Proceedings of Grammar Systems Week 2004, edited by E. Csuhaj-Varjú and Gy. Vaszil,
MTA SZTAKI, Budapest, pages 113 - 125, 2004.

On the Time, Space and Communication Complexity of Cooperating Distributed Grammar Systems

Liliana Cojocaru

Rovira i Virgili University of Tarragona
Pl. Imperial Tàrraco 1, 43005 Tarragona, Spain
liliana.cojocaru@estudiants.urv.es

Abstract

In this paper we deal with trade-offs between *time, space,* and *communication complexity* of languages generated by Cooperating Distributed Grammar Systems (henceforth CDGS) with regular, linear and context-free components. We propose two types of communication structures. The first structure is determined by the *communication graph* of the CDGS, i.e, a directed graph in which the vertices are labeled by the CDGS components and the directed edges correspond to pairs of grammars (G_a, G_b), $a \neq b$, that communicate with each other. The communication is done through those nonterminals that appear on the right side of a production of G_a and on the left side of a production of G_b, according to the protocol of cooperation used by the system. We will refer to these nonterminals as *communicational nonterminals.* The second structure is given by the communicational protocol tree attached to the generated language, i.e. the derivation tree of a special kind of *Szilard language* introduced in this paper, called *communicational Szilard language.* A *communication complexity* measure, i.e. how many times the system components communicate with each other using a minimal number of communicational nonterminals, is defined and studied depending on modes of derivation, weak and strong fairness conditions. We found that the communication complexity of weakly and strongly q-fair languages in the case of grammar systems with regular or linear components is linear. These languages can be accepted by a nondeterministic multitape Turing machine in linear space and linear time. In the case of languages generated by CDGS with context-free components we show that the communication complexity varies from linear to logarithmic. The space required by a nondeterministic multitape Turing machine to accept these languages equals to the communication complexity, while the time is linear in all the cases.

1 Introduction

Communication complexity is one of the youngest branches of complexity theory. It has been inspired from the inter-processor communication, in which the input, viewed as sequences of messages distributed among different parts of a system, has to be split in different partitions in order to allow an optimal communication. It has been introduced in 1979 by Yao [17], and it studies the communication exchanged during a computational process by *minimizing the amount of information* exchanged between the system components. The flow of information (the number of exchanged bits) is measured by ignoring the other cost, such as time and space.

This paper is devoted to the communication complexity of Cooperating Distributed Grammar Systems, with regular, linear and context-free components, but concerns also other computational resources used by the system, such as time and space. We deal with trade-offs between this three measures in order to control the generative process underlined by CDGS.

Investigations related to the communication complexity of distributed grammar systems have been done so far in several papers, e.g. [9], [10], [11], [12], [13]. In these papers the authors focus on the communication complexity of *Parallel Communicating Grammar Systems* (henceforth *PCGS*). They have considered two kinds of communication complexity measures. The first one is the *communication structure of PCGS*, i.e. the shape of the communication graph, consisting of directed communication links between the grammars, while the second one is a *communication complexity measure*, i.e. the number of exchanged messages during the computational process.

For the case of CDGS we propose two types of communication structures. The *first structure* is determined by the *communication graph* of CDGS, which is a directed graph where the vertices are labeled by the CDGS components and the directed edges correspond to pairs (G_a, G_b), $a \neq b$, of grammars that communicate with each other. The communication is done through those nonterminals that appear on the right side of a production of G_a and on the left side of a production of G_b, according to the protocol of cooperation used by the system. We refer to these nonterminals as *communicational nonterminals.* The *second structure* is a protocol tree determined by the interconnection between the system components, i.e. the way in which they bring consecutive contributions on the sentential form during the language generation process. We define a new kind of language, called *communicational Szilard language.* The derivation tree of a certain word γ_w^c from this language is the *communicational protocol tree* attached to the corresponding word w from the language generated by the system, for which γ_w^c is the *communicational control word* of w. A *communication complexity measure*, i.e., how many times the system components communicate with each other using a minimal number of communicational nonterminals, is defined and studied depending on modes of derivation, weak and strong fairness conditions.

2 Prerequisites

Grammar systems are sets of grammars that function together according to a specified protocol of cooperation. In CDGS all the components have a common axiom, all grammars have the same working tape, each of them making their own contribution on the common sentential form. At each moment only one grammar is active. Which component of the system is active at a given moment, and when a grammar stops to be active, is decided by the protocol of cooperation. This protocol consists in stop conditions such as modes of derivation (how many times a rewriting rule of the same component can be applied), in weak fairness conditions (each component has to be activated almost the same number of times) or in strong fairness conditions (each component has to be activated almost the same number of times, by taking into account the number of internal productions that are applied for each grammar). CDGS simulates the blackboard model of problem solving, in which the blackboard is the common working tape, and the components $G_1, ..., G_r$ are the knowledge sources (agents, processors, etc.). They have been introduced and studied in [2], [3], [4] and [8], with forerunner related papers [14] and [15]. Formally a CDGS is defined as follows:

Definition 1 A **Cooperating Distributed Grammar System** of degree r, $r \geq 1$ is a construct of the form: $$\Gamma = (N, T, S, P_1, \ldots, P_r), \quad (1)$$
where the set N and T are disjoint finite alphabets, the *nonterminal* and the *terminal* alphabet, respectively. $S \in N$ is the system axiom, and $P_1, P_2, \ldots, P_r$ are finite sets of rewriting rules over $N \cup T$.

