Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

Weighted Contextual Grammars

Master Thesis

2021
Dávid Barbora, Bc

Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

Weighted Contextual Grammars

Master Thesis

Study Programme:	Computer Science
Field of Study:	Computer Science
Department:	Department of Computer Science
Supervisor:	RNDr. Peter Kostolányi, PhD.

Bratislava, 2021
Dávid Barbora, Bc

59383663

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta:	Bc. Dávid Barbora
Študijný program:	informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor:	informatika
Typ záverečnej práce:	diplomová
Jazyk záverečnej práce:	anglický
Sekundárny jazyk:	slovenský

Názov: Weighted Contextual Grammars
Kontextuálne gramatiky s váhami

Anotácia: V práci sa pre niekoľko vybraných variantov Marcusových kontextuálnych gramatík zadefinujú ich obdoby rozšírené o váhy z polokruhu. Preskúmajú sa základné vlastnosti kontextuálnych gramatík s váhami a nimi realizovaných formálnych mocninových radov.

Vedúci:	RNDr. Peter Kostolányi, PhD.
Katedra:	FMFI.KI - Katedra informatiky
Vedúci katedry:	prof. RNDr. Martin Škoviera, PhD.
Dátum zadania:	11.12.2018

Dátum schválenia: 15.12.2018 prof. RNDr. Rastislav Kráľovič, PhD.
garant študijného programu

....

študent vedúci práce

59383663

Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Bc. Dávid Barbora
Study programme:	Computer Science (Single degree study, master II. deg., full time form)
Field of Study:	Computer Science
Type of Thesis:	Diploma Thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Weighted Contextual Grammars

Annotation: Counterparts incorporating weights from a semiring will be defined in the thesis for several selected variants of Marcus contextual grammars. Basic properties of weighted contextual grammars and of formal power series they realise will be explored.

Supervisor:	RNDr. Peter Kostolányi, PhD.
Department:	FMFI.KI - Department of Computer Science
Head of department:	prof. RNDr. Martin Škoviera, PhD.
Assigned:	11.12.2018
Approved:	15.12.2018

prof. RNDr. Rastislav Kráľovič, PhD.
Guarantor of Study Programme

Student

Supervisor

Acknowledgments: I would like to thank my supervisor RNDr. Peter Kostolányi, PhD. for valuable advice, literature and motivation. I would like to thank my girlfriend Emma and to my family for all the support, mental and financial during the last year. I would like to thank all my friends from the non-profit organization Trojsten without which I would not get to this final step of study at this school and maybe not even to the first. In the end I want to express my gratitude to all the medical workers in the fight with pandemic for the possibility to defend this thesis in person.

Abstrakt

Pre štyri základné varianty kontextuálnych gramatík sme zadefinovali ich obdoby rozšírené o váhy z polokruhu. Zadefinovali sme formálne mocninové rady realizované týmito gramatikami a preskúmali ich základné vlastnosti. Ukázali sme, že formálne mocninové rady realizované externými kontextuálnymi gramatikami s voľbou sú uzavreté na súčet. Venovali sme sa tiež jednoznačnosti týchto gramatík a dokázali sme, že táto vlastnosť závisí od vlastností zvoleného polokruhu. Dokázali sme, že postačujúca algebraická štruktúra ktorá rozširuje polokruh je pre túto vlastnosť polopole a všetky formálne mocninové rady realizované externými kontextuálnymi gramatikami s volbou s váhami nad polopolom sú jednoznačné. V závere poskytujeme krátky prehľad otvorených otázok o kontextuálnych gramatikách s váhami hodných preskúmania.

Kľúčové slová: kontextuálne gramatiky s váhami, formálny mocninový rad, polokruh

Abstract

We defined weighted contextual grammars over semiring in four basic variants as extension of contextual grammars. Behavior of these grammars was defined as formal power series and basic properties were studied. We show that formal power series realized by weighted contextual grammars with choice in external mode are closed under sum. We also study ambiguity of such grammars and prove that it depends on properties of semiring. We proven that sufficient algebraic structure that extends semiring is semifield and all weighted contextual grammars over semifield with choice are unambiguous. In the end of this thesis we provide a short overview of open questions about weighted contextual grammars that are worth of study.

Keywords: weighted contextual grammars, formal power series, semiring

Contents

Introduction

This thesis combine two big fields in formal language theory - Contextual grammars and theory of weighted automata. Many generative and computational models have been studied in combination with weights, however Contextual grammars lack this extension. In this thesis we decided to define weighted variant of selected Contextual grammars in the general form with weights from semiring as the appropriate algebraic structure. Our goal is open field of weighted contextual grammars with the definitions and generalization of known results to weighed variant.

Contextual grammars are particularly interesting because they are very different from standard models from Chomsky hierarchy. There are multiple variants of contextual grammars that are known to model properties of natural languages that are in sense of Chomsky hierarchy non-context free. Other derived models are studied as a good model for biological operations over DNA. All the studys of formal languages have origins in understanding of natural languages or artificial languages. With the same goal was firstly introduced weighted variant of finite state automaton. We believe that introduction of weighted Contextual grammars will be a great contribution to both contextual grammars theory and theory of weighted automata.

