Communication complexity hierarchies for distributive generation of languages

Dana Pardubská*†‡
Faculty of Mathematics and Physics
Comenius University
842 15 Bratislava
Slovakia

September 21, 1993

Abstract

Communication complexity of language generation by Parallel Communicating Grammar Systems (PCGS) is investigated. The only results on communication complexity obtained till now are some hierarchies for PCGS with some additional restrictions on the communication structure of PCGS. The first results for PCGS with no restriction on their communication structure are presented here. The main ones are:

(i) There is an infinite hierarchy of constant communication complexity for PCGS without restrictions on their communication graph; i.e. $k+1$ communications are more powerful than k.

(ii) There is a gap between the communication complexity $O(1)$ and $\Omega(\log n)$ and $\Theta(\log n)$ is necessary and sufficient number of communications to get a nonregular language over one-letter alphabet.

(iii) There is an infinite hierarchy of communication complexity functions $f_k(n) = \sqrt[k]{n}$ for $k \in I\!N$.

---

*This work has been done during the author's stay at the Dept. of Mathematics & Informatics of the University of Paderborn

†A part of results has been presented at the conference IMYCS'92

‡The author was partially supported by MSV SR 1/990123/93 and by EC Cooperative Action IC 1000: project ALTEC

1 Introduction.

This paper is devoted to the computational aspects of Parallel Communicating Grammar systems ($PCGS$) introduced in [PS 89]. The concept of $PCGS$s differs from the previous models of parallel derivation of words (languages) like Lindenmayer systems [HR 75, RS 80] in that $PCGS$s can be viewed as typical distributive systems consisting of a number of independent elements cooperating by the exchange of information via communication links. Thus, the derivation of a $PCGS$ is a sequence of parallel derivation steps and communication steps. As typical for distributive systems there are several complexity measures of importance for $PCGS$. The number of grammars (elements) of $PCGS$ (called the degree of $PCGS$) has been considered as a descriptional complexity measure already in [SK 92] and investigated in [HKK 93]. Another descriptional complexity measures corresponding to the communication structure (a directed graph whose nodes are elements of a given $PCGS$ and edges correspond to the communication links between the elements of $PCGS$) has been introduced and investigated in [HKK 93, Par 93].

Communication complexity as a computational complexity measure corresponding to the number of communications between the component grammars of $PCGS$ is studied in this paper. The basic questions concerning the power of communication complexity are investigated here. It is easy to see that only one communication is sufficient to generate nonregular language $\{a^n b^n \mid n \geq 1\}$. In [HKK 93] the power of $PCGS$'s with special communication structures were studied from computational complexity point of view. It was shown there, that there is an infinite hierarchy of constant communication complexity for $tree - PCGS$ and $c - PCGS$ (PCGS whose communication structure is a tree, resp. tree with a sole nonleaf vertex). To get the hierarchy for unrestricted communication structure has been left open there. Our first result solves this problem by showing that $k + 1$ communications are more powerful than k communications for unrestricted $PCGS$, and for $dag - PCGS$(PCGS whose communication structure is a directed acyclic graph).

The situation is not the same considering only the generation of languages over one-letter alphabet. We prove that every language over one-letter alphabet, that can be generated with $PCGS$ with a constant communication complexity, is regular. We know [San 90] that there are non-regular languages over one-letter alphabet generable by $PCGS$. That means, there are languages, for which there is not any constant bounding the amount of communication complexity necessary for their generation. So, it is quite natural to ask, what is the less possible amount of the communication complexity measure that allow the $PCGS$'s to generate nonregular language over one-letter alphabet. The answer to this question is given here, too. We prove, that if communication complexity cannot be bounded by any constant, then it is at least $\Omega(\log n)$. In fact, we even prove the existence of the gap between $O(1)$ and $\Omega(\log n)$ for languages over arbitrary finite alphabet.

There have been only hierarchies of constant communication complexity obtained, till now. But, in fact, the communication complexity is defined as a function depending on the length of the generated word. We give here first results showing that there is hierarchy of communication complexity for some special class of functions. The only known result of this kind has been established in [HKK 93], where a language requiring linear $\Omega(n)$ communication complexity to be generated by $tree - PCGS$s is constructed.

The paper is organized as follows. Next section contains the basic definitions, notions and denotations. The results referred to constant communication complexity are presented in Section 3. Section 4 contains the results concerning nonconstant communication complexity.

2 Preliminaries

We assume the reader to be familiar with basic definitions and notations in formal language theory and we specify only some of them related to the PCGS. We denote by ε the empty symbol (word) and, for any word x, $|x|$ denotes the length of x. For a set K of symbols and a word x, $|x|_K$ denotes the number of occurrences of symbols of K in x. Let R denote the set of regular languages.

First, we give the definition of Parallel Communicating Grammar System ($PCGS$).

Definition 1 *A PCGS of degree m, $m \geq 1$, is an $(m+1)$-tuple $\Pi = (G_1, \dots, G_m, K)$, where*

- *$G_i = (N_i, T, S_i, P_i)$ are regular grammars satisfying*
  - *$N_i \cap T = \emptyset$ for all $i \in \{1, \dots, m\}$*
  - *$P_i \subset N$ x $T^* N \cup N$ x T^+*
- *$K \subseteq \{Q_1, \dots, Q_m\} \bigcap \cup_{i=1}^{m} N_i$ is a set of special symbols, called communication symbols, $K_i = K \cap N_i$ is the set of communication symbols of G_i.*

Now, we describe the work of *PCGS*s. The possible communications in PCGS Π are determined by the communication graph. The vertices of this directed graph $G(\Pi)$ corresponds to the individual component grammars and are labeled by their names $G_1, \dots, G_m$. The directed edges describe the possibility of inquiry. The edge (G_i, G_j) is presented in the $G(\Pi)$ iff the communication symbol Q_j belongs to the nonterminal alphabet of G_i.

An m-tuple $(x_1, \dots, x_m)$, $x_i = \alpha_i A_i$, $\alpha_i \in T^*, A_i \in (N_i \cup \varepsilon)$, is called configuration. With every configuration $C = (\alpha_1 A_1, \dots, \alpha_m A_m)$ its nonterminal cut $N(C) = (A_1, A_2, \dots, A_m)$ is associated. If the nonterminal cut of the configuration contains at least one communication symbol, then so called communication cut, that is m-tuple $(B_1, B_2, \dots, B_m)$, where $B_i = A_i$ for $A_i \in K$ and $B_i = \varepsilon$ for $A_i \notin K$, is associated with it, too.

We say a configuration $(x_1, \dots, x_m)$ directly derives a configuration $(y_1, \dots, y_m)$ and write $(x_1, \dots, x_m) \rightarrow (y_1, \dots, y_m)$, if one of the next two cases holds:

1. $|x|_K = 0$ for all i, $1 \leq i \leq m$, and either $x_i \rightarrow y_i$ in G_i when x_i contains nonterminal or x_i is the terminal word and $y_i = x_i$.
2. if $|x_i|_K > 0$ for some i, $1 \leq i \leq m$, then, for each i such that $x_i = z_i Q_{j_i}$, for some $z_i \in T^*, \forall Q_{j_i} \in K$, the following happens:
   (a) If $|x_{j_i}|_K = 0$ then $y_i = z_i x_{j_i}$ and $y_{j_i} = S_{j_i}$
   (b) If $|x_{j_i}|_K > 0$ then $y_i = x_i$.

For all remaining indeces t, for which x_t does not contain communication symbol and Q_t has not occured in any of x_i's, we put $y_t = x_t$.

A derivation of a $PCGS$ Π is a sequence of configurations $X_1, X_2, \ldots, X_t$, where $X_i \rightarrow X_{i+1}$ is in Π. It can be viewed as a sequence of *rewriting* and *communication steps*, too.

If no communication symbol appears in any of the component grammars then we perform a rewriting step consisting of rewriting steps synchronously performed in each of the grammars. If some of the commponents is a terminal string, it is left unchanged. If some of the component grammars contains nonterminal that cannot be rewritten, the derivation is blocked. If the first grammar G_1 contains a terminal word y, the derivation is finished and y is the word generated by Π in this derivation. If a communication symbol is present in any of the components, then a communication step is performed. It consists of replacing all communication symbols with the phrases they refer to under condition these phrases do not contain communication symbol. If some communication symbols are not satisfied in this communication step, they may be satisfied in one of the next ones. Communication steps are performed until no more communication symbols are present or the derivation is blocked because no communication symbol has been satisfied in the last communication step.

The language generated by a $PCGS$ consists of the terminal words generated in G_1 (in the cooperation with the other grammars).