(1) can be rewritten equivalently: $$\Gamma = (N, T, S, G_1, \ldots, G_r), \quad (1')$$
in which $G_i = (N, T, S, P_i)$, for all i, $1 \leq i \leq r$, are Chomsky grammars, called the *components* of Γ. For $X \in \{REG, LIN, CF\}$ we denote by CDGS$_r$X, $r \geq 1$, CD grammar systems with r components, that have regular, linear and context-free components, respectively. The language generated by these systems depends on the way in which the internal rules of each component bring their own contribution on the sentential form. This can be done with respect to several modes of derivation, defined below:

Definition 2 Let $\Gamma = (N, T, S, P_1, \ldots, P_r)$ be a CDGS, $x, y \in (N \cup T)^*$, and $i \in \{1, ..., r\}$. The **terminating derivation** (denoted by $\Rightarrow^t_{P_i}$), the **k-steps derivation** (denoted by $\Rightarrow^{=k}_{P_i}$), **at most k-steps derivation** (denoted by $\Rightarrow^{\leq k}_{P_i}$), **at least k-steps derivation** (denoted by $\Rightarrow^{\geq k}_{P_i}$), and the ***-mode of derivation** (denoted by $\Rightarrow^*_{P_i}$), represent modes of derivations that allow for each component P_i to consecutively activate: as many rules as possible, exactly k rules, at most k rules but at least one rule, at least k rules, and arbitrarily many rules, respectively.

Let $\Gamma = (N, T, S, P_1, \ldots, P_r)$ be a CDGS and $M = \{t, *\} \cup \{\leq k, = k, \geq k | k \geq 1\}$.

Definition 3 The language generated by Γ in *f-mode*, $f \in M$ is defined as:
$$L_f(\Gamma) = \{w \in T^* | S = w_0 \Rightarrow^f_{P_{i_1}} ... \Rightarrow^f_{P_{i_m}} w_q = w, m \geq 1, 1 \leq i_j \leq r, 1 \leq j \leq m\}.$$

For $X \in \{REG, LIN, CF\}$, $f \in M$ we denote by $CD_rX(f)$, $r \geq 1$, the family of languages generated by CDGS with r components, that have only regular, linear, and context-free rules activated in the f-mode of derivation.

Definition 4 Let $\Gamma = (N, T, S, P_1, \ldots, P_r)$, be a CDGS. The **control word** of w, with respect to the system components that have been applied in f-mode, $f \in M$, for a terminal derivation: $S = w_0 \Rightarrow^f_{P_{i_1}} w_1 \Rightarrow^f_{P_{i_2}} w_2 ... \Rightarrow^f_{P_{i_m}} w_q = w$ is defined as $\gamma_w = P_{i_1} P_{i_2} ... P_{i_m}$. The **Szilard language** associated to the derivation in the f-mode, in Γ is: $Sz(\Gamma, f) = \{\gamma_w | w \in L_f(\Gamma), f \in M\}$.

We denote by $SZ(f)$ the *family of Szilard languages* $Sz(\Gamma, f)$ for any grammar system Γ, in the f-mode of derivation. For more properties of these languages the reader is referred to [7].

Definition 5 Let $\Gamma = (N, T, S, P_1, \ldots, P_r)$, be a CDGS. The **communicational control word** of w, that is a control word built with respect to the communicational nonterminals used during a terminal derivation of the form $S = w_0 \Rightarrow^f_{P_{i_1}} w_1 \Rightarrow^f_{P_{i_2}} w_2 ... \Rightarrow^f_{P_{i_m}} w_q = w$, in f-mode, $f \in M$, is defined as $\gamma^c_w = P^{n_1}_{i_1} P^{n_2}_{i_2} ... P^{n_q}_{i_m}$, where n_j, is the *number of communicational nonterminals* rewritten during the application of rules of the component P_{i_j}, $1 \leq j \leq m$. The **communicational Szilard language** associated to a terminal derivation in the f-mode, in Γ is defined as:

$$Szc(\Gamma, f) = \{\gamma^c_w | w \in L_f(\Gamma), f \in M\}.$$

We denote by $SZC(f)$ the *family of communicational Szilard languages* $Szc(\Gamma, f)$ for any grammar system Γ, in f-mode of derivation. Note that in the case of grammar systems with regular and linear components the languages $Sz(\Gamma, f)$ and $Szc(\Gamma, f)$ are equal. They can be different only in the case of grammar systems that contains at least one non-linear rule.