In the first chapter we introduce concept of context in sense of contextual grammars and define four basic classes of languages generated by the internal and external contextual grammars both with and without choice. These variants are presented with simple examples to provide intuition behind context adjoining. Generative capacity is presented for the reference without proofs, as they are either trivial, complicated or not so important for this thesis. Results about closure properties are also presented without proofs as most of proofs are just counterexamples. Two more modes of derivation in contextual grammars are presented next; being maximal global and maximal local use of selectors. These modes, however more complicated, model properties of natural languages even better. The second part of this chapter is dedicated to standard formalism in study of weighted automata. First we define formal power series with multiple non-commuting variables and coefficients from semiring. This is the best formalism to describe behavior of some weighted model, or in other words to describe set of words with weights from semiring. One of the advantages of this formalism is that we can simply extend union of languages to sum of formal power series.

Weighed contextual grammars are defined in the second chapter in the four variants as well. Weighted contextual grammars with choice are defined as extension of contextual grammars with choice and weighted contextual grammars without choice are defined as extension of non-weighted variant without choice, however we also present it as special case of weighed contextual grammars with choice. Behavior of these grammars is defined in both external and internal mode. In both modes we defined formal power series realized by weighted contextual grammar with choice as sum of all derivations in the grammar in corresponding mode. We also show that weight of word in such grammar can be calculated from weights of words it can be derived from in one step. Behavior of weighted contextual grammars without choice can be viewed as special case of variant with choice, so we do not give much attention to it.

Basic properties of formal power series realized by grammars as we defined them are studied in the third chapter. In particular we provide a proof of closure of formal power series realized by weighted contextual grammars with choice in external mode over sum. This result can be also interpreted as closure of non-weighted variant over union which is one of a few positive closure results in the field of contextual grammars. The proof we present is also different from the standard proof in non-weighted variant. Another property we decided to study is ambiguity. It is particularly interesting in the linguistic view of language and also results into some interesting outcomes. Languages generated by external contextual grammars with choice are in general unambiguous, which mean that there is grammar that can generate each word in only one way. We show that in weighted variant it depends on semiring. We present example and proof of inherently ambiguous formal power series realized by weighted contextual grammar with choice in external mode over standard semiring of natural numbers as well as sufficient condition on semiring for all power series to be unambiguous with respect to weighted contextual grammars over such semiring in external mode.

Chapter 1

Preliminaries

In this thesis, we define and study weighted variants of selected contextual grammars. The first part of this chapter presents the basic variants - contextual grammars with and without choice in both internal and external mode of derivation, as well as newer variant being contextual grammars with maximal use of selectors. Four basic variants are a good starting point as they are well studied and maximal use of selectors is a good extension as they suit natural language even better. In the second part of this chapter we present formalism of formal power series over semiring with non-commuting variables that is widely used in study of weighted variants of formal languages. The reader is not expected to have any prior knowledge of contextual grammars or formal power series, but formal languages and Chomsky hierarchy should not be a new topic.

1.1 Contextual Grammars

In this section, we will introduce Contextual Grammars [4], introduced by Solomon Marcus in 1969. The contextual grammars are sometimes referred as Marcus Contextual grammars, to avoid ambiguity between numerous generative models dealing with context also called grammars like all variations of Context-sensitive grammars [1]. In this section, most of names and definitions are based on monograph Marcus Contextual Grammars [7] by Gheorghe Păun written in 1997 which contains all important definitions and results about Contextual grammars known at that time.

The motivation behind all contextual grammars comes from linguistics, and properties from natural languages that make them non Context free in sense of Chomsky hierarchy. On the other side, context free languages have some additional properties that are not required by natural languages. As a result of these observations, contextual grammars are incomparable with Chomsky hierarchy, in sense of inclusion. We will present examples to give better understanding where this incomparability comes from.

1.1.1 Definitions and variants

All kinds of contextual grammars deal with context. Basic idea of context is, that substring w of string s is in context of words u and v if $s = xuwvy$ for some words x and y. Contextual grammars always start with a finite set of words, called axioms, and generate new words by iteratively adding context from a finite set of contexts.

The most of variability comes in context adjoining. The basic idea is to add context only externally without any restrictions, so that $x \Longrightarrow y$ if $y = uxv$ for some context (u, v). To be consistent with further definitions, we will define contextual grammar more generally, with a selection mapping φ that restricts usage of contexts only to specified words.

Definition 1 (Contextual Grammar with choice) *We say $G = (V, A, C, \varphi)$ is a Contextual Grammar with choice, where V is an alphabet, A is a finite language over V called axioms, C is a finite subset of $V^* \times V^* - \{(\varepsilon, \varepsilon)\}$ called contexts, and $\varphi : V^* \to 2^C$ is the selection/choice mapping.*

The absence of empty context $(\varepsilon, \varepsilon)$ is not contained in the original definition for simplicity, however it will simplify further definitions in this case. The contextual grammar without empty context can be also considered as normal form, but it is such a small difference that we will use it as part of definition. Next we will define two variants of derivation step. The external and the internal mode of derivation.

Definition 2 (Internal and external derivation step) *Let $\Longrightarrow_{ex}$ and $\Longrightarrow_{in}$ be relations on V^* such that:*

$$\begin{aligned} x \Longrightarrow_{ex} y \quad iff \quad & y = uxv, \;\; for\ a\ context\ (u, v) \in \varphi(x) \\ x \Longrightarrow_{in} y \quad iff \quad & x = x_1x_2x_3, \;\; y = x_1ux_2vx_3, \;\; for\ any \\ & x_1, x_2, x_3 \in V^*, (u, v)\ in\ \varphi(x_2) \end{aligned}$$

Now we can associate two languages with contextual grammar with choice G, based on the derivation step used.