Definition 2 *For any $PCGS$ Π, $L(\Pi) = \{\, \alpha \in T^* | \, (S_1, \ldots, S_m) \rightarrow^* (\alpha, \beta_2, \ldots, \beta_m)\}$*

Now, to illustrate the definition of PCGS and to motivate some new notions, the example of PCGS follows.

Example 1 *Let us have a $PCGS$ Π (unambiguously) given by the sets of rules of its component grammars*

$G_1 \quad : S_1 \rightarrow aS_1|aQ_2, \quad Z_2 \rightarrow aZ_2|aQ_3, \quad Z_3 \rightarrow b$

$G_2 \quad : S_2 \rightarrow bZ_2 \qquad\quad Z_2 \rightarrow bZ_2$

$G_3 : \quad S_3 \rightarrow Z_3 \qquad\quad Z_3 \rightarrow bZ_3.$

$$G_2 \longleftarrow G_1 \longrightarrow G_3$$

Figure 1: communication graph of Π.

The communication graph of Π, $G(\Pi)$, is depicted on the Figure 1.

The derivation of PCGS Π can proceed as follows.

$(S_1, S_2, S_3) \xrightarrow{1} (aS_1, bZ_2, Z_3) \xrightarrow{2} (a^2S_1, b^2Z_2, bZ_3) \xrightarrow{3} (a^3Q_2, b^3Z_2, b^2Z_3) \xrightarrow{4}$
$\xrightarrow{4} (a^3b^3Z_2, S_2, b^2Z_3) \xrightarrow{5} (a^3b^3aZ_2, bZ_2, b^3Z_3) \xrightarrow{6} (a^3b^3a^2Q_3, b^2Z_2, b^4Z_3) \xrightarrow{7}$
$\xrightarrow{7} (a^3b^3a^2b^4Z_3, b^2Z_2, S_3) \xrightarrow{8} (a^3b^3a^2b^5, b^3Z_2, Z_3)$

The first, second and third steps of the derivation are generative, because their nonterminal cuts do not contain any communication symbol. The fourth step is communication because of Q_2 in component grammar G_1. Since component grammar G_2 does not contain communication symbol, this communicative step can be succesfuly performed. Then the 5-th and the 6-th steps are generative and the 7-th step corresponds to the succesfuly performed communication between the grammars G_1 and G_3. From the nonterminal cut point of view this derivation can be written as follows:

$$\begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} \begin{pmatrix} S_1 \\ Z_2 \\ Z_3 \end{pmatrix} \begin{pmatrix} S_1 \\ Z_2 \\ Z_3 \end{pmatrix} \begin{pmatrix} Q_2 \\ Z_2 \\ Z_3 \end{pmatrix} \begin{pmatrix} Z_2 \\ S_2 \\ Z_3 \end{pmatrix} \begin{pmatrix} Z_2 \\ Z_2 \\ Z_3 \end{pmatrix} \begin{pmatrix} Q_3 \\ Z_2 \\ Z_3 \end{pmatrix} \begin{pmatrix} Z_3 \\ Z_2 \\ S_3 \end{pmatrix} \begin{pmatrix} \varepsilon \\ Z_2 \\ Z_3 \end{pmatrix}$$

The described $PCGS$ generates, as can easily be shown, the language $L_2 = \{a^{i_1}b^{i_1}a^{i_2}b^{i_1+i_2} \ |i_1, i_2 \in I\!N\}$.

Now, let us stress our attention in more details to the derivations of PCGSs. Let $\mathcal{D}(w) = C_1, C_2, \ldots, C_t$ be a derivation of a word w. With this derivation two sequences of nonterminal cuts could be associated. First one is that of all nonterminal cuts of this derivation. The second one is a sequence containing only communication cuts of the derivation. We will call the sequence of nonterminal cuts the trace of the derivation (resp. trace) and denote $T(\mathcal{D}(w))$. The sequence containing only communication cuts of the derivation will be called the communication sequence of the derivation and will be denoted by $C(\mathcal{D}(w))$.

Realize that for a given $PCGS$ Π it is meaningful to speak about the set of all derivations of Π - $\underline{\mathcal{D}(\Pi)}$, the set of all traces of Π - $\underline{\mathcal{T}(\Pi)}$ and the set of all communication sequences of Π - $\underline{\mathcal{C}(\Pi)}$. For a natural number k, $\mathcal{C}(\Pi, k)$ denotes the set of all communication sequences of Π with at most k communications in it. We note that there is not one-to-one relation between the sets $\mathcal{D}(\Pi)$ and $\mathcal{T}(\Pi)$. To every $d \in \mathcal{D}(\Pi)$ the $T(d) \in \mathcal{T}(\Pi)$ is given unambiguously. But, for any $t \in \mathcal{T}(\Pi)$, there is a set $T^{-1}(t)$ such that for every $d \in \mathcal{T}^{-1}(\Pi)(t)$ $T(d) = t$. The cardinality of the set $\mathcal{T}^{-1}(\Pi)(t)$ can be bounded by a constant that depends on Π and t. The relation between the sets $\mathcal{D}(\Pi)$ and $\mathcal{C}(\Pi)$ is not unambiguous, too. For every $d \in \mathcal{D}(\Pi)$ there exists precisely one $C(d) \in \mathcal{C}(\Pi)$. But there are $c \in \mathcal{C}(\Pi)$ for which the cardinality of the set $\mathcal{C}^{-1}(\Pi)(c) = \{d \in \mathcal{D}(\Pi) | C(d) = c\}$ is bounded by infinity only.

The communication sequence of the derivation from Example 1 is $(Q_2, \varepsilon, \varepsilon)$, $(Q_3, \varepsilon, \varepsilon)$. Let $N(\mathcal{D}(w)) = N(C_1), N(C_2), \ldots, N(C_t)$ be a sequence of nonterminal cuts of a computation $\mathcal{D}(w)$. Let $i, j \in \{1, 2, \ldots, t\}, i < j$ and $N(C_i) = N(C_j)$ holds. Then the sequence of steps corresponding to the subderivation $C_i \rightarrow C_{i+1}, \ldots, C_{j-1} \rightarrow C_j$ form a <u>cycle of the derivation</u>. If none of the nonterminal cuts $N(C_i), \ldots, N(C_j)$ contains a communication symbol, then the cycle of the derivation is called the <u>generative cycle of the derivation</u> (resp. generative cycle).

Let $I = \{t_1, t_2, ..., t_k\}, t_1 < t_2 < ... < t_k$, be the set of all communication steps of $\mathcal{D}(w)$.

- the <u>i-th generative section</u> of $\mathcal{D}(w)$, $1 \leq i \leq k+1$, is the subsequence $t_{i-1}+1$, $t_{i-1}+2, \ldots, t_i - 1$ of derivation steps, $t_0 = 0, t_{k+1} - 1 = t$.
- Let $C_{t_j+1} = (\alpha_1 A_1, \alpha_2 A_2, \ldots, \alpha_n A_n)$ be the configuration just at the beginning of the $j-th$ generative section (after the succesfuly performed $j-th$ communication step; resp. at the beginning of the derivation $\mathcal{D}(w)$).

Let $C_{t_{j+1}} = (\alpha_1 \beta_1 B_1, \alpha_2 \beta_2 B_2, \ldots, \alpha_n \beta_n B_n)$ be that of the end of the j-th generative section.

Then $\underline{g(i,j)(\mathcal{D}(w))}$ is the substring β_i of $\alpha_i \beta_i B_i$. We prefer the abbreviation $g(i,j)$ if it is not misleading. The terminal word w generated in the derivation $\mathcal{D}(w)$ can be composed using some of $g(i,j)(\mathcal{D}(w))'s$. Thus, for a given $\mathcal{D}(w)$ and $i, j \in I\!N$ one can speak about the number of occurrences of $g(i,j)(\mathcal{D}(w))$ in the word w. But, in fact, this number does not depend on the whole derivation. It depends on the

$$\boxed{w = a^3b^3a^2b^5}$$

$$\boxed{g(1,1) = a^3} \boxed{g(2,1) = b^3} \boxed{g(1,2) = a^2} \boxed{g(3,1) = b^2} \boxed{g(3,2) = b^2} \boxed{g(1,3) = b}$$

Figure 2: Values $g(i,j)$ for the word $a^3b^3a^2b^5$

communication sequence $C(\mathcal{D}(w))$ only, so the following denotation is correct.