Besides modes of derivation other restrictions that control the generative process are given by fairness conditions. Informally, these conditions require that all components of the system have approximately the same contribution on the common sentential form. They have been introduced in [6], in order to control and to increase the generative capacity of grammar systems. Formally they are defined as follows:

Definition 6 Let $\Gamma = (N, T, S, P_1, \ldots, P_r)$, be a CDGS, and

$$\text{D: } S = w_0 \Rightarrow^{=n_1}_{P_{i_1}} w_1 \Rightarrow^{=n_2}_{P_{i_2}} w_2 ... \Rightarrow^{=n_q}_{P_{i_m}} w_q = w$$

be a derivation in f-mode, where P_{i_j} performs n_j steps, $1 \leq j \leq m$. For any $1 \leq p \leq r$, we set

$$\psi_D(p) = \sum_{i_j = p} 1 \qquad \text{and} \qquad \varphi_D(p) = \sum_{i_j = p} n_j \tag{2}$$

- *the weak maximal difference between the contribution of two components involved in the derivation* D is defined as:

$$dw(D) = max\{|\psi_D(i) - \psi_D(j)| \, |1 \leq i, j \leq r\},$$

- *the strong maximal difference between the contribution of two components* is:

$$ds(D) = max\{|\varphi_D(i) - \varphi_D(j)| \, |1 \leq i, j \leq r\}.$$

Let $u \in \{w, s\}$, $x \in (N \cup T)^*$, $f \in M$ and

$$du(x, f) = min\{du(D)| \text{ where D is a derivation of } x \text{ in } f\text{-mode}\},$$

for a fixed natural number $q \geq 0$,

- the **weakly q-fair** language generated by Γ in the f-mode is defined as:

$$L_f(\Gamma, w - q) = \{x | x \in L_f(\Gamma) \text{ and } dw(x, f) \leq q\}$$

- the **strongly q-fair** language generated by Γ in the f-mode as:

$$L_f(\Gamma, s - q) = \{x | x \in L_f(\Gamma) \text{ and } ds(x, f) \leq q\}.$$

For $X \in \{REG, LIN, CF\}$ and $f \in M$, $M = \{t, *\} \cup \{\leq k, = k, \geq k | k \geq 1\}$ we denote by $CD_rX(f, w - q)$ and $CD_rX(f, s - q)$, $r \geq 1$, the family of weakly and strongly q-fair languages, respectively, generated by CDGS with r components, that have regular, linear, and context-free components in the f-mode of derivation.

3 The Amount of Communication

In this section, two communication structures are proposed in order to investigate the communicational process of distributed generation of languages for the case of CDGS. The first structure is given by the *communication graph* of the CDGS, while the second structure is given by the *protocol of collaboration* between the system components, and it is strictly related to the structure of the communicational Szilard language associated to a language generated by CDGS. We call this structure the *communicational protocol tree* underlied by the grammar system. Another measure deals with the number of communicational steps spent during the computational process. We call it *communication complexity.* In what follows we describe how these complexity measures work together, and how they can be used to characterize the communicational process of CDGS.

3.1 Measures of Communication

Let $\Gamma = (N, T, S, G_1, \ldots, G_r)$ be a CDGS and $M = \{t, *\} \cup \{\leq k, = k, \geq k | k \geq 1\}$.

Definition 7 The **communication graph** of a CDGS crossed in the f-mode of derivation, $f \in M$, is a directed graph in which the vertices are labeled by the CDGS components that communicate with each other. Each directed edge, from a node labeled by G_a to another node labeled by G_b, $a \neq b$, corresponds to a communication step from the component G_a to the component G_b, done during the derivational process, i.e., there exists at least one nonterminal that appears on the right side of a production from G_a and of the left side of a production from G_b, rewritten at least one time during the derivational process in the f-mode.

Definition 8 The **communicational protocol tree** attached to a word $w \in L_f(\Gamma)$, $f \in M$ is the derivation tree attached to the communicational control word of w, i.e. γ_w^c, in the f-mode.

Note that the number of sons of a given node in the communicational protocol tree depends on the type of the rule through which the communication is performed. In the case of regular or linear rules, a grammar G_a communicates with another grammar G_b through only one nonterminal, so that the corresponding protocol tree will be a simple tree (each node has only one son). In the case of non-linear rules the number of sons equals the number of communicational nonterminals from the right side of the rule. Consequently, the shape of the communicational protocol tree depends not only on modes of derivation, but also on the number of communicational nonterminals that exist on the right side of the same production. Therefore, it might exist grammar systems with non-linear rules for which only one nonterminal from the right side of each production is a communicational one. In this case the protocol tree will be a simple tree, too. The communicational protocol tree is not a simple tree, only in the case of CDGS for which there exists at least one non-linear rule that has at least two communicational nonterminals at the right side of it.

