On Weak Growing Context-Sensitive Grammars*

Gerhard Buntrock and Gundula Niemann

Institut für Informatik, Universität Würzburg, D–97072 Würzburg, Germany; E-mail: [bunt, niemann] informatik.uni-wuerzburg.de

Abstract. We introduce the weak growing context-sensitive grammars (WGCSG): grammars that are growing with respect to a valuation of the positions inside a rule as well as the symbols. We evaluate a string by summing up every product of a symbol value with its position value. For every position valuation which does not agree with the beginning of any exponential function, the corresponding class of WGCSGs characterizes CSL. Such a valuation function is called unsteady. On the other hand all WGCSG related to steady position valuations are linearly inclusion-ordered and characterize the hierarchy of exponential time bounded languages in CSL. This hierarchy collapses iff
- any class defined by an unsteady position valuation has a normal form of order 2 (e.g. Cremers' or Kuroda's normal form)
- any class defined by a steady position valuation is closed under inverse homomorphisms.

1 Introduction

A beautiful theory has always simple concepts and remarkable results. Here we investigate a widely unknown but nice part of the class of context-sensitive languages (CSL).

Our simple concept consists in a valuation of strings and a concentration on context-sensitive grammars that are growing under such valuation, and the remarkable results are in form of characterizations of known dual bounded complexity classes by this quite different concept. Paul asks in general which kind of speed up theorem does exist for space bounded computations ([Pau78]). Only little is known about such dual bounded classes, see e.g. [Pip79]. Here we characterize the open problem, whether all linear space-bounded automata can be simulated by linear space-bounded automata with an additionally universal exponential time bound, in the concepts of formal language theory. There it is questioned, whether a class of languages has a normal form of bounded order or whether another class is closed under inverse homomorphisms.

Noam Chomsky already has observed in his famous work that CSL is characterized by monotone grammars [Cho59]. We will use this characterization for definition: A context-sensitive grammar (CSG) is a quadruple \((N, T, S, P)\), where

* The results of our investigations first were developped in [Nie94]. Informations where the journal version will be published can be obtained via ftp/WWW service from haegar.informatik.uni-wuerzburg.de/pub/TRs/bunt-niem95.dvi.gz.
N and T are finite disjoint alphabets of nonterminal and terminal symbols, respectively, S ∈ N is a start symbol and P is a finite set of rules (productions) of the form \( \alpha \to \beta \) with \( \alpha, \beta \in (N \cup T)^* \), where \( \alpha \) contains at least one nonterminal symbol, and \( |\alpha| \leq |\beta| \).

Dahlhaus and Warmuth investigated the complexity of growing context-sensitive languages, an important subclass of CSL, since it is contained in \( \text{LOGCFL} \), which is a subclass of \( P \) ([DW86], for \( \text{LOGCFL} \) see [Sud78],[RuzS0]). GCSL is defined with growing context-sensitive grammars (GCSG), that is done by replacing the inequality in the definition of CSG by a sharp one\(^2\). If we valuate the symbols of a CSG with a homomorphic mapping into the natural numbers with addition and demand that in every rule the sum of the values of the symbols increases, this type of grammars (quasi GCSG, QGCSG for short) characterizes GCSL ([BL92]).

Unfortunately, the class GCSL is very weak. Not even the language \( \text{COPY} := \{ww \mid w \in \{0,1\}^*\} \) is contained in GCSL [Bun93]. Weak growing context-sensitive grammars (WGCSG) are a generalization of QGCSG and much more powerful. In a WGCSG we valuate the symbols as well as the positions inside a rule. In the following we will refer to \( \max\{\max(|\alpha|,|\beta|) \mid (\alpha \to \beta) \in P\} \) as the order of \( G \).

**Definition 1.1**: A position valuation is a function \( s: \mathbb{N}^+ \to \mathbb{N}^+ \) with:

- If \( s \) is defined for a \( j \in \mathbb{N}^+ \), it is also defined for every \( i \in \mathbb{N}^+ \) with \( i < j \).
- If \( s \) is defined for all \( i \in \mathbb{N}^+ \), we say \( s \) is an infinite position valuation — in the other case, \( s \) is a finite position valuation. We call each \( i \in \mathbb{N}^+ \) where \( s(i) \) is defined a valuated position of \( s \).

We restrict ourselves to the most interesting position valuations which have at least two valuated positions.

**Definition 1.2**: Let \( s \) be a position valuation. A grammar \( G = (N, T, S, P) \) is weak growing context-sensitive related to the position valuation \( s \) or also \( s \)-weak growing context-sensitive (an \( s \)-WGCSG for short), if the following holds:

(i) \( G \) is context-sensitive

(ii) Let \( l \) be the order of \( G \). Then \( s \) has at least \( l \) valuated positions.

(iii) There exists a function \( f: N \cup T \to \mathbb{N} \) such that for every rule \( (\alpha_1 \ldots \alpha_n \to \beta_1 \ldots \beta_m) \in P \) holds:

\[
(*) \quad \sum_{i=1}^{n} s(i) \cdot f(\alpha_i) < \sum_{i=1}^{m} s(i) \cdot f(\beta_i).
\]

Such a function \( f \) is called symbol valuation for \( G \). If the inequality applies for a rule \( (\alpha \to \beta) \in P \), we say it is \( s \)-weak growing with \( f \).

The sets of corresponding grammars and languages are denoted by \( \text{WGCSG}_s \) and \( \text{WGCSL}_s \), respectively.

The valuation of positions gives the possibility to interchange two symbols by a rule. This can be used to prove \( \text{COPY} \in \text{WGCSL}_s \) for any nonconstant

---

\(^2\) Note, we differ between the name of the class CSG and a member CSG.