• Let $CC = (C_1, C_2, \ldots, C_p)$, $C_r \in \{\varepsilon, Q_1, \ldots, Q_m\}^m$ for some natural numbers r, m, p (where m can be considered as the number of component grammars of some $PCGS$ and CC can be considered as communication sequence of $PCGS$ mentioned). Then we will denote by $\underline{n(i,j)(CC)}$ (resp. $n(i,j)$) the number of occurrences of $g(i,j)(\mathcal{D}(w))$ (resp. $g(i,j)$) in w for every $\mathcal{D}(w)$ with communication sequence equal to CC.

The situation for the word $a^3b^3a^2b^5$ from the Example 1 is described on the Figure 2.

The last notion recalled here is the notion of communication complexity measure. This measure corresponds to the number of communications between the component grammars necessary to generate the language. In [San 90] this complexity measure was defined as a number of communication symbols in the derivation. But, such definition does not reflect just the number of communications between the grammars. It may differ from it by a constant factor (that is bounded by the number of component grammars decremented by 1). This difference is caused by such "calls for a communication" in which the "called grammar" obtains the communication symbol, too. We prefer to count the number of realized communications here. There is one more difference in the definition of communication complexity comparing that of [San 90] and ours. In [San 90] it is defined as usual - minimizing over all derivations for a given word and then maximizing over all words of the given length. We require all of the derivations, not only the minimal one, to be bounded by the considered amount of complexity. This definition is essential for our proof of the lower bound on nonconstant communication complexity. All other proofs works also for the above mentioned definition. Now, the formal definition of communication complexity follows.

Definition 3 *Let* Π *be a* $PCGS(m)$ *,* $L = L(\Pi)$ *and* $\mathcal{D}(\Pi)$ *be the set of all derivations of* Π*. Let* $\mathcal{D}(\Pi, w)$ *be the set of all derivations of a word* w *by* Π*. Let* $D_\Pi(w) = C_0, C_1, C_2, \ldots, C_t, C_i = (C_{i1}, C_{i2}, \ldots, C_{im})$ *be a derivation of a terminal word* w *in* Π *and* $I = \{t_1, t_2, ..., t_k | t_1 < t_2 < ... < t_k\}$ *be the set of those communication steps of the derivation* $D_\Pi(w)$ *for which the* $(t_i - 1)$*-st step,* $i \in \{1, \ldots, k\}$*, is generative. Then*

$$\begin{aligned}
\underline{com(D_\Pi(w))} &= \sum_{i=1}^{k} |C_i|_K \\
\underline{com(w, \Pi)} &= max\{com(D) \mid D \in \mathcal{D}(\Pi, w)\} \\
\underline{com(n, \Pi)} &= max\{com(w, \Pi) \mid |w| = n\}.
\end{aligned}$$

Finally,

$$\underline{COM(f(n))} = \{L(\Pi) | \forall n \in I\!N : com(n, \Pi) \leq f(n)\}.$$

Let us denote by $\underline{x - PCGS(m) - f(n)}$ the $PCGS$ of degree m with the communication graph in the class of graphs x and at most $f(n)$ communications during the generation of any word of the length n. We consider $x \in$ \{ *tree, dag(directed acyclic graph), one-way array, c*\} where c is a class of trees with only one non-leaf node. $c - PCGS$s are called centralized [San 90, HKK 93]. We use the notation $x - PCGS - k$ instead of $x - PCGS - g(n)$ if $g(n) = k$ for every $n \in I\!N$.

3 Constant communication complexity

The aim of this section is to show how powerful $PCGS$s are when only constant number of communications are allowed. Here it is shown, that in general, there is an infinite hierarchy of constant communication complexity. The situation is different in the class of languages over one-letter alphabet. We will show that every language over one-letter alphabet that is generated in $PCGS$ with a communication complexity bounded by a constant is regular.

First, we stress our attention to the infinite hierarchy of constant communication complexity. This problem was studied in [HKK 93] for $PCGS$'s with special communication graphs. It was shown there, that there is an infinite hierarchy of constant communication complexity for $PCGS$ whose communication graph are treesx or dags. Here, we give the result for $PCGS$'s with no restriction on communication graph.

Theorem 1 *Let $k \in N$. Then L(PCGS-(k)) - L(PCGS-(k-1)) $\neq \emptyset$.*

Proof. To show the above infinite hierarchy of constant communication complexity, the following language L_k is used for every natural k.

$$L_k = \{a^{i_1}b^{i_1}a^{i_2}b^{i_1+i_2}\dots a^{i_k}b^{i_1+i_2+\cdots+i_k}|i_j \in N\}.$$

The language L_2 of Example 1 is a special case of L_k for $k = 2$. First, the $PCGS\Pi$ generating the language L_k is given by the rewriting rules of the following $k + 1$ component grammars of Π :

$$\begin{aligned}
G_1 \quad &: S_1 \to aS_1|aQ_2 \\
&\quad Z_j \to aZ_j|aQ_{j+1} \\
&\quad Z_k \to b \\
G_i \quad &: S_i \to bZ_{i-1} \\
&\quad Z_j \to bZ_j \\
G_{k+1} &: S_{k+1} \to Z_k \\
&\quad Z_k \to bZ_k, \text{ where } i = 2, \cdots, k,\ j = 1, \cdots, k-1
\end{aligned}$$

As can be easily seen, assumed PCGS Π is centralized. To prove $L(\Pi) = L_k$ is left to the reader. Note only, that the first component grammar G_1 generates the groups of a's, while the component grammar G_{j+1} generates the subword $b^{i_1+i_2+\ldots+i_j}$, $j \in \{1, 2, \ldots, k\}$.

We start the proof of the lower bound giving the following notations describing some important features of the derivation of $PCGS$. For any $PCGS(m)$ Π and any constant $h \in I\!N$, we define:

$\mathcal{S}(h, \Pi)$ is the set of all traces of Π with the following two propereties:

- there are at most h communications in every sequence of $\mathcal{S}(h, \Pi)$;
- each of the nonterminal cuts occurs at most once in one generative section of any sequence of $\mathcal{S}(h, \Pi)$.

$\underline{\mathcal{D}(h,\Pi)}$ is the set of all derivations of Π which have their sequences of nonterminal cuts in $\mathcal{S}(h,\Pi)$.

It is clear that the cardinality of $\mathcal{D}(h,\Pi)$ can be bounded by some constant depending only on Π and h, because the length of derivations of $\mathcal{D}(h,\Pi)$ is bounded by $(h+1)\cdot$ the number of different nonterminal cuts of Π.

Now, for every $PCGS(m)$ Π, $h \in I\!N$ and the word w, $\underline{A(w,h)} = \{w \mid \exists \mathcal{D}(w) \in \mathcal{D}(h,\Pi)\}$.

The constant $D(h)$ bounding the length of words that can be generated by Π with at most h communications and without any repetition of nonterminal cuts during one generative section of any derivation is:

$$D(h) = \max_{w \in A(w,h)} |w|.$$

With respect to this fact, for every word $\omega \in L(\Pi)$, $|\omega| > D(h)$, the minimal derivation using at most h communications has to contain a "cycle of nonterminal cuts" corresponding to one generative section of the derivation.

Based on this idea we will show, that in those steps of the derivation $\mathcal{D}(w)$ (of some Π with $L(\Pi) = L_k$) in which the part of a^{i_1} is generated, subwords of $b^{i_1}, b^{i_1+i_2}, \ldots, b^{i_1+i_2+\cdots+i_k}$ have to be generated in other grammar (or grammars), too. Then, a simple consideration sufficies to realize that at least k communications are needed to have all of this subwords in one -namely the first - component grammar of Π.

Let $PCGS$ Π_k generating L_k have the communication complexity bounded by some $p \in I\!N$. We show that $p \geq k$.

Let $u = a^{i_1}b^{i_1}a^{i_2}b^{i_1+i_2}\ldots a^{i_k}b^{i_1+i_2+\cdots+i_k}$, $i_j > D(p)$ for every $j \in \{1,\ldots,k\}, i_r \neq i_s$ for $r \neq s$ and $\mathcal{D}(u)$ be one of its minimal derivations containing at most p communications. In order to simplify the manipulation with u, let us write it as follows:

$$u = \alpha_1\beta_1\alpha_2\beta_2\ldots\alpha_k\beta_k,\ \alpha_j = a^{i_j},\ \beta_j = b^{i_1+i_2+\cdots+i_j}.$$

Since $i_1 > D(p)$ there exist $j \in \{1,\ldots,m\}$ and the following part P of the generative section of the derivation $\mathcal{D}(u)$

$$P = \begin{pmatrix} w_1M_1 \\ \vdots \\ w_jM_j \\ \vdots \\ w_mM_m \end{pmatrix} \cdots \begin{pmatrix} w_1w_1^{'}M_1 \\ \vdots \\ w_jw_j^{'}M_j \\ \vdots \\ w_mw_m^{'}M_m \end{pmatrix}$$

such that $w_j^{'} \neq \varepsilon$ becomes a part of α_1. So, we can write $\mathcal{D}(u) = XPY$ for some parts of the derivation X, Y.