The *communication complexity* measure represents the number of communications between different components, during the generative process, by using a minimal number of communicational nonterminals in a specified mode of derivation. This communication measure represents in fact how many times, in a terminal derivation, different components can be applied (this fact can be restricted only by the f-mode of derivation or fairness conditions). Due to the above observation the communication complexity is a function of the length of the generated word depending on $\psi_D(p)$, $1 \leq p \leq r$, introduced in (2). Let $\Gamma = (N, T, S, P_1, \ldots, P_r)$, be a CDGS, and D a derivation in Γ, such that $D : S \Rightarrow^f_{P_{i_1}} w_1 \Rightarrow^f_{P_{i_2}} w_2 ... \Rightarrow^f_{P_{i_m}} w_q = w$.

Definition 9 We denote by $Com(D) = \sum_{p=1}^{r} \psi_D(p)$, where $\psi_D(p) = \sum_{i_j = p} 1$, the number of communication steps used during the derivation D.
The **communication complexity of a word** w, $w \in L_f(\Gamma)$ is defined as:
$$Com(w, \Gamma) = min\{Com(D) | D : S \Rightarrow^* w\}.$$
The **communication complexity of** Γ over all words of length n is:
$$Com_\Gamma(n) = sup\{Com(w, \Gamma)|, w \in L_f(\Gamma), |w| = n\}.$$
The **class of languages that can be generated within communication f** by a CDGS, is defined as: $COM(g) = \bigcup_\Gamma \{L_f(\Gamma) | Com_\Gamma = O(g)\}$.

To be observed that in the case of CDGS with regular or linear components the control language and the communicational control language are equal. Furthermore, due to the fact that the height of the communicational protocol tree of a certain word w is equal with the length of the derivation of the control word associated to w, Definition 9 of the communication complexity can be equivalently redefined as follows.

Definition 10 The number of communication steps, i.e $Com(D)$, used during a derivation D of a particular word w is the height of the communicational protocol tree attached to the communicational control word.

The **communication complexity of a word w**, $w \in L_f(\Gamma)$, i.e $Com(w, \Gamma)$, is the minimum of the heights over all communicational protocol trees attached to each communicational control words of w.

The **communication complexity of** Γ over all words of length n, i.e $Com_\Gamma(n)$, is the supremum of the heights over all minimal communicational protocol trees.

Let Γ be a CDGS, and D be a (minimal) terminal derivation of w, where $w \in L_f(\Gamma)$, $f \in M$, $M = \{t, *\} \cup \{\leq k, = k, \geq k | k \geq 1\}$. We denote by $|\gamma_w(D)|$ and $|\gamma_w^c(D)|$ the length of the derivation of the control word, $|\gamma_w|$, and of the communicational control word, $|\gamma_w^c|$, associated to w, respectively. Then, the next theorem holds.

Theorem 1 For each grammar system Γ, that has only useful components[1] and $w \in L_f(\Gamma)$, we have:
1. $|\gamma_w(D)| = Com(|w|)$.
2. There exist two positive constants a and b, such that $a|\gamma_w^c(D)| \leq |w| \leq b|\gamma_w^c(D)|$.

Proof. The first claim is a direct consequence of Definitions 9 and 10. To prove the second claim we consider firstly the case of regular and linear components. In this

[1]That is each component brings contributions on the sentential form, directly through terminal symbols (in the case of regular or linear rules), or indirectly through non-terminal symbols (in the case of context-free components).

case at each step between two consecutive communications at least one terminal symbol is brought into the sentential form. Therefore, the generated word will contain at least the number of communication steps performed during the generative process. In this case we have $a = 1$, and b equals the maximum number of the terminal symbols brought into the sentential form by each rule of the system.

In the case of grammar systems with (non-linear) context-free components, let us consider γ_w^c of the form $\gamma_w^c = P_{i_1}^{n_1} P_{i_2}^{n_2} ... P_{i_k}^{n_k}$. The leaves in the associated protocol tree correspond to those components that contain rules that have on their right side only terminals. They contribute in the sentential form with substrings of w. The worst case happens when no terminal symbol is brought into the sentential form between any two consecutive communication steps. Therefore, the number of the leaves, i.e. n_k, is less than or equal to the length of w, and the length of w cannot be more then n_k multiplied by a constant c. So that we have

$$n_k \leq |w| \leq cn_k \leq c_{max} |\gamma_w^c(D)|. \tag{3}$$

On the other hand there are situations when between each two consecutive communication steps each component contributes in the sentential form with a constant number of terminals. Each communicational nonterminal might bring its own contribution on the sentential form, too. That is why we also have

$$c_{min} |\gamma_w^c(D)| = c_{min}(n_1 + n_2 + ... + n_k) \leq c_1 n_1 + c_2 n_2 + ... + c_k n_k \leq |w| \tag{4}$$

From (3) and (4) we have $a = c_{min}$ and $b = c_{max}$. □

Due to the above result the communication complexity of CDGS is strictly related not only to modes of derivation or fairness conditions but also on the types of the rules of the system components, and on the number of non-terminals of a non-linear rule that are communicational non-terminals. In the next subsection, we state several results that concern the time, space and the communication complexity of CDGS with regular, linear and context-free components.

3.2 CDGS with Regular and Linear Components

Theorem 2 For each grammar system $CDGS_r X$, with $X \in \{REG, LIN\}$ and $r \geq 2$, there exists a $CDGS$ with only one component that will generate the same language, and vice-versa, independently of modes of derivation.