Definition 3 (Generated language) *For $\alpha \in \{in, ex\}$ we say:*

$$L_\alpha(G) = \{x \in V^* | w \Longrightarrow_\alpha^* x, \;\; for\ any\ w \in A\}$$

is internal/external language generated by G.

Now we shall return to the simplest form of contextual grammars, which is called Contextual grammars without choice, where we ignore the selection function φ.

Definition 4 (Contextual grammar without choice) *Let $G = (V, A, C, \varphi)$ be contextual grammar with choice as defined earlier. If $\varphi(x) = C$ for all $x \in V^*$ we say that grammar G is without a choice.*

As we can see in this definition is φ completely ignored, so we can write contextual grammar without choice simply as $G = (V, A, C)$ without rewriting all other definitions.

Now we have sufficient definitions for four basic variants of contextual grammars:

ECC - the family of languages externally generated by contextual grammars with choice

ICC - the family of languages internally generated by contextual grammars with choice

EC - the family of languages externally generated by contextual grammars without choice

IC - the family of languages internally generated by contextual grammars without choice

In case of grammars with choice, another natural restriction is on selectors (language of words x on which $\varphi(x)$ is defined the same). For F being a family of languages, we shall write ECC(F) (ICC(F) respectively). For corresponding class with restricted language of selectors from family F.

We will not give much attention to this restricted variants, however the notation might be sometimes used. The family F is usually considered being finite (FIN), regular (REG), context-free (CF), context-sensitive (CS) or recursively enumerable (RE) with emphasis on FIN and REG.

1.1.2 Examples of contextual grammars

In this section we present four simple examples of contextual grammars to present some basic or interesting properties.

Example 1 *Every finite language L over alphabet V, belongs to every family of contextual languages. We just consider this language to be language of axioms of grammar, and contexts to be empty set. We can see that $L_{ex}(G) = L_{in}(G) = L$.*

$$G = (V, L, \emptyset, \emptyset)$$

Example 2 *For each alphabet V, the language V^* belongs to every family of contextual languages.*

$$G = (V, \{\varepsilon\}, \{\varepsilon\} \times V, \varphi)$$

Where G is without choice.

The next example will show the difference between internal and external mode of derivation.

Example 3 *Consider the grammar without choice* $G = (\{a, b\}, \{\varepsilon\}, \{(a, b)\}, \varphi)$*.*

$$L_{ex}(G) = \{a^n b^n | n \geq 0\}$$

$$L_{in}(G) = D_{\{a,b\}}$$

where $D_{\{a,b\}}$ *is the Dyck language over alphabet* $\{a, b\}$*.*

The second equality can by proven by induction on the length of the string.

Note that $L_{ex}(G)$ is not regular and that $L_{in}(G)$ is not linear.

Example 4 *For the grammar*

$$G = (\{a, b, c, d\}, \{abcd\}, \{(a, c), (b, d)\}, \varphi)$$

$$\varphi(ab^+c) = \{(a, c)\}$$

$$\varphi(bc^+d) = \{(b, d)\}$$

we have

$$L_{in}(G) = \{a^n b^m c^n d^m | n, m \geq 1\}$$

The concept described by this language is in linguistic often and it is called crossed dependencies. Note that the selectors of G are regular languages.

1.1.3 Generative capacity

In this section we will present some basic results about comparability of internal and external contextual grammars as well as position in the Chomsky Hierarchy. Results are presented without proof, for reference.

Lemma 1 *Each family IC, ICC is incomparable with each family EC, ECC.*

Lemma 2 $FIN \subset IC \cap EC$

Let LIN_1 be the family of languages generated by linear grammars with only one nonterminal.

Lemma 3 $EC = LIN_1$

Lemma 4 $REG \subset ICC$

Lemma 5 $REG \subseteq ICC(FIN)$

Lemma 6 $IC \subset CS$

Lemma 7 *REG is incomparable with each of the families IC, EC, ECC*
LIN and CF are incomparable with each of the families IC, ICC, ECC
CS and RE are incomparable with each of the families ICC, ECC

Closure properties are presented in the following table[7]:

	IC	*EC*	*ICC*	*ECC*
Union	No	No	No	Yes
Intersection	No	No	No	No
Complement	No	No	No	No
Concatenation	No	No	No	No
Kleene +	No	No	No	No
Morphisms	No	Yes	No	Yes
Finite substitution	No	Yes	No	Yes
Substitution	No	No	No	No
Intersection with regular languages	No	No	No	No
Inverse morphisms	No	No	No	No
Shuffle	No	No	No	No
Mirror image	Yes	Yes	Yes	Yes

Figure 1.1: Closure properties of contextual grammars.

1.1.4 Maximal use of selectors

Another version of derivation step in contextual grammars was supposed by Martín-Vide in 1995 [6]. This mode was considered to suit natural languages very well as it models many properties of natural and artificial languages (programming languages) that lead to non-context-freeness [5]. As an example of these properties author states reduplication, crossed dependencies, and multiple agreements.

Another significant step towards description of natural languages was made by Karin Harbusch in 2004 when polynomial parser was introduced for a contextual grammars with $CL_\alpha(F)(F \in \{FIN, REG, CF\}, \alpha \in \{in, Ml, Mg\})$.

Now we will define derivation step with maximal local (Ml) and maximal global (Mg) use of selectors. The key difference is, that when using selector on some word, there can not be another longer selector that can be used. In local mode we chose maximal selector among single production and in global mode among all selectors from all productions.