Let $I = \{s \mid w_s^{'}$ be a part of $u\}$.

By the repetition of the above mentioned part of the derivation (that is possible, because from the nonterminal cut point of view this part forms a cycle), the following derivation $\mathcal{D}(u^{'})$

$$\mathcal{D}(u^{'}) = X \begin{pmatrix} w_1M_1 \\ \vdots \\ w_jM_j \\ \vdots \\ w_mM_m \end{pmatrix} \cdots \begin{pmatrix} w_1w_1^{'}M_1 \\ \vdots \\ w_jw_j^{'}M_j \\ \vdots \\ w_mw_m^{'}M_m \end{pmatrix} \cdots \begin{pmatrix} w_1w_1^{'}w_1^{'}M_1 \\ \vdots \\ w_jw_j^{'}w_j^{'}M_j \\ \vdots \\ w_mw_m^{'}w_m^{'}M_m \end{pmatrix} Y$$

of a terminal word $u^{'} \in L(\Pi_k)$ is obtained. According to our assumption $L_k = L(\Pi_k)$, $u^{'}$ has to be of the form $u^{'} = \alpha_1^{'}\beta_1^{'}\alpha_2^{'}\beta_2^{'} \ldots \alpha_k^{'}\beta_k^{'}$, where

$$|\alpha_1^{'}| - |\alpha_1| = |\beta_1^{'}| - |\beta_1| > 0, \ |\alpha_i^{'}| \geq |\alpha_i|, \ |\beta_i^{'}| > |\beta_i|, \ i \in \{2, \ldots, m\}.$$

Moreover, for every $i \in I, \ w_i^{'} \in \{a\}^* \cup \{b\}^*$. (In the case it is not true, the unbounded repetition of this cycle leads to the unbounded number of "borders" between the a's and b's in the word generated and this contradicts to the structure of L_k.)

The only difference between these two words u, u' is in the number of repetitions of $w_i^{'}, \ i \in I$, in them. Whenever $w_i^{'}, \ i \in I$, is a part of u, $(w_i^{'})^2$ is the part of u'. Since $|\beta_r^{'}| > |\beta_r|, \ r \in \{1, \ldots, k\}$, for every r there is a $j_r, \ j_r \in \{1, \ldots, m\}$ such that $w_{j_r}^{'}$ is a part of β_r.

Because $w_j^{'}$ ($w_j^{'}$ is a part of α_1), $w_{j_1}^{'}, \ldots, w_{j_k}^{'}$ are positioned in others grammars in the configuration $(w_1w_1^{'}w_1^{'}M_1, \ldots, w_mw_m^{'}w_m^{'}M_m)$ of $\mathcal{D}(u)$, at least k communications are necessary for them to become a part of the word u. It implies that $p \geq k$, which completes the proof.

$\square$

Since we do not have any restriction on the type of the communication graph of $PCGS$ Π in the lower bound part of the proof of Theorem 1, we have the following corollary.

Corollary 1 *Let $k \in I\!N$, x be a type of communication graphs and* Π *be a* $x-PCGS$ *generating L_k with communication complexity k. Then*
$\mathcal{L}(x - PCGS - (k)) - \mathcal{L}(x - PCGS - (k-1)) \neq \emptyset$.

Realize that the language L_k is generated in $PCGS$ with communication graph that is centralized, in our proof. Thus, Theorem 4 of [HKK 93] providing the hierarchy of constant communication complexity for $x - PCGS$'s, where $x \in \{tree, c\}$ is a special case of Corollary 1. Moreover, we have the following.

Corollary 2 *Let $k \in I\!N$. Then,* $\mathcal{L}(dag - PCGS(k)) - \mathcal{L}(dag - PCGS(k-1)) \neq \emptyset$.

The next theorem contrasts with the infinite hierarchy proved above. It also documents that, from communication complexity point of view, the gap between regular and nonregular languages is greater when restricted to languages over one-letter alphabet.

Theorem 2 *For every* $k \in I\!N, \mathcal{L}(PCGS - (k)) \cap a^* \subset R$.

<u>Proof.</u> For a given $PCGS - k$ generating a language over one-letter alphabet, the construction of an equivalent regular grammar $\mathcal{G}$ is given. First we recall that, for every derivation, $g(i,j)$ denotes the terminal part x of the word generated in the i-th component grammar during the j-th generative section of the derivation and $n(i,j)$ is the number of occurrences of x in resulting terminal word. Now, the construction of a regular grammar is based on the following fact.

Fact 1 $\forall k \in I\!N$, *$PCGS$ Π and $i, j \in I\!N$ there is a constant $N(k,\Pi)$ such that $n(i,j)(CC) \leq N(k,\Pi)$ for every $CC \in \mathcal{C}(\Pi)$ that contains at most k communications.*

Let $\Pi = (G_1, G_2, \ldots, G_m, K), G_i = (N_i, \{a\}, P_i, S_i)$, be a $PCGS(m) - k$. Then the cardinality of the set of all possible communication sequences $|\mathcal{S}(k, \Pi)|$ is bounded by the number $(\prod_{i=1}^{m}(|N_i \cap K| + 1))^k$. We can suppose there is some numbering on elements of this set.

Informally, the idea of the simulation is the following one. At the beginning $\mathcal{G}$ nondeterministically guesses the communication sequence CC of the simulated derivation by a choice of some rules for its axiom S. Note that it is possible because the communication sequence is of a length bounded by a constant. Obviously, for each symbol a generated by a component grammar in any generative section of the computation, CC provides the information how many times a is included in the resulting word. So, the generative steps of Π can be directly simulated by regular rules producing the corresponding number of a's when the actual nonterminal cut of Π is included in the nonterminal of $\mathcal{G}$. The simulation of communication steps of Π is used to check the correctness of the nondeterministic decision of $\mathcal{G}$ at the beginning of the simulation.

Now, the formal description of the regular grammar $\mathcal{G}$ is given in two steps. First, the set of nonterminal symbols of $\mathcal{G}$ follows, then the set of rewriting rules of $\mathcal{G}$ is described.

Nonterminals of $\mathcal{G}$

The set of nonterminals is composed from the following three subsets:

- $\{S\}$, where S is the axiom of $\mathcal{G}$
- N_1
- $\{[NC, NN, CC, p]$, where

  $NC \in (N_1 \cup \varepsilon) \text{ x } (N_2 \cup \varepsilon) \text{ x } \ldots \text{ x } (N_m \cup \varepsilon)$

  $NN \in \{0, 1, \ldots, N(k, \Pi)\}^m$

  $CC \in \{K \cup \varepsilon\}^k$

  $p \in \{0, 1, \ldots, k+1\}\}$

The meaning of a nonterminal $[NC, NN, CC, p]$ describing an actual configuration of Π is

NC is a nonterminal cut of Π just simulated

NN is an m-tuple of $n(i,j)'s$ corresponding to the generative section of the derivation of Π just simulated

CC is the sequence of communication cuts of the derivation of Π just simulated;

p is the number of the generative section of the derivation of Π just simulated

Rewriting rules of the grammar $\mathcal{G}$

The set of rules P of $\mathcal{G}$ consists of the following sets:

- $\{S \to [(S_1, S_2, \ldots, S_m), (n_1, n_2, \ldots, n_m), CC, t, 0] \mid CC \in \mathcal{C}(\Pi, k),$
  $n_i \in \{0, \ldots, N(k, \Pi)\}$ for $i \in \{1, \ldots, m\}$ and n_i is the value of $n(i, 0)$ for the communication sequence CC and t is the length of $CC\}$.
  This corresponds to guessing the communication sequence of the derivation of Π generating the terminal word. The derivation with this communication sequence will be simulated by the grammar $\mathcal{G}$.