Corollary 1 $CD_* X = X$, for $X \in \{REG, LIN\}$, independently of modes of derivation.

Corollary 2
$SZC(CDGR_* X, f) = SZ(CDGR_* X, f) = REG$, for $X \in \{REG, LIN\}$, and $f \in \{*, t\} \cup \{\geq k, = k, \leq k | k \geq 1\}$.

Corollary 3 The communication complexity of $CD_* X(f)$, $X \in \{REG, LIN\}$, $f \in \{t, *\} \cup \{\leq k, = k, \geq k | k \geq 1\}$ is 0.

Corollary 4 $LIN \subseteq COM(0)$.

It is well known that the communication complexity divides languages in small complexity classes. The above results show that the communicational process of

CDGS is a lazy one. So that the process of communication in these system is not so powerful as it has been proved to be for the case of PCGS, where several hierarchies of very (small) complexity classes have been found. Due to the fact that fairness conditions increase the generative power of a grammar system, the above theorem and corollaries do not hold for the case of q-fair languages. A CDGS with arbitrary number of components cannot be "compressed" into a single grammar that generates the same q-fair language, by preserving the mode of derivation, too. Even for these types of languages in the case of a constant communication, the class of weakly q-fair languages generated by CDGS with regular or linear components coincide with the languages generated by the same grammar without any weak fairness condition, so that due to Corollary1 we have:

Corollary 5 $CD_{r,c}X(f, w-q) \subseteq COM(0)$ and $CD_{r,c}X(f, s-q) \subseteq COM(c)$ where $X \in \{REG, LIN\}$ and c is the constant number of communicational steps performed during the derivation.

Corollary 6 The communication complexity of weakly/strongly q-fair languages, $L_f(\Gamma, w-q)/L_f(\Gamma, s-q)$, $f \in \{t\} \cup \{\leq k, = k, \geq k | k \geq 1\}$, generated by $CDGS_*X$, $X \in \{REG, LIN\}$, for which the communication graph is a tree or a dag is 0/constant.

Nevertheless the above results do not hold for the case of strongly q-fair languages generated by CDGS with non-constant communication. That is why the communication complexity of q-fair languages deserves to be studied separately. In what follows we focus on the time, space and communication complexity of these very particular class of languages. Below we give one of the most classical example, see [6], related to these types of languages.

Example 1 Let us consider the CDGS

$$\Gamma_1 = (\{S, A, A', B, B'\}, \{a, b\}, S, P_1, P_2, P_3, P_4)$$

with the components:

$$P_1 = \{S \to aA', A \to aA'\}, \quad P_2 = \{A' \to aA\},$$
$$P_3 = \{A \to bB', B \to bB'\}, \quad P_4 = \{B' \to bB, B' \to b\}.$$

The communication graph is displayed in Figure 1.

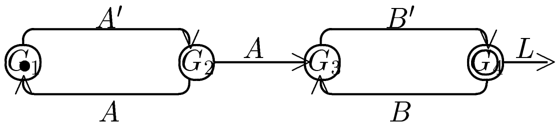

Figure 1: The communication graph associated to the CDGS, Γ_1.

The language generated by Γ_1, in the f mode of derivation, where $f \in \{t, = 1, \geq 1\} \cup \{\leq k | k \geq 1\}$ is: $L_f(\Gamma_1) = \{a^{2n}b^{2m} | n \geq 1, m \geq 1\} \in REG$. The *Szilard* language is $Sz(\Gamma_1) = \{(P_1P_2)^n(P_3P_4)^m | n \geq 1, m \geq 1\}$. If we impose the fairness conditions then we have:
$L_f(\Gamma_1, w-q) = L_f(\Gamma_1, s-q) = \{a^{2n}b^{2m} | n \geq 1, m \geq 1, |n-m| \leq q\} \notin REG$. Not being regular this language cannot be inferred from the communication graph

displayed in Figure 1. That is why we need another kind of machine to check the fairness conditions. Next we show that if the communication is not constant then the communication complexity of q-fair languages cannot be more than linear. Furthermore, fairness conditions can be checked in linear time and space by a multitape Turing machine, and in linear space and quadratic time by a one tape Turing machine. For the definitions of one tape or multitape Turing machine, the reader is referred to [16].

Theorem 3 $CD_rX(f, w-q) \cup CD_rX(f, s-q) \in COM(n)$.