In this definition we will use modular form of contextual grammar with choice which is equvalent to the one presented earlier.

Definition 5 (Contextual Grammar with choice in modular form) *We say $G = (V, A, P)$ is a Contextual Grammar with choice in modular form, where V is an alphabet, A is a finite language over V called axioms, P is a finite set of pairs (S_i, C_i) where $S_i \subseteq V^*$ and $C_i \subseteq V^* \times V^* - \{(\varepsilon, \varepsilon)\}$ called productions, where S_i is selector and C_i is set of allowed contexts.*

Definition 6 (Maximal local use of selectors) *$x \Longrightarrow_{Ml} y$ iff $x = x_1x_2x_3$, $y = x_1ux_2vx_3$, for $x_2 \in S_i$, $(u,v) \in C_i$, for some $1 \leq i \leq n$, and there are no $x_1', x_2', x_3' \in V^*$ such that $x = x_1'x_2'x_3'$, $x_2' \in S_i$ and $|x_1'| \leq |x_1|$, $|x_3'| \leq |x_3|$, $|x_2'| > |x_2|$.*

Definition 7 (Maximal global use of selectors) *$x \Longrightarrow_{Mg} y$ iff $x = x_1x_2x_3$, $y = x_1ux_2vx_3$, for $x_2 \in S_i$, $(u, v) \in C_i$, for some $1 \leq i \leq n$, and there are no $x'_1, x'_2, x'_3 \in V^*$ such that $x = x'_1x'_2x'_3$, $x'_2 \in S_j$, for some $1 \leq j \leq n$, and $|x'_1| \leq |x_1|$, $|x'_3| \leq |x_3|$, $|x'_2| > |x_2|$.*

1.2 Formal power series

Power series as a formalism state a great role in all forms of weighted computational models. In this section, we provide definition of semiring and formal power series with multiple non-commuting variables and coefficients from semiring. Semiring is appropriate algebraic structure as it lets us use summation and multiplication. Typical definition of weight of word consists of summation over weights of all derivations of such word and weight of derivation consist of multiplication of weights of individual derivation steps. The main goal of extending model with weights is to get weights of words. To describe this property of weighted language, we need more than just set of words. Formal power series provide us with good formalism as on basic level it is just mapping of words and coefficients, same as some weighted language.

Definition 8 (Semiring) *Tuple $(S, +, \cdot, 0, 1)$ is a semiring where:*

- *S is set*
- *$(S, +)$ is a commutative monoid with identity element 0:*
  - $(a + b) + c = a + (b + c)$
  - $0 + a = a + 0 = a$
  - $a + b = b + a$
- *$(S, \cdot)$ is a monoid with identity element 1:*
  - $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
  - $1 \cdot a = a \cdot 1 = a$
- *Multiplication left and right distributes over addition:*
  - $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$
  - $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$
- *Multiplication by 0 annihilates:*
  - $0 \cdot a = a \cdot 0 = 0$

Operation $+$ is called addition and $\cdot$ is called multiplication. The symbol $\cdot$ is usually omitted from the notation; that is, $a \cdot b$ is just written ab. Similarly, an order of operations is accepted, according to which $\cdot$ is applied before $+$; that is, $a + bc$ is $a + (bc)$.

In the next part, we will provide a definition and basic properties of formal power series. The basic definition is usually given as series of coefficients a_n of x^n in $\sum a_n x^n$. In this case, we have formal power series of multiple non-commuting variables, so number series are no longer sufficient. To describe weighted language we need universal mapping of coefficients to words of variables.

Definition 9 (Formal power series) *A formal power series r over a semiring S and over an alphabet Σ is mapping from Σ^* to S. We denote as $r = \sum_{w \in \Sigma^*}(r, w)w$ where (r, w) is used instead of $r(w)$ for coefficient of word $w \in \Sigma^*$ in formal power series r.*

The set of all formal power series over a semiring S and an alphabet Σ is denoted $S\langle\langle\Sigma^*\rangle\rangle$.

When formal power series is used as weighted language, we shall use term monomial as weighted word. Monomial is a formal power series such, that there is only one $w \in \Sigma^*$ such that $(r, w) \neq 0$. In case of monomials we will omit the sum in notation and write just cw wherem w is a word and $c \in S$.

The equivalent of union of languages in weighted variant is sum of formal power series defined as follows.

Definition 10 (sum of formal power series) *Sum of formal power series is binary operation $+$ over set $S\langle\langle\Sigma^*\rangle\rangle$ where $(r_1 +_1 r_2, w) = (r_1, w) +_2 (r_2, w)$ for every w in Σ^*, where $+_1$ is sum of series and $+_2$ is sum in semiring S.*

Chapter 2

Weighted Contextual Grammars

In this chapter we start our study by defining weighted variant of internal and external contextual grammars with and without choice presented in previous chapter. As inseparable part of weighted grammars we also define behavior as a formal power series.

2.1 Definitions

The first question to be answered is, what should be weighted in contextual grammars. In finite state automata or Chomsky grammars, there is a set of possible transitions or rewriting rules and the weight is assigned to each transition from this set. Contextual grammars are originally defined in two equivalent forms. One consider choice as mapping from words to sets of allowed contexts (as we defined it) and the other one defines choice as set of pairs of language and the set of allowed contexts. The definitions are equivalent however they show different view of choice.

In this thesis, two following variants will be considered . Either we can give weight to context as a pair of words, or we can consider context as interconnected with the word it is applied to. The second one is obviously more general and appears to be more suitable for contextual grammars with choice, as one context can be used on different words with different selectors. It seems natural to assign different weight to different usages of context.