- $\{[(M_1, M_2, \ldots, M_m), (n(1,p), n(2,p), \ldots, n(m,p)), CC, t, p] \to$
  $a^s[(M_1', M_2', \ldots, M_m'), (n(1,p), n(2,p), \ldots, n(m,p)), CC, t, p] \in P$
  for every $p \in \{0, \ldots, t\}$, $CC \in \mathcal{C}(\Pi, k)$, $s \in I\!N$ and nonterminal cuts $(M_1, M_2, \ldots, M_m), (M_1', M_2', \ldots, M_m')$ such that the next five conditions are fulfiled:

  - The nonterminal cut $(M_1, M_2, \ldots, M_m)$ does not contain any communication symbol,
  - $n(i,p) = n(i,p)(CC)$ for $i \in \{1, \ldots, m\}$
  - $M_j \to a^{i_j} M_j' \in P_j$,
  - $s = \sum_{M_j \in N_j - K} (n(p,j) \cdot i_j)$, and

- $M'_{r_z} = M_{r_z}$ for $M_{r_z} = \varepsilon\}$.

This corresponds to the simulation of one generative step in Π according to the nonterminal cut $(M_1, M_2, \ldots, M_m)$.

- $\{[(M_1, M_2, \ldots, M_m), (n(1,p), n(2,p), \ldots, n(m,p)), CC, t, p] \rightarrow$ $[(M'_1, M'_2, \ldots, M'_m), (n(1,p{+}1), n(2,p{+}1), \ldots, n(m,p{+}1)), CC, t, p{+}1]$ for every $p \in \{0, \ldots, t\}$ and nonterminal cuts $(M_1, M_2, \ldots, M_m), (M'_1, M'_2, \ldots, M'_m)$ such that the next five conditions are fulfiled:

  - $n(i,p) = n(i,p)(CC)$ for $i \in \{1, \ldots, m\}$
  - the nonterminal cut $(M_1, M_2, \ldots, M_m)$ corresponds to the $p-th$ member of the communication sequence CC,
  - $M'_j = M_j$ if $(M_j \in \{N_j \cup \varepsilon\}$ & $M_s \neq Q_j \forall s)$,
  - $M'_j = S_j$ if $\exists s\ M_s = Q_j$ & $M_s \notin K$, and
  - $M'_j = M_r$ if $M_j = Q_r$ & $M_r \notin K\}$.

This corresponds to the succesfuly performed communication step of the simulated derivation of Π.

- $\{[(M_1, M_2, \ldots, M_m), (n(1,p), n(2,p), \ldots, n(m,p)), CC, t, t+1] \rightarrow a^{i_1} M'_1 | a^{i_2}$ for every $p \in \{0, \ldots, t\}, CC \in \mathcal{C}(\Pi, k)$ and nonterminal cut $(M_1, M_2, \ldots, M_m)$ such that the following is true:

  - $n(i,p) = n(i,p)(CC)$ for $i \in \{1, \ldots, m\}$
  - $M_1 \rightarrow a^{i_1} M'_1 | a^{i_2} \in P_1\}$

This corresponds to the simulation of the first generative step in the last generative section of the derivation simulated.

- $\{[(\varepsilon, M_2, \ldots, M_m), (n(1,t+1), n(2,t+1), \ldots, n(m,t+1)), CC, t, t+1] \rightarrow \varepsilon$ for $CC \in \mathcal{C}(\Pi, k)$, nonterminal cut $(\varepsilon, M_2, \ldots, M_m)$ and $n(i,t+1) = n(i,t+1)(CC), i \in \{1, \ldots, m\}\}$.

This corresponds to the situation that the derivation finishes with a communication step in which the communication symbol of the first component grammar is replaced by a terminal word.

- P_1.

¿From the above construction of the grammar $\mathcal{G}$ it is clear that $L(\mathcal{G}) = L(\Pi)$ holds. Hence, $PCGS$'s with constant communication complexity and one terminal symbol can generate only a subclass of regular languages.

□

Since every regular language can be generated by some PCGS(1), we have the following.

Corollary 3 $\bigcup_{k \in I\!N} \mathcal{L}(PCGS-(k)) \ \cap a^* = R \cap a^*$.

4 Nonconstant communication complexity

In this subsection we are interested in the communication complexity that cannot be bounded by any constant. First, the bound separating the regular and nonregular languages over one-letter alphabet in terms of communication complexity, is given. This bound - $\Omega(\log n)$ - is shown to be the sufficient and necessary condition to obtain a nonregular language that cannot be generated with communication complexity bounded by any constant.

We stress our attention to a derivation of $PCGS$ in order to make clear the relation of individual $g(i,j)$'s to the terminal word generated. For our purposes the most important moments of the derivation are the communication steps. In every communication step some of $g(i,j)$'s change their position by moving from one component grammar to another one, some of them multiply their occurrences (when moved to at least two component grammars in one communication step). Moreover, in every component grammar the new one $g(i,j)$ starts its existence.The terminal word is composed from those $g(i,j)$'s (resp. their manifold copies) that are positioned at the first component grammar when the derivation of the terminal word finishes.

For a given derivation $\mathcal{D}(w)$ the motion of a single $g(i,j)$'s during $\mathcal{D}(w)$ can be expressed by the directed graph. The graph containing the motion of all $g(i,j)$'s during $\mathcal{D}(w)$ is an analogy to that of a computational tree. On the level l the vertices of the graph correspond to individual occurrences of $g(i,j)$'s. The names of the vertices reflect to the number of component grammar at which the represented $g(i,j)$ is located after $l-th$ communication step (during the $(l+1)-st$ generative section of the derivation $\mathcal{D}(w)$). Since after every communication step the new $g(i,j)$ starts to be generated in every component grammar, new m vertices are added to this graph on every level. For every vertex v with the name $new(x)$, $x \in I\!N$ the path from v to a leaf vertex corresponds to a path in the communication graph of Π and reflects to the movement of that $g(i,j)$ thats creation is represented by v according to the derivation $\mathcal{D}(w)$.

We call this graph the communication dag of the derivation. For a $PCGS$ of m components, the communication dag is acyclic with only m vertices on each "level" with the input degree greater than one. We use the notion "root" to indicate the (unique) vertex with the indegree 0, and the notion "leaf" when refering to the vertex with zero outdegree.

Now, we explain the meaning of some denotations used in the following formal definition of the communication dag. All verteces, except of the root, are named by natural numbers resp. by $new(s), s \in I\!N$. Then, $num(u)$ indicates the number assigned to the vertex u, resp. value s in the case u is named $new(s)$. As usual, $d(x,y)$ denotes the distance from a vertex x to a vertex y.

Definition 4 *Let* $\Pi = (G_1, G_2, \ldots, G_m, K)$ *be a* $PCGS$*, and let* $\mathcal{D}(w)$ *be a derivation of a terminal word* $w \in L(\Pi)$*. Let* $CC(w)$ *be the sequence of communication cuts of* $\mathcal{D}(w)$ *and* $|CC(w)| = l(w)$*. Then the communication dag of the derivation* $\mathcal{D}(w)$ *-* $cdag(\mathcal{D}(w))$ *is the directed graph (with the edges directed from the fathers to the sons) obtained by the following procedure:*

input*:* $\mathcal{D}(w)$*, resp.* $CC(w)$

output*: the communication dag* $cdag(\mathcal{D}(w))$

1 take a new vertex r *and assign the name root to it;*

2 make m *new verteces* $v_1, v_2, \ldots, v_m$ *with the names* $new(1), new(2), \ldots, new(m)$ *the sons of the vertex* r *(with the name root);*

3 `for` *h:=1* `to` $l(w)$

`do`

$S(h) = \{v \mid d(v,r) = h\}$;

$\{(C_{1h}, C_{2h}, \ldots, C_{mh})$ *be the h-th member of the sequence* $CC(w)\}$

for every $u \in S(h)$

`do`

- $K(u) = \{k \mid C_{kh} = Q_{num(u)}\}$
- `if` $K(u) \neq \emptyset$ & $C_{num(u)h} \notin K$
  `then` $\forall k \in K(u)$ *make new vertex with the name* k *a son of the vertex* u
  `else` *make new vertex with the name* $num(u)$ *a son of* u

`od`*;*

`for` $s := 1$ `to` m

`do`

- *take the new one vertex* $v(s)$ *and assign the name* $new(s)$ *to it;*
- $\forall u \in S(h) : num(u) = s$ *make* u *a father of* $v(s)$

`od`

`od`

How are the $g(i,j)$'s related to the terminal word generated and the communication dag described? Every leaf with the name 1, resp. $new(1)$, corresponds to some $g(i,j)$, that is a part of the terminal word generated. The appropriate values i, j can be found as follows:

Let v be a leaf with the name $new(1)$. Then it corresponds to unique occurrence of $g(1, l(w)+1)$.

