}(\mathrm{pSCG})$ and let G be a pSCG with $L(G) = L$. Let moreover $m = \max_{pw(G)}$. Since L is not a context-free language, we can assume that $m \geq 1$. Then let M be the number of Lemma 3, and $u = a^{2^{2^{mMN}}}$, where $N = \max\{|\mathrm{rhs}(r)| \mid r \in R\}$. Let moreover $der : S = \alpha_0 \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = u$ be one of the shortest derivations of G from S to u. Clearly $n \geq M$. Thus, by Lemma 3, there are $i < j$ in $[M]$ such that $\alpha_i \leadsto_m \alpha_j$. 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 2, implies that $\alpha_i = \alpha_j$. This yields that $der' : \alpha_0 \Rightarrow \ldots \Rightarrow \alpha_i \Rightarrow \alpha_{j+1} \Rightarrow \ldots \Rightarrow \alpha_n$ is also a derivation of G from S to u with $|der'| < n$. However this contradicts the assumption that der is a shortest derivation from S to u. Applying Lemma 2 we get that there is an $u' \in \Sigma^*$ such that $\alpha_i \Rightarrow^* u'$ and $|u| < |u'| \leq (m+1)|u|$. Since $S \Rightarrow^* \alpha_i$, also $S \Rightarrow^* u'$ holds. Consequently, $u' \in L$.

Clearly, the shortest word $v \in L$ with $|u| < |v|$ is $a^{2^{2^{mMN+1}}}$. On the other hand, $|u'| \leq (m+1)2^{2^{mMN}} < 2^{2^{mMN}} \cdot 2^{2^{mMN}} = 2^{2^{mMN+1}} = |v|$. Thus $|u'| < |v|$ yielding $u' \notin L$ which is a contradiction. Therefore $L \notin \mathcal{L}(\mathrm{pSCG})$.

4   Conclusions

In this paper we have investigated permitting semi-conditional grammars introduced by Kelemen [13]. We showed that these grammars are strictly weaker than context-sensitive grammars when erasing rules are not allowed. However, it is still open whether this remains true if erasing rules are allowed. In [19] it was shown that allowing erasing rules does not increase the generative power of permitting random context grammars. To decide whether this holds also for permitting semi-conditional grammars is a possible topic for future work. It is also an interesting question, for example, whether the inclusion $\mathcal{L}(\mathrm{pRCG}) \subseteq \mathcal{L}(\mathrm{pSCG})$ depicted in Fig. 1 is strict or not.

References

1. Bordihn, H., Fernau, H.: Accepting grammars and systems: an overview. In: Developments in Language Theory, Magdeburg, Germany (1995)
2. Dassow, J., Masopust, T.: On restricted context-free grammars. J. Comput. Syst. Sci. **78**(1), 293–304 (2012)
3. Dassow, J., Păun, G.: Regulated Rewriting in Formal Language Theory. Springer, New York (1989)
4. Dickson, L.E.: Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors. Am. J. Math. **35**(4), 413–422 (1913)
5. Ewert, S., van der Walt, A.: A shrinking lemma for random forbidding context languages. Theor. Comput. Sci. **237**(1–2), 149–158 (2000)
6. Ewert, S., van der Walt, A.: A pumping lemma for random permitting context languages. Theor. Comput. Sci. **270**(1–2), 959–967 (2002)
7. Gazdag, Z.: A note on context-free grammars with rewriting restrictions. In: Brodnik, A., Galambos, G. (eds.) Proceedings of the 2010 Mini-Conference on Applied Theoretical Computer Science. University of Primorska Press, Koper (2011)
8. Gazdag, Z.: Remarks on some simple variants of random context grammars. J. Autom. Lang. Comb. **19**(1–4), 81–92 (2014)
9. Gazdag, Z., Tichler, K.: On the power of permitting semi-conditional grammars, extended version. https://www.researchgate.net/publication/312587701_On_the_Power_of_Permitting_Semi-conditional_Grammars
10. Higman, G.: Ordering by divisibility in abstract algebras. Proc. Lond. Math. Soc. **3**(1), 326–336 (1952)
11. Joshi, A.K.: Tree adjoining grammars: how much context-sensitivity is required to provide reasonable structural descriptions? In: Dowty, D.R., Karttunen, L., Zwicky, A.M. (eds.) Natural Language Parsing, pp. 206–250. Cambridge University Press, Cambridge (1985)
12. Jurafsky, D., Martin, J.H.: 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. Kelemen, J.: Conditional grammars: motivations, definitions, and some properties. In: Proceedings of the Conference on Automata, Languages and Mathematical Sciences, Salgótarján, pp. 110–123 (1984)
14. Kruskal, J.B.: Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. Trans. Am. Math. Soc. **95**(2), 210–225 (1960)
15. Masopust, T.: Simple restriction in context-free rewriting. J. Comput. Syst. Sci. **76**(8), 837–846 (2010)
16. Păun, G.: A variant of random context grammars: semi-conditional grammars. Theor. Comput. Sci. **41**, 1–17 (1985)
17. Salomaa, A.: Formal Languages. Academic Press, New York, London (1973)
18. van der Walt, A.: Random context languages. Inf. Process. **71**, 66–68 (1972)
19. Zetzsche, G.: On erasing productions in random context grammars. In: Abramsky, S., Gavoille, C., Kirchner, C., Meyer auf der Heide, F., Spirakis, P.G. (eds.) ICALP 2010. LNCS, vol. 6199, pp. 175–186. Springer, Heidelberg (2010). doi:10.1007/978-3-642-14162-1_15