On the other hand, in case of contextual grammars without choice, a context can be used on any word without any restriction so giving power of choice to weight seems like artificial extension. We will define weighted contextual grammars without choice specially with weighted contexts instead of restricted weighted transitions as it is cleaner.

In sense of weights the advantage of contextual grammars is, that there is no use of "empty transitions" as ε-transitions in automata or nonterminal to nonterminal transitions in chomsky grammars. It is obvious, that in non-weighted version of contextual grammars it makes no sense to use empty context. It adds nothing and result into absolutely same state. In weighted form, however, we explicitly mention that empty context is forbidden, as it could result into infinite derivation and infinite product or sum of weights which is not defined on semiring and would make our definitions invalid.

Definition 11 (Weighted Contextual Grammar with choice) *Let S be a semiring. A Weighted Contextual grammar over S with choice is a tuple $G = (\Sigma, A, C, \varphi, \nu, \iota)$, where (Σ, A, C, φ) is a Contextual Grammar with choice (without weights), ι is initial weight mapping of axioms $\iota : \Sigma^* \to S$ where $\iota(w) = 0$ if $w \notin A$ and ν is a weight mapping of transitions $\nu : \Sigma^* \times (\Sigma^* \times \Sigma^*) \to S$ where $\nu(w, (u, v)) = 0$ if $(u, v) \notin \varphi(w)$.*

The last condition is not necessary in the definition, however it simplifies notation in some later statements. Similarly ι could be defined as $A \to S$ but it would complicate later notation.

Weighted contextual grammar without choice will be first defined as extension of non-weighted variant and then we will show that it is special case of weighted grammars with choice which will allow us to adapt the definition of behavior.

Definition 12 (Weighted Contextual Grammar without choice) *Let S be a semiring. A Weighted Contextual grammar over S without choice is a tuple $G = (\Sigma, A, C, \nu, \iota)$, where (Σ, A, C) is a Contextual Grammar without choice (without weights), ι is initial weight mapping of axioms $\iota : \Sigma^* \to S$ where $\iota(w) = 0$ if $w \notin A$ and ν is a weight mapping of contexts $\nu : C \to S$.*

Let $G = (\Sigma, A, C, \nu, \iota)$ be a Weighted Contextual Grammar without choice over semiring S. Let $\varphi(w) = C$ for all $w \in V^*$ and $\nu'(w, c) = \nu(c)$ for all $(w, c) \in V^* \times C$ and zero otherwise. Grammar $G' = (\Sigma, A, C, \varphi, \nu', \iota)$ is a Weighted contextual grammar with choice that is equivalent to G.

2.2 Behavior

To define behavior of weighted contextual grammar, we have to consider the mode of derivation. We will define four families of weighted languages (formal power series).

WECC - the family of power series realized by weighted contextual grammars with choice in external mode

WICC - the family of power series realized by weighted contextual grammars with choice in internal mode

WEC - the family of power series realized by weighted contextual grammars without choice in external mode

WIC - the family of power series realized by weighted contextual grammars without choice in internal mode

2.2.1 External mode

First we will define behavior of grammars in external mode. We defined grammars without choice as special case of grammars with choice, so the defined behavior stands for both.

For this definition, lets consider γ to be series of contexts in order of application on some axiom w in external mode.

$$\gamma = ((u_1, v_1), (u_2, v_2), \dots, (u_n, v_n))$$

The word derived by this series of contexts will be $\lambda(\gamma) = w_n$ where $w_0 = w$ and $w_k = u_k w_{k-1} v_k$.

γ is the derivation in grammar G if each context satisfies condition to be applied in this grammar $(u_k, v_k) \in \varphi(w_{k-1})$. The set of all derivations in grammar G in external mode will be called $\mathcal{D}_{ex}(G)$. Now we can define weight of the derivation γ as $\sigma(\gamma) = \iota(w_0)\nu(w_0, (u_1, v_1))\nu(w_1, (u_2, v_2)) \cdots \nu(w_{n-1}, (u_n, v_n))$.

The monomial realized by derivation γ can be now defined as

$$||\gamma|| := \sigma(\gamma)\lambda(\gamma)$$

Definition 13 (Behavior of weighted contextual grammar in external mode) *Let S be a semiring and G be a Weighted Contextual Grammar over S with choice. A formal power series $||G||_{ex}$ realized by grammar G in external mode is a series in $S\langle\langle\Sigma^*\rangle\rangle$ given as*

$$||G||_{ex} = \sum_{\gamma \in \mathcal{D}_{ex}(G)} ||\gamma||$$

In this definition, we have used the advantage of formal power series to avoid explicit definition of weight of word. Sometimes, however we will need to consider calculation of weight of some word in different way.

Lemma 8 (Weight of w in $||G||_{ex}$) *Let $Par(w)$ be set of all partitions of word w to three words $(w', (u, v))$ where $w = uw'v$ and $|uv| > 0$. Weight of word w in formal power series realized by grammar G is sum*

$$(w, ||G||_{ex}) = \iota(w) + \sum_{(w',c) \in Par(w)} (w', ||G||_{ex})\nu(w', c)$$

We can define the sum in this general way over all partitions, as we said in the definition of weighted grammar, that $\nu(w, c)$ is zero if context can not be adjoined to w.