Let v be a leaf with the name 1 and $v = v_{l(w)+1}, v_{l(w)}, \ldots, v_t$ be the shortest possible sequence of the vertices such that

- v_t is named $new(s)$ for some $s \in \{1, 2, \ldots, m\}$.
- v_i is a father of v_{i+1} in $cdag(\mathcal{D}(w)), i \in \{t, ..., l(w)\}$.

Then v corresponds to one of the occurrences (possible not unique) of $g(s,t)$.

$cdag(\mathcal{D}(w))$ is the graph with the distance from the leaf to the sole root bounded by the number of communication steps. The uotdegree of every vertex of $cdag(\mathcal{D}(w))$ is bounded by the number of components grammar increased by one. Hence, the following observation holds.

Observation 1 *Let Π be a $PCGS(m) - f(n)$ and let $\mathcal{D}(w)$ be one of its derivation of the terminal word w. Then there are at most $(m+1)^{f(|w|)+1}$ leaves in the communication graph $cdag(\mathcal{D}(w))$.*

Lemma 1 *Let Π be a $PCGS(m) - f(n), f(n) \in o(\log n)$. Then there is $n_0 \in I\!N$ such that for every word $w \in L(\Pi), |w| \geq n_0$ each its derivation of $\mathcal{D}(\Pi, w)$ uses at least one generative cycle.*

Proof. We start the proof of the lemma by bounding the maximal length of the word that can be generated in the generative sequence without the repetition of nonterminal cuts.

Let n_i denote the number of noncommunication nonterminals in the $i-th$ component grammar of Π. By d we denote the maximal length of the right side of the rewriting rules taken over individual component grammars. Since $B = \prod_{i=1}^{m}(n_i + 1)$ is the number of all nonterminal cuts, B is the maximal number of generative steps without repetition of some nonterminal cut, too. So, the length of the word generated in one component grammar during one generative section without the repetition of nonterminal cuts is bounded by

$$d \cdot \prod_{i=1}^{m}(n_i + 1). \tag{1}$$

The length of the terminal word generated can be expressed as follows:

$$|w| = \sum_{\substack{i=1,\ldots,m \\ j=0,\ldots,l(w)+1}} n(i,j) \cdot |g(i,j)|. \tag{2}$$

With respect to (1), $|w|$ can be bounded by

$$|w| \leq \sum_{\substack{i=1,\ldots,m \\ j=0,\ldots,l(w)+1}} n(i,j) \cdot \underbrace{(d \cdot \prod_{i=1}^{m}(n_i+1))}_{D} = \sum_{\substack{i=1,\ldots,m \\ j=0,\ldots,l(w)+1}} n(i,j) \cdot D. \tag{3}$$

According to the Observation 1 $\sum_{\substack{i=1,\ldots,m \\ j=0,\ldots,l(w)+1}} n(i,j)$ is bounded by $(m+1)^{f(|w|)+1}$. So we have

$$n \leq D \cdot (m+1)^{f(|w|)+1}, D = d \cdot \prod_{i=1}^{m}(n_i+1) \tag{4}$$

$$\log n \leq \log D + (f(n)+1) \cdot \log(m+1) \leq 2 \cdot (\log D + \log(m+1)) \cdot f(n)$$

$$\frac{\log n}{2 \cdot (\log D + \log m)} \leq f(n). \tag{5}$$

The following fact follows from (5). If the derivation contains only $f(n)$ communications and every nonterminal cut is at most once in one generative section, then $f(n) \in \Omega(\log n)$. Hence, if $f(n) \in o(\log n)$, then $\exists\ n_0 \in I\!N$ such that $\forall n \geq n_0$

$$n > D \cdot (m+1)^{f(n)+1}$$

and there is no way to generate the word with the length at least n without any generative cycle.

□

Now, we are able to show that there are no languages with the nonconstant communication complexity in $COM(o(\log n))$.

Theorem 3 *For every $PCGS(m) - f(n)$ generating a language $L \notin COM(O(1))$, $f(n) \in \Omega(\log n)$.*

Proof. To prove Theorem 3 we use the idea used to show, that there is a gap between $SPACE(1)$ and $SPACE(\log\log n)$ [HU 69]. Let Π be a $PCGS$ generating a language $L(\Pi) \notin COM(O(1))$. In what follows only minimal derivations of Π, i. e. the derivations for which every repetition of any cycle of nonterminal cuts leads to a longer terminal word, are considered.

Since $L(\Pi) \notin COM(O(1))$, L is an infinite language. Moreover, for every $k \in I\!N$ there exists at least one word $w \in L(\Pi)$ such that one of its derivations contains at least k communications. For every $k \in I\!N$ one can fixed one of them in the following way:

w_k is one of the shortest words with $\mathcal{D}(\Pi, w_k)$ involving a derivation $\mathcal{D}(w_k)$ containing at least k communications .

Suppose, for the contradiction, $f(n) = o(\log n)$. Then, according to Lemma 1, there is a constant $n_0 \in I\!N$ such that $\forall w \in L(\Pi)$, $|w| \geq n_0$, every derivation of w has to contain at least one generative cycle.

So, let $k_0 \in I\!N$ be such that $|w_{k_0}| \geq n_0$. Then $\mathcal{D}(w_{k_0})$ has to contain a generative cycle. Since only minimal derivations are considered, removing one repetition of the mentioned cycle the word $w_{k'_0}$ that is shorter than w_{k_0} is generated. But its derivation -not necessary the maximal one- contains the same number of communications as $\mathcal{D}(w_{k_0})$. This contradicts to the assumption that w_{k_0} is one of the shortest words using k_0 communications in one of its derivations.

$\square$

Corollary 4 *Let* Π *be a* $PCGS(m) - f(n)$ *generating a nonregular language over one-letter alphabet. Then* $f(n) \in \Omega(\log n)$.

To show hierarchy of nonconstant communication complexity the following language $L(k)$ is used:
$L(k) = \{w \in \{a\}^+ \mid |w| = d^k, d \in I\!N\}$

Theorem 4 *For every* $k \in I\!N$ *and every function* $f(n) \in o(\sqrt[k]{n})$,

$$L(k) \notin COM(f(n)).$$

Proof: Suppose, for a contradiction, $f(n) = o(\sqrt[k]{n})$, and there is a $PCGS(m) - f(n)$ Π generating the language $L(k)$. For a long enough word $w = a^{d^k}$ let $\mathcal{D}(w) = C_0, C_1, \ldots, C_t$ be the sequence of the configurations that corresponds to one of its derivations. Denote by $T = \{i_1, i_2, \ldots, i_d\}$ the sequence of those time units that $i_j = \min\{t \mid C_{t-1} \rightarrow C_t$ is a communication step and $|C_t|_a \geq j^k\}$ (we denote here by

$|C|_a$ the number of a's in the configuration C). T can be composed from $f(n)$ (not necessary nonempty) sets $T_1, \ldots, T_{f(n)}$ for which

- $T_t = \{i_{t_1}, i_{t_1+1}, \ldots, i_{t_1+\Delta_t}\}$

- $i_{t_1} = i_{t_1+\Delta_t}$

- $t \in \{1, \ldots, f(n)\}$

is true. Since $f(n) = o(\sqrt[k]{n}) = o(d)$, it is not possible to bound $\Delta_t, t \in \{1, \ldots, f(n)\}$ and $n \in I\!N$, by some constant that depends on k only. Hence, for every $c \in I\!N$ there exists $d \in I\!N$ such, that

$$\exists j, t \in \{1, \ldots, f(d^k)\} : T_t = \{i_j, \ldots, i_{j+\Delta_t}\} \texttt{ and } \mathtt{\Delta_t} \geq \mathtt{c}.$$

It implies that

$$|C_{i_{j-1}}|_a < j^k \texttt{ and } |C_{i_j}|_a \ > (j+c)^k. \tag{6}$$

Since in one communication step the number of symbols in $PCGS$ can be increased by at most $(m-1)\cdot$ the length of the longest component of the configuration just before this step, $|C_j|_a \leq (m-1) \cdot |C_{i_{j-1}}|_a$. With respect to this it follows from (6) that, there were some generative steps between the time units i_{j-1} and i_j. In these generative steps at least $(j+c)^k - m \cdot (j^k)$ symbols where generated. If D is the constant bounding the right side of the rewriting rule of Π, then it is possible to choose the constant c in such a way that

$$\frac{(j+c)^k - m \cdot (j^k)}{D \cdot m} > \textit{number of different nonterminal cuts of } \Pi. \tag{7}$$

¿From (7) it follows that there is generative cycle contained in the derivation $\mathcal{D}(w)$. This cycle lies somewhere between the configurations $C_{i_{j-1}}$ and C_{i_j}. One repetition of this cycle increases the length of the terminal word generated by $\triangle$. Hence, for every $p \in I\!N$, the word $a^{d^k + p\cdot\triangle}$ belongs to $L(\Pi)$. But $\exists p_0$ $a^{d^k + p_0\cdot\triangle} \notin L_k$. This contradiction completes our proof.