Proof. In the case of grammar systems with regular and linear rules the Szilard and the communicational Szilard languages are equal. With respect to Theorem 1, for any word w that belongs to the weakly or strongly q-fair language, we have: $|\gamma_w(D)| = Com(w, \Gamma) = |\gamma_w^c(D)|$, where D is a minimal derivation of w. Hence, $sup\{|\gamma_w(D)|| \ w \in L_f(\Gamma, u-q), |w| = n\} = sup\{Com(w, \Gamma)| \ w \in L_f(\Gamma, u-q), |w| = n\} = Com_\Gamma(n) = sup\{|\gamma_w^c(D)||w \in L_f(\Gamma, u-q), |w| = n\} = O(n)$, where $u \in \{w, s\}$. □

Theorem 4 Let Γ be a CDGS$_r$X, for $X \in \{REG, LIN\}$. The weakly q-fair language generated by Γ, i.e., $L_f(\Gamma, w-q)$, can be accepted by a nondeterministic Turing machine with $r+1$ tapes in linear time and space. Moreover the next relation hold:

$$Space_T(n) \in O(Com_\Gamma), \quad Time_T \in O(Com_\Gamma).$$

Proof. Let $\Gamma = (N, T, S, P_1, \ldots, P_r)$ be a CDGS$_r$X, for $X \in \{REG, LIN\}$. Let $L_f(\Gamma, w-q)$ be the weakly q-fair language generated by Γ. Next we describe a nondeterministic $(r+1)$-tape Turing machine, with left, right, and stationary movements, that accepts $L_f(\Gamma, w-q)$. For the beginning the machine has on the first tape an input string w of length n, generated by the above CDGS$_r$X, for $X \in \{REG, LIN\}$, in f-mode of derivation, $f \in \{t\} \cup \{\leq k, = k, \geq k | k \geq 1\}$, followed by q symbols \$. Symbols from the right side of a production of the form $A \to xA'y$, or of the form $A \to A'y$, are marked in the input string by an index r, while the others are left unchanged. Each tape of the machine corresponds to one of the system components. At the start of the computation all heads are placed at the beginning of each tape. The machine starts nondeterministically with the component that contains the axiom. Let us assume that this is the component $i-1$. Therefore, the i^{th} head of the machine writes the nonterminal symbol that appears on the right side of the starting production on the i^{th} tape, letting the first head to read the symbol x from the input string, corresponding to a starting rule of the form $S \to xS'$, $S \to xS'x_r$, or to read no symbol in the case of starting rules of the form $S \to S'$, $S \to S'x_r$. From now on the machine simulates on the i^{th} tape the derivation process done by the component $i-1$, in the f-mode, as follows. The first head reads the symbols x (or x_r) from the input tape, with respect to rules of the form $A \to xA'$, $A \to xA'x_r$ (or $A \to A'x_r$), and does not read any symbol with respect to rules of the form $A \to A'$, $A \to A'x_r$ of the component $i-1$. In the meantime, the i^{th} head rewrites on the same cell the nonterminal reached by the rule, i.e. A'. All the other $(r-1)$ heads make no moves and write no symbols. When a communicational nonterminal has been reached, and when the f-mode of derivation of the component $i-1$ ends, the machine nondeterministically jumps to the tape corresponding to the component to which the current component is communicating. Let us consider that this is the component $j-1$. In this moment the machine replaces the communicational nonterminal from the i^{th} tape with the \$ symbol, in order to mark that the i^{th} component has been applied once in the generative process. From now on the machine simulates the work of the component $j-1$ on the j^{th} tape, in the same way as before. The process is repeated until the whole input string is read[2], i.e., the generative process performed by the CDGS$_r$X, for $X \in \{REG, LIN\}$ in the f-mode of derivation has been accomplished. In this moment the first head of the machine will be located on the symbol \$. On each i^{th} tape there will be $\psi_D(i-1)$ symbols \$, that correspond to the number of contributions brought on the sentential form by the component $(i-1)$, $2 \le i \le r+1$. From now on the checking of the weak fairness condition, is performed as follows. All the heads, excepting the first one, simultaneously delete one by one the \$ symbols that have been written on each tape. When one of the heads gets at the beginning of the tape, i.e. the minimum number of the contributions of a component had been deleted, the first head starts to read the q symbols \$ from the first input tape. If these symbols have been read before the deletion of all \$ symbols from the other tapes it means that the weak q-fair restriction is not accomplished, therefore w will not be accepted by the machine. If the deletion of the \$ symbols from all tapes ends before or at the same time with the reading of the q symbols \$ from the first tape, w satisfies the weak q-fair restriction. Therefore the word will be accepted by the machine, as belonging to $L_f(\Gamma, w-q)$. The space needed by this machine to perform the computation cannot be more than the number of the communication steps performed by the grammar system, that is $Space_T(n) \in O(Com_\Gamma(n))$. The time required by the machine to accept and to check the weak q-fair condition is $O(Com_\Gamma(n))$, too. Due to Theorem 3, $Com_\Gamma \in O(n)$, so that $Space_T \in O(n)$ and $Time_T \in O(n)$. □

Theorem 5 Let Γ be a CDGS$_r$X, for $X \in \{REG, LIN\}$. The weakly *q-fair* language generated by Γ, i.e. $L_f(\Gamma, w-q)$, can be accepted by a nondeterministic Turing machine with one tape in linear space and quadratic time, i.e.

$$Space_T \in O(n), \quad Time_T \in O(n^2).$$

The above theorem is a direct consequence of Theorem 4 and of Savitch's theorem.