Proof 1 *Let D_w be a set of derivations $\delta \in \mathcal{D}_{ex}(G)$ such that $\lambda(\delta) = w$. We can say that weight of word w is sum of weights of derivations from D_w. We will now show that for calculation of weight, this set of derivations can be simplified to axiom and pairs* *consisting of word and last context adjoined to this word. Each derivation* $\delta \in D_w$ *is either axiom or can be divided to derivation of previous word* w' *and last adjoined context* (u, v)*. In* D_w *there can be many derivations of word* w *that end with the same context. Lets say, that derivations* $\delta_1 \dots \delta_n$ *for some* n *end with the same context* (u, v)*. Each of these derivations will be divided to the same pair of word* w' *and context* (u, v)*. We know that each of these derivations of* w' *is different, and we also know that there has to be every possible derivation of* w'*. If there was some derivation* δ' *of* w' *that is not contained in* $\delta_1 \dots \delta_n$ *we could adjoin* (u, v) *and we have new derivation of* w *which is contradiction.*

Now, as we know that there are all derivations of w'*, we also know that sum of the weights of these derivations is weight of word* w'*. In semiring multiplication distributes over summation so we can write* $(w', ||G||_{ex})\nu(w', (u, v))$ *instead of* $\sum_{\delta \in D_{w'}} \sigma(\delta)\nu(w', (u, v))$

$$
\begin{aligned}
(w, ||G||_{ex}) &= \sum_{\delta \in D_w} \sigma(\delta) \\
&= \iota(w) + \sum_{(w', c) \in Par(w)} \sum_{\delta \in D_{w'}} \sigma(\delta)\nu(w', c) \\
&= \iota(w) + \sum_{(w', c) \in Par(w)} (w', ||G||_{ex})\nu(w', c)
\end{aligned}
$$

So the weight of word w *in* $||G||_{ex}$ *is equal to the sum as we stated in 8.*

2.2.2 Internal mode

Behavior of Weighted Contextual Grammars in internal mode is more tricky, because derivation can not be simply written as series of contexts, as it would be ambiguous where was the context applied and our definition of weight function depends on selection. We have to include in definition the division of word to left, middle and right part, when context is applied. The initial definition of derivation is inspired by study on ambiguity of contextual grammars as it is closely related. The ambiguity is defined in five different levels of detail on the derivation of word. We will consider the last of them, originally called the description of derivation which has enough information for our weight function. The ambiguity considers three factors being order of application of contexts, selected word for application and the exact position of application. Study shows, that each of them matters and each level of ambiguity defines different set of languages[7].

Definition 14 (derivation in internal mode) *Let* $G = (\Sigma, A, C, \varphi, \nu, \iota)$ *be a weighted contextual grammar over semiring* S*. The sequence*

$$\delta = w_0, x_{1,1}(u_1)x_{2,1}(v_1)x_{3,1}, x_{1,2}(u_2)x_{2,2}(v_2)x_{3,2}, \dots, x_{1,n}(u_n)x_{2,n}(v_n)x_{3,n},$$

is derivation in internal mode if
$w_0 \in A$
$w_i = x_{1,i}x_{2,i}x_{3,i}$ *for* $x_{1,i}, x_{2,i}, x_{3,i} \in V^*$
$w_{i+1} = x_{1,i}u_i x_{2,i}v_i x_{3,i}$, *where*
$(u_i, v_i) \in \varphi(x_{2,i}), 1 \leq i \leq n-1$

For the compatibility with previous definition, the word derived by δ will be $\lambda(\delta) = w_n$ The set of all internal derivations in grammar G will be called $\mathcal{D}_{in}(G)$. Now we can define weight of the derivation δ as

$$\sigma(\delta) = \iota(w_0)\nu(x_{2,1}, (u_1, v_1))\nu(x_{2,2}, (u_2, v_2)) \cdots \nu(x_{2,n-1}, (u_{n-1}, v_{n-1}))$$

.

The monomial realized by derivation δ can be now defined as

$$||\delta|| := \sigma(\delta)\lambda(\delta)$$

Definition 15 (Behavior of weighted contextual grammar in internal mode) *Let S be a semiring and G be a Weighted Contextual Grammar over S with choice. A formal power series $||G||_{in}$ realized by grammar G in internal mode is a series in $S\langle\langle\Sigma^*\rangle\rangle$ given as*

$$||G||_{in} = \sum_{\delta \in \mathcal{D}_{in}(G)} ||\delta||$$

In the same sense as in external mode, we can now define weight of word in $||G||_{in}$. This definition is a little more complicated as word consists of five parts in each derivation step. That also means, that different partitions no longer mean different contexts but also different places of application of the same context. For example, if we have word $w = aaa$ and $\varphi(a) = (a, \varepsilon)$ there are three ways to apply it that all result into word $aaaa$ but the derivation and partition is different.

Lemma 9 (Weight of w in $||G||_{in}$) *Let $Par_5(w)$ be set of all partitions of word w to five words (w_1, u, w_2, v, w_3) where $w = w_1uw_2vw_3$ and $|uv| > 0$. Weight of word w in formal power series realized by grammar G is sum*

$$(w, ||G||_{in}) = \iota(w) + \sum_{(w_1,u,w_2,v,w_3) \in Par_5(w)} (w_1w_2w_3, ||G||_{in})\nu(w_2, (u, v))$$

We can define the sum in this general way over all partitions, as we said in the definition of weighted grammar, that $\nu(w, c)$ is zero if context can not be adjoined to w.