$\square$

Theorem 5 *For every $k \in I\!N, L(k) \in COM(O(\sqrt[k]{n}))$.*

To prove this theorem some new notions and technical lemmas are given.

Definition 5 *Let $f(p)$ be an increasing function from $I\!N$ to $I\!N$. We say $f(p)$ is* *<u>PCGS-countable</u> if there exists a $PCGS$ Π such that the following conditions are true:*

1, $L(\Pi) = \{a^{f(p)} \, | p \in I\!N\}$

2, $\forall p \in I\!N$ *there is unique derivation* $\mathcal{D}(p)$ *of the word* $a^{f(p)}$

3, there exists $r \in I\!N$ *such that for every* $p > r$
$\mathcal{D}(p) = F_0, F_1, \ldots, F_s, (C_1, \ldots, C_d)^{p-r}, E_0, E_1, \ldots, E_e$, *where*

- F_0 *is starting configuration of* $PCGS$ Π
- $F_1, \ldots, F_s, E_0, E_1, \ldots, E_e$ *and* $C_1, \ldots, C_d$ *are nonterminal cuts*
- $s, d, e \in I\!N$ *are constant dependidng on* Π, f

4, There is a constant z such that after i repetitions of the cycle $(C_1, \ldots, C_d)$ of $\mathcal{D}(p)$ satisfying 3, the terminal part of the component grammar G_1 is $a^{f(i+r)-z}$.

The mentioned $PCGS$ is said to count the function $f(p)$.

Informally, there is a possibility for Π to "decide" to finish the derivation of the word without the next repetition of the cycle or to continue in the derivation of the "longer" terminal word with the repetition of the cycle. The sequence of nonterminal cuts $l_1, ..., l_{j_n}$ causes the terminal part of component grammar G_1 will be prolonged by z symbols and become the terminal word of the right length.

Fact 2 *Let $f(p)$ is $PCGS-countablefunction$ and Π is a $PCGS$, that counts $f(p)$. Then there is a constant $d(\Pi, f)$ such, that for every a^n, $n = f(p)$, number of communications in the derivation $\mathcal{D}(w, \Pi)$ is bounded by $d(\Pi, f) \cdot p$.*

Lemma 2 *Let Π be the $PCGS$ counting the function $f(p)$ and $k \in I\!N$ be a constant. Then the function $g(p) = k \cdot f(p)$ is PCGS-countable.*

<u>Proof:</u> It is sufficient to replace every rewriting rule of Π of the form $A \to a^i B$ by the rule $A \to a^{i \cdot k} B$, where A, B are nonterminals and a is the terminal of Π.

$\square$

Lemma 3 *Let $f(p)$ be a PCGS-countable function and $g(p)$ be the function given by the following prescription:*

$g(1) = c$

$g(p+1) = g(p) + f(p) + d$, *for some* $c, d \in I\!N$.

Then the function $g(p)$ is PCGS-countable.

Proof. Since the formal description of the $PCGS$ Π_g would be confusing we omit it here. We prefer to informally describe how such a system can be constructed. Let $\Pi_f = (G'_1, \ldots, G'_m, K)$ be a $PCGS(m)$, that counts the function $f(p)$. We shall describe the $PCGS(m)$ Π_g, that counts the function $g(p)$. The work of this system will simulate the system Π_f in one group of the grammars, in other group of the grammars the actual value of $g(p)$ will be remembered, the third group of the grammars will secure the addition of $f(p) + g(p)$ and the fourth one helps the grammars to contain their starting symbol at the moment it is necessary, as described on the Figure 3.

G_1	is the grammar for remembering the value of $g(p)$
$F_1, \ldots, F_m$	are the grammars simulating the modificated work of the system Π_f
P_1, P_2	are the grammars used to remember the value of $f(p)$
$C_1, \ldots, C_t$	are the grammars used to secure some of the component grammar to contain its starting symbol
$\boxed{x}$	represents the terminal string a^x
Q_{name}	indicates the communication symbol that implies the communication with the commponent grammar *name* (analogous for nonterminal symbols)
nonterminals	are depicted only in the case, they are necessary.

On the Figure 3 only those parts of the derivation are described that are necessary to explain the derivation as a whole. Since the rewriting rules of the type $A \to B$, where A, B are nonterminals, are often refered in what follows, we will say the grammar is active standing when applying the rewriting rule of that type.

The derivation of the terminal word $a^{g(p)}$ (resp.$a^{f(p)}$) can be according to the Figure 3 described as follows:

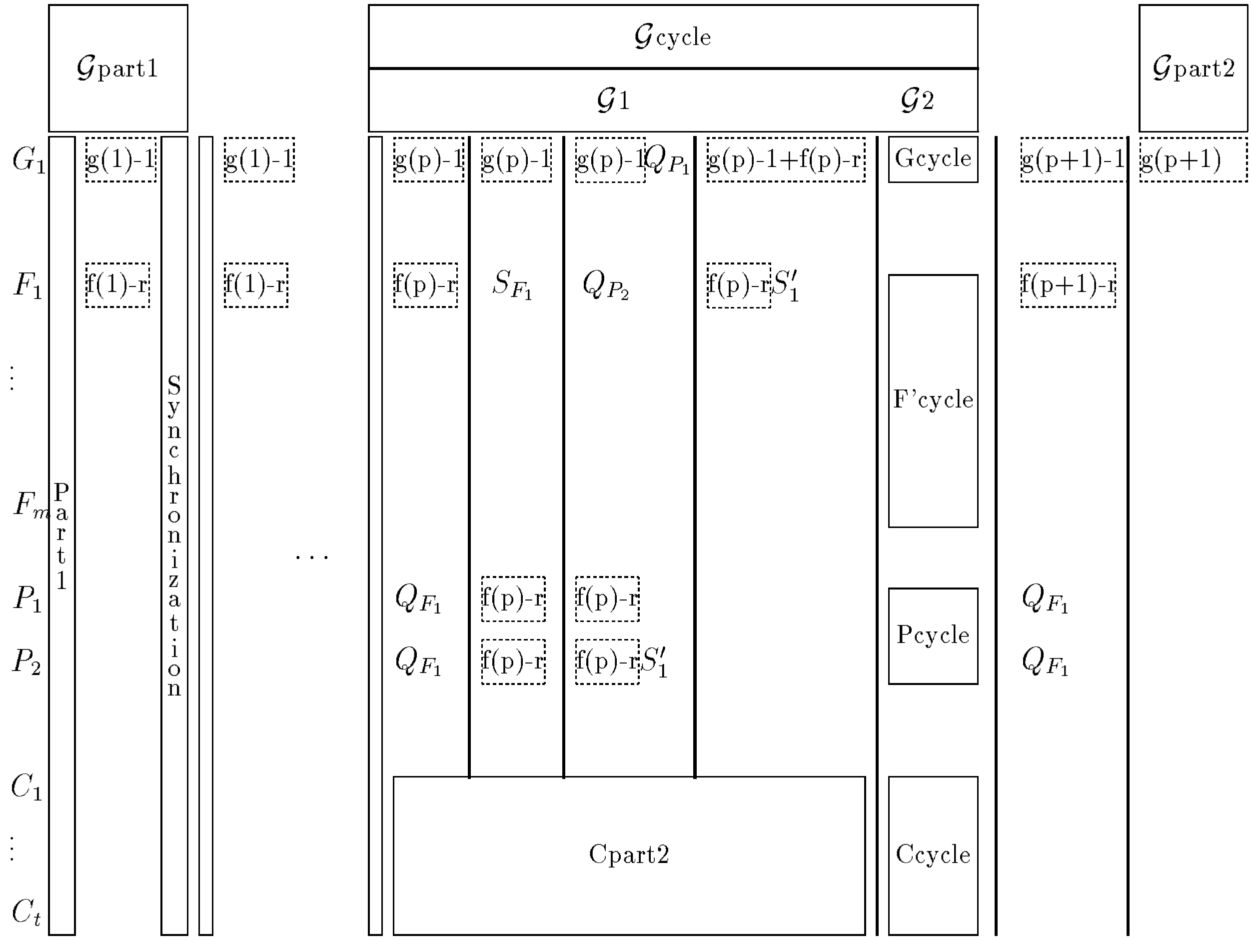

Figure 3: Schema of a derivation of Π_g.