3.3 CDGS with Context-free Components

Theorem 6 $CD_*CF(f) = CF$, for $f \in \{*, =1, \ge 1\} \cup \{\le k | k \ge 1\}$.

Corollary 7 $SZ(CDGR_*CF, f) = REG$, for $f \in \{*, =1, \ge 1\} \cup \{\le k | k \ge 1\}$.

Corollary 8 $CF \subseteq COM(0)$.

Corollary 9 $CD_{r,c}CF(f, w-q) \in COM(0)$, where $f \in \{*, =1, \ge 1\} \cup \{\le k | k \ge 1\}$, and c is the constant number of communication steps spent during the computation.

[2]Note that when all the communicational nonterminals have been spent, the symbols x_r left on the input tape, will be read by forbidding the other heads, excepting the first one, to move.

Corollary 10 The communication complexity of $L_f(\Gamma, w-q)$, $f \in \{*, =1, \geq 1\} \cup \{\leq k | k \geq 1\}$ generated by $CDGS_*CF$, for which the communication graph is a tree or a dag is 0.

Corollary 11 $CD_*LIN(f) \cup CD_*CF(f_1) \cup CD_{*,c}LIN/CF(f/f_1, w-q) \subset COM(0)$, $CD_{*,c}CF(f_2, w-q) \cup CD_{*,c}X(f, s-q) \subset COM(c)$, $X \in \{REG, LIN, CF\}$, for $f \in \{t, *\} \cup \{\leq k, = k, \geq k | k \geq 1\}$, $f_1 \in \{*, = 1, \geq 1\} \cup \{\leq k | k \geq 1\}$ and $f_2 \in \{t\} \cup \{= k, \geq k | k \geq 1\}$.

In the case of non-constant communication we have.

Theorem 7 For each grammar system Γ with context-free components, and $w \in L_f(\Gamma)$, there exists a bijection $h : N \to N$ such that $|\gamma_w^c(D)| = h(|\gamma_w(D)|)$, where D is the the minimal derivation of w, $\gamma_w(D)$ and $\gamma_w^c(D)$ are the length of the derivation of the control word, γ_w, and of the communicational control word, γ_w^c, associated to w, respectively.

Proof. Let $\gamma_w^c = P_{i_1}^{n_1} P_{i_2}^{n_2} ... P_{i_k}^{n_k}$, be the communicational control word attached to w. In the case of CDGS with context-free components, γ_w^c is developed as a protocol tree attached to w. The number of sons at each level of this tree depends recursively on the number of sons of the previous levels, because communicational rules from different components are applied recursively, by using each time the same type of rules, that increases (linearly or exponentially) the number of communicational nonterminals used during the derivation. Consequently, at the end of the generative process the sum $n_1 + n_2 + ... + n_k$, will be a linear, polynomial or exponential function that depends on the length of the generated string. □

Next, we call the function h the *characterization function of the communicational Szilard language* $Szc(\Gamma, f)$.

Theorem 8 The class of languages generated by CDGS$_r$CF in f-mode of derivation, $f \in \{t\} \cup \{= k, \geq k | k \geq 1\}$, for which the characterization function of the $Szc(\Gamma, f)$ language is linear, polynomial of rank p or exponential with the base p, has the communication complexity in $O(n)$, $O(\sqrt[p]{n})$, or $O(log_p n)$, respectively.

Proof. With respect to Theorem 1, for any word $w \in L_f(\Gamma)$ of length n, we have

$$|\gamma_w(D)| = Com_\Gamma(|w|) = Com_\Gamma(n) \text{ and } |\gamma_w^c(D)| = O(|w|) = O(n).$$

With respect to Theorem 7, there exists a generative bijection $h : N \to N$, such that $|\gamma_w^c(D)| = h(|\gamma_w(D)|)$. Consequently, we have $O(n) = h(Com_\Gamma(n))$, so that $Com_\Gamma(n) = h^{-1}(O(n))$. Therefore, if h is a linear function then $Com_\Gamma(n) \in O(n)$, in the case that h is a polynomial function of rank p, then $Com_\Gamma(n) \in O(\sqrt[p]{n})$, while in the case that h is an exponential function of base p, then $Com_\Gamma(n) \in O(log_p n)$. □

Theorem 9 The class of languages generated by a CDGS$_r$CF in f-mode of derivation, $f \in \{t\} \cup \{= k, \geq k | k \geq 1\}$, are generated by a nondeterministic Turing machine, with $r+1$ tapes, within $Space_T \in O(Com_\Gamma)$ and $Time_T \in \Theta(n)$.

Corollary 12 The class of languages generated by a $CDGS_rCF$ in f-mode of derivation, $f \in \{t\} \cup \{= k, \geq k | k \geq 1\}$ for which the characterization function of the $Szc(\Gamma, f)$ is linear, polynomial of rank p, or exponential with the base k, are recognizable by a nondeterministic Turing machine, with $r + 1$ tapes, within $Space_T$ $O(n)$, $O(\sqrt[p]{n})$, or $O(log_p n)$, respectively, and in $Time_T \in \Theta(n)$.