Proof 2 *Let D_w be a set of derivations $\delta \in \mathcal{D}_{in}(G)$ such that $\lambda(\delta) = w$. We can say that weight of word w is sum of weights of derivations from D_w. We will now show that for calculation of weight, this set of derivations can be simplified to axiom and pairs consisting of word divided to three parts and last context adjoined between those parts. Each derivation $\delta \in D_w$ is either axiom or can be divided to derivation of previous word divided to three parts (w_1, w_2, w_3) and last adjoined context (u, v). In D_w there can be many derivations of word w that end with the same context in the same position. Lets say, that derivations $\delta_1 \dots \delta_n$ for some n end with the same partition $w_1(u)w_2(v)w_3$. Each of these derivations will be divided to the same pair of divided word (w_1, w_2, w_3) and context (u, v). We know that each of these derivations of $w' = w_1w_2w_3$ is different, and we also know that there has to be every possible derivation of w'. If there was some derivation δ' of w' that is not contained in $\delta_1 \dots \delta_n$ we could adjoin (u, v) in the right place and we have new derivation of w which is contradiction.*

Now, as we know that there are all derivations of w', we also know that sum of the weights of these derivations is weight of word w'. In semiring multiplication distributes over summation so we can write $(w', ||G||_{in})\nu(w_2, (u, v))$ instead of $\sum_{\delta \in D_{w'}} \sigma(\delta)\nu(w_2, (u, v))$. Note that this is just sum of weights of all derivation with the same last step.

$$(w, ||G||_{in}) = \sum_{\delta \in D_w} \sigma(\delta)$$

$$= \iota(w) + \sum_{(w_1,u,w_2,v,w_3) \in Par_5(w)} \sum_{\delta \in D_{w_1w_2w_3}} \sigma(\delta)\nu(w_2, (u, v))$$

$$= \iota(w) + \sum_{(w_1,u,w_2,v,w_3) \in Par_5(w)} (w_1w_2w_3, ||G||_{in})\nu(w_2, c)$$

So the weight of word w in $||G||_{in}$ is equal to the sum as we stated in 9.

2.3 Examples

Example 5 *Let $G = (\{a, b\}, \{\varepsilon\}, \{(a, b)\}, \varphi, \nu, \iota)$ where*
$\varphi(a^nb^n) = \{(a, b)\}$
$\iota(\varepsilon) = 1$
$\nu(a^nb^n, (a, b)) = 2$

$||G||_{ex}$ is a WECC over semiring $S = (\mathbb{N}, +, \cdot, 0, 1)$.

This grammar has only one context and the only derivation step that can be always done is to add this context. The language generated by this grammar is $L(G) = \{a^nb^n\}$ and for each word there is only one derivation. Each application of context (a, b)

multiplies weight by two which means that weight of each word $a^n b^n$ is 2^n. The formal power series realized by this grammar is

$$||G||_{ex} = \sum_{n \in \mathbb{N}} 2^n a^n b^n$$

In the internal mode, each word in the form $a^n b^n$ can be derived in n ways from the word $a^{n-1}b^{n-1}$. That means that weights of first few words are $1, 2, 2{\cdot}2^2, 3{\cdot}2{\cdot}2^3, 4{\cdot}3{\cdot}2{\cdot}2^4$ so as we can see

$$||G||_{in} = \sum_{n \in \mathbb{N}} n! 2^n a^n b^n$$

Example 6 *Let* $G = (\{a, b\}, \{\varepsilon\}, \{(\varepsilon, a), (\varepsilon, b)\}, \varphi, \nu, \iota)$ *where*
$\varphi(w) = \{(\varepsilon, a), (\varepsilon, b)\}, \forall w \in \{a, b\}^*$
$\iota(\varepsilon) = 0$
$\nu(w, c) = 1, \forall w \in \{a, b\}^*, \forall c \in \{(\varepsilon, a), (\varepsilon, b)\}$

$||G||_{ex}$ *is a WECC over semiring* $S = (\mathbb{N}, \cdot, +, 1, 0)$.

The semiring S has switched operations and both contexts add one character to the end of the word. Each word has only one derivation, and weight of derivation is sum of ones for each character, or in other words length of word. As we will later mention, for languages in ECC there always exist a contextual grammar where each word has only one derivation. By extending such grammar to weighted variant over this semiring in the same way, with weight of context being its length, the weights of words will be again lengths of them.

Chapter 3

Properties of formal power series realized by weighted Contextual Grammars

In this chapter, we will prove some results that are known from contextual grammars without weight. Most of contextual grammars are not closed under classical operations like union, intersection, complement, concatenation, substitution or intersection with regular language. Mirror image is the only exception in all four classes (EC, ECC, IC, ICC), but there is not much to prove about it, as the construction is trivial in all four variants. The other exception is closure of ECC under union, that will result into closure of formal power series realised by WECC under sum.

3.1 Closure of WECC over sum

Theorem 1 (WECC is closed under sum) *Let G_1 and G_2 be weighted contextual grammars over semiring S. There is a weighted contextual grammar G over semiring S such that $||G|| = ||G_1|| + ||G_2||$.*

The closure of ECC under union is usually proven as a simple corollary of property of ECC languages called "EBS" which however is not sufficient to weighted variant. In weighted case we have to take care of weights and make sure that for the union language there is a grammar that sums weights of words from two languages in the right way. We will provide a constructive proof.

The weighted contextual grammar as defined earlier can contain contexts that are never used as well as selection mapping can contain selectors that can never be used (they select word that can not be derived). This property could result into new unwanted derived words when two grammars are unioned as such selector could be usable in the second grammar.