$\boxed{\text{Fpart1}}\ \boxed{\text{Fcycle}}^{\star}\ \boxed{\text{Fpart2}}$

$\boxed{\mathcal{G}part1}\ \boxed{\mathcal{G}cycle}^{\star}\ \boxed{\mathcal{G}part2}$, where

- $\mathcal{G}part1(Fpart1)$ corresponds to the sequence of nonterminal cuts $F_0, F_1, \ldots, F_s$
- $\mathcal{G}cycle(Fcycle)$ corresponds to the sequence of nonterminal cuts creating the cycle $C_0, C_1, \ldots, C_d$
- $\mathcal{G}part2(Fpart2)$ corresponds to the sequence of nonterminal cuts $E_0, E_1, \ldots, E_e$ from the Definition 5.

While on the Figure 3 the last repetition of the $\mathcal{G}cycle$ is depicted, the more precise description of the derivation of the word $a^{g(p)}$ is following:

$$\boxed{\mathcal{G}part1} \cdots \boxed{\mathcal{G}cycle}\boxed{\mathcal{G}part2}\ .$$

After the $\mathcal{G}$part1 G_1 contains $(g(1)-1)$ a's, $F_1,\ldots,F_m$ have the same contents as $G'_1,\ldots,G'_m$ generating in Π, e.g. F_1 has $(f(1)-r)$ a's generated. The grammars $P_1,P_2,C_1,\ldots,C_t$ are active standing during this steps.

Synchronization is a part of the derivation in which all grammars are active standing. This part of the derivation is used to obtain nonterminal cut corresponding to the beginning of the $\mathcal{G}cycle$. This nonterminal cut is indicated by $\Box$.

$\mathcal{G}$cycle is composed from two parts. The first one - $\mathcal{G}1$ - corresponds to the adition of $g(n)$ and $f(n)$. The second one - $\mathcal{G}2$ - guarantees the system of grammars $F_1,\ldots,F_m$ work in such a way, that the result of this part of the derivation is the same as the repetition of the cycle $Fcycle$ in the derivation of Π_f.

The last step of the derivation (that is named $\mathcal{G}part2$) is generative and leads to a final terminal word.

Now, let us go to the $F'cycle$, $Ppart$s and $Cpart$s in more details.

Why the grammars $F_1,\ldots,F_m$ have to work in a different way during the $Fcycle'$ than during the $Fcycle$? The reason is, that whenever the grammar F_1 is called (the nonterminal Q_{F_1} is present in a nonterminal cut) the rewriting rule $S_{F_1}\rightarrow Q_{P_2}$ is apllied according to the $\mathcal{G}1$. But this rewriting step was not present in the $Fcycle$ and we have to change the rewriting rule of the $PCGS$ Π_f $S_{G'_1}\rightarrow Z'_1$ to $S_{F_1}\rightarrow Q_{P_2}$. To be able to continue in the modified $Fcycle$ (in Figure 4 it is $Fcycle'$), the grammar P_2 has to contain exactly the string Z as its phrase at that moment. This can be guaranted by active standing of the grammar P_2 (that is active standing whole derivation) and the group of the grammars $C_1,\ldots,C_t$, that are active standing all the time, too. When it is necessary, calling P_1, resp. P_2, they caused P_1, resp. P_2 to start from its starting symbol. All others grammars $F_2,\ldots,F_m$ must have their rewriting rules changed (in comparison with $G'_2,\ldots,G'_m$), too. Let $(S_1,Z_2,\ldots,Z_m)$ is the nonterminal cut corresponding to the situation G'_1 contains its nonterminal during the $Fcycle$ and $(Z'_1,Z'_2,\ldots,Z'_m)$ the next one. Let in this generative step rewriting rules $Z_i \rightarrow a^{j_i}Z'_i, i \in \{2,\dots,m\}$ $Z_i \neq \varepsilon$ *and* $j_i \in I\!N$ were applied. Then those of the grammars $F_2,\dots,F_m$ for which $Z_i \neq \varepsilon$ will have the rewriting rules $Z_i \rightarrow Z''_i, Z''_i \rightarrow a^{j_i}Z'_i$ instead of mentioned rules of $PCGS$ Π_f, where $Z''_2,\dots,Z''_m$ are new (with respect to the nonterminals of $G'_j, j \in \{1,\dots,m\}$) nonterminals.

During the *Gcycle* r a's are generated in G_1.

□

Fact 3 *There are PCGS's Z_1,resp. $\overline{Z}_1$, counting the function $f(p) = p$,resp. $f(p) = 2p$.*

Proof of the Theorem 5. Now, for a given $k \in I\!N$, we are ready to describe the $PCGS - h(n)$ Z_k generating the language L_k. Then, the function $h(n)$ bounding the communication complexity of the constructed Z_k will be shown to be in $O(\sqrt[k]{n})$. The inductive construction of Z_k is based on the binomial theorem:

$$(p+1)^{k+1} = p^{k+1} + a_1p^k + \cdots + a_kp + 1, a_i = \binom{\text{k+1}}{\text{i}}$$

According to this theorem the value of the function x^{k+1} at the point $x = p+1$ can be count knowing the values of the functions $x^{k+1}, x^k, \dots, x$ at the point $x = p$.

For $k = 1$ we have $(p+1)^2 = p^2 + 2p + 1$. Use the Lemma 3 with $f(p) = 2p, c = 1, d = 1$. Then there is a $PCGS$ Z_2 counting the function p^2. Suppose, the systems $Z_1, Z_2, \dots, Z_{k-1}$ are those counting the functions $p, p^2, \dots, p^{k-1}$, respectively. Based on the Lemma 2 there are $PCGS$'s $Z'_1, Z'_2, \dots, Z'_{k-1}$ counting the functions $\binom{\text{k-1}}{1}p^{k-1}, \binom{\text{k-1}}{2}p^{k-2}, \dots, \binom{\text{k-1}}{\text{k-1}}p$, respectively. Following the Lemma 3 the system Z_{k+1} counting the function p^{k+1} can be obtained.

With respect to the Fact 2 the number of communications during the derivation of the terminal word linearly depends on the number of the repetition of the cycle. In the case of our language the cycle is repeated $(d-1)$ times, where $d = \sqrt[k]{n}$. So the communication complexity can be bound by some function $h(n) \in O(\sqrt[k]{n})$.

□

Theorems 4 and 5 show we cannot find the function bounding the communication complexity of all languages that are $PCGS$-generable. Nor in the general nor in the case of the languages over one-letter alphabet. But we see the number of component grammars rapidly increases with k. It would be interesting to show a hierarchy of communication complexity inside the $\mathcal{L}(PCGS(m))$ for some $m \in I\!N$. Another one problem that remains open is to show that there is (if any) such condition α that for arbitrary two functions $f_1(n), f_2(n)$ satisfying α (e.g. $f_2(n) = o(f_1(n))$) $COM(O(f_1(n))) - COM(O(f_2(n))) \neq \emptyset$ holds.

Acknowledgements

The author wish to thank J.Hromkovič for helpful discussions and comments to the final form of the paper.

References

[PS 89] Gh.Paun,L.Santean:Parallel communicating grammar systems: the regular case. *Ann.Univ.Buc.Ser.Mat.-Inform.*37 vol.2(1989),pp.55-63.

[HU 69] J.E.Hopcroft, J.D.Ullman: Formal Languages and Their Relation to Automata. Addison-Wesley Publishing Company, Reading, Massachusetts, 1969.

[HR 75] G.T.Herman, G.Rozenberg: Developmental Systems and Languages, North-Holland, Amsterdam, 1975.

[HKK 93] J.Hromkovič, J.Kari, L.Kari: Some hierarchies for the communication complexity measures of cooperating grammar systems. *Theoretical Computer Science*, to appear (extended abstract in: Proc. of the MFCS'93 Lecture Notes in Comp. Science 711, pp.495 - 505.)

[Par 92] D.Pardubská: The communication complexity hierarchy of parallel communicating systems. Presented at IMYCS'92.

[Par 93] D.Pardubská: On the power of communication structure for distributive generation of languages. In: Developments in Language Theory. Preproceedings, University of Turku, Turku 1993, pp.30 - 32.

[RS 80] G.Rozenberg, A.Salomaa: The Mathematical Theory of L Systems. Academic Press, 1980.

[San 90] L.Santean:Parallel Communicating Systems. *EATCS Buletin*, Num.42 (1990), 160 - 171.

[SK 92] L.Santean, J.Kari: The impact of the number of cooperati ve grammars on the generative power, *Theoretical Computer Science* (1992) pp.249-263.