Furthermore, for the case of q-fair languages, the next theorem holds, [1].

Theorem 10 The class of q-fair languages generated by a $CDGS_rCF$ in f-mode of derivation, $f \in \{t\} \cup \{= k, \geq k | k \geq 1\}$ for which the characterization function of $Szc(\Gamma, f)$ is linear, polynomial of rank p, or exponential with the base p, are recognizable by a $(k + 1)$-tape nondeterministic Turing machine in $Space_T \in O(n)$, $O(\sqrt[p]{n})$, or $O(log_p n)$, respectively, and $Time_T \in \Theta(n)$.

4 Conclusion

The process of communication in the case of CDGS with regular and linear components is a lazy one. Furthermore, the communication can be lost, i.e., it has no efficiency in building communication complexity classes, in the case of CDGS with regular or linear components, without any fairness conditions. In the case of non-constant communication, for the case of weakly and strongly q-fair languages, the communication complexity cannot be more than linear. The communication can be lost too, for the case of languages generated by CDGS with context-free components, in the f mode of derivation, $f \in \{*, = 1, \geq 1\} \cup \{\leq k | k \geq 1\}$, or in the case of weakly q-fair languages generated by CDGS with constant communication. The communication is preserved in the case of CDGS with context-free rules, non-constant communication, weakly and strongly q-fair condition and in $f \in \{t\} \cup \{= k, \geq k | k \geq 1\}$ mode of derivation. Moreover, in this last situation we reached several complexity classes depending on the characterization function of the communicational Szilard language associated to a CDGS.

References

[1] L. Cojocaru. On the Time, Space, and Communication Complexity of q-fair Languages. In *Preproceedings of the 6^{th} Descriptional Complexity of Formal Systems Workshop*, DCFS 2004, pp. 154-163, London, Ontario, Canada, July 26-28, 2004.

[2] E. Csuhaj-Varjú, J. Dassow. On cooperating/distributed grammar systems. *Journal of Information Processing and Cybernetics EIK* 26, 49-63, 1990.

[3] E. Csuhaj-Varjú, J. Kelemen. Cooperating grammar systems: a syntactical framework for the blackboard model of problem solving. In *Proceedings AICSR'89*, I. Plander, Ed. North Holland, Amsterdam, 121-127, 1989.

[4] E. Csuhaj-Varjú, J. Dassow, J. Kelemen, Gh. Păun, *Grammar Systems. A Grammatical Approach to Distribution and Cooperation,* Gordon and Breach, Yverdon, 1994.

[5] J. Dassow, Gh. Păun, G. Rozenberg, Grammar Systems. *Handbook of Formal Languages*, Volume II, Chapter 4, edited by G. Rozenberg and A. Salomaa, Springer, Berlin, 155-213, 1997.

[6] J. Dassow, V. Mitrana. Fairness in Grammar Systems. Acta Cybernetica 12, 331-345, 1996.

[7] J. Dassow, V. Mitrana, Gh. Păun. Szilard Languages Associated to Cooperating Distributed Grammar Systems. *Studii si Cercetari Matematice*, 45, 403-413, 1993.

[8] M. Gheorghe, Gh. Păun. Further remarks on cooperating distributed grammar system. *Bull. Math. Soc. Sci. Math. Roumanie.* 34(82), 232- 245, 1990.

[9] J. Hromkovič, J. Kari, L. Kari, D. Pardubská, Two lower bounds on distributive generation of languages. *Fundamenta Informatica*, 25, 271-284, 1996.

[10] J. Hromkovič, J. Kari, L. Kari, Some hierarchies for the communication complexity measures of cooperating grammar systems. *Theoretical Computer Science*, 127, 123-147, 1994.

[11] D. Pardubská, On the power of communication structure for distributed generation of languages. *In Developments in Language Theory at the Crossroads of Mathematics, Computer Science and Biology*, Editors G. Rozenberg and A. Salomaa, Turku, Finland, 12-15, July, 1993.

[12] D. Pardubská, The Communication complexity hierarchy of parallel parallel communicating systems. IMYC'92, 1992.

[13] M. Lukàč. About two communication structures of PCGS. Master Thesis, Department of Computer Science, Faculty of Mathematics and Physics, Comenius University, Bratislava, 1992.

[14] R. Meersman, G. Rozenberg. Cooperating grammar systems. In: *Proc. MFCS'78, LNCS 64*, Springer Verlag, Berlin, 364-374, 1978.

[15] R. Meersman, G. Rozenberg, D. Vermeir. Cooperating grammar systems. In: *Techn. Report 78-12*, Univ. Antwerp, Dept. Math., 1978.

[16] I. Sudkamp. Languages and Machines. An Introduction to the Theory of Computer Science. Addison-Wesley Ed., 1997.

[17] A. C. Yao. Some complexity questions related to distributed computing. In *Proceedings of the 11^{th} Annual ACM Symposium on Theory of Computing (STOC)*, 209-213, 1979.