To avoid this, we define normal form which avoids such selectors. This has to be done separately for external and internal mode of derivation, as they can generate very different languages so they might use very different selectors.

Definition 16 (Normal form of weighted contextual grammar with choice) *Let $G = (\Sigma, A, C, \varphi, \nu, \iota)$ be a weighted contextual grammar over semiring S. We say that G is in normal form in external (internal) mode if each $c \in C$ and each $c \in \varphi(w), \forall w \in \Sigma^*$ is used in some derivation in $\mathcal{D}_{ex}(G)$ ($\mathcal{D}_{in}(G)$).*

We can see from the definition, that for each weighted contextual grammar there are equivalent grammars in normal form for external and internal derivation. This definition might look as construction, however it is only existentional as the procedure described in definition can be infinite.

Proof 3 *Let $G_1 = (\Sigma_1, A_1, C_1, \varphi_1, \nu_1, \iota_1)$ and $G_2 = (\Sigma_2, A_2, C_2, \varphi_2, \nu_2, \iota_2)$ be weighted contextual grammars over semiring S in the normal form in external mode.*
$\Sigma = \Sigma_1 \cup \Sigma_2$
$A = A_1 \cup A_2$
$C = C_1 \cup C_2$
$\varphi(w) = \varphi_1(w) \cup \varphi_2(w), \ \forall w \in \Sigma^*$
$\nu(w, c) = \nu_1(w, c) + \nu_2(w, c), \ \forall w \in \Sigma^*, \forall c \in C$
$\iota(a) = \iota_1(a) + \iota_2(a), \ \forall a \in A$

We will show that $G = (\Sigma, A, C, \varphi, \nu, \iota)$ is such that $||G|| = ||G_1|| + ||G_2||$

The weight of some word $w \in \Sigma^*$ in $||G||$ is sum of weights of all derivations of w in G. We will show that these derivations correspond to derivations in G_1 and G_2.

We want to show, that $(w, ||G_1||) + (w, ||G_2||) = (w, ||G||)$. From the definition, weight of the word $(w, ||G||)$ is sum of weights of the derivations of this word in grammar G. As shown earlier, this sum can be rearranged to the following formula:

$$(w, ||G||) = \iota(w) + \sum_{(w',c') \in Par(w)} (w', ||G||)\nu(w', c)$$

We will rearrange the sum of $(w, ||G_1||)$ and $(w, ||G_2||)$ using the same formula and prove that it equals the formula for $(w, ||G||)$.

$$(w, ||G_1||) + (w, ||G_2||) = \iota_1(w) + \sum_{(w',c') \in Par(w)} (w', ||G_1||)\nu_1(w', c)$$

$$+\iota_2(w) + \sum_{(w',c') \in Par(w)} (w', ||G_2||)\nu_2(w', c) =$$

$$\iota_1(w) + \iota_2(w) + \sum_{(w',c') \in Par(w)} (w', ||G_1||)\nu_1(w', c) + (w', ||G_2||)\nu_2(w', c) =$$

Using the fact that $|w'| < |w|$ *we have* $(w', ||G_1||) + (w', ||G_2||) = (w', ||G||)$ *by induction on length of the word and* $\iota_1(w) + \iota_2(w) = \iota(w)$ *from the definition.*

$$\iota(w) + \sum_{(w',c') \in Par(w)} (w', ||G||)(\nu_1(w', c) + \nu_2(w', c)) =$$

And from the definition we finally get

$$\iota(w) + \sum_{(w',c') \in Par(w)} (w', ||G||)\nu(w', c) = (w, ||G||)$$

3.2 Ambiguity of WECC

Another property, that holds specially for ECC is unambiguity. Language L is unambiguous with respect to external contextual grammars with choice, if there is grammar G such that $L_{ex}(G) = L$ and for each word in L there is only one derivation in G.

Ambiguity is very relevant linguistic property of grammar or language, as it tells whether we can say the syntactic structure of derivation from the word. Sadly, it is undecidable whether contextual grammar is ambiguous or not.

We will present simple counterexample that in WECC are not all power series unambiguous. On the other hand, we can easily see, that over some semirings, all power series are unambiguous with respect to weighted contextual grammars with choice. We will present some sufficient conditions and examples.

Definition 17 (Ambiguity of weighted contextual grammars with choice) *We say that grammar* G *is ambiguous in external (internal) mode if there is a word* $w \in \Sigma^*$ *such that cardinality of set of all derivations of* w *in external (internal) mode is greater than one* $|D_w| > 1$.

Definition 18 (Ambiguity of formal power series) *Formal power series* $F \in S\langle\langle\Sigma^*\rangle\rangle$ *is ambiguous with respect to weighted formal power series with choice in external (internal) mode, if every such grammar* G *that realizes* F *in external (internal) mode is ambiguous.*

Theorem 2 (Ambiguous formal power series in WECC) *There are infinitely many formal power series in WECC that are ambiguous.*

Proof 4 *Let* $G = (\{a, b\}, \{\varepsilon, b\}, \{(b, \varepsilon), (a, \varepsilon), (b, b)\}, \varphi, \nu, \iota)$ *where*
$\varphi(a^n b) = (a, \varepsilon), (b, \varepsilon)$
$\varphi(a^n) = (a, \varepsilon), (b, b)$
$\nu(a^n b, (a, \varepsilon)) = 1$
$\nu(a^n b, (b, \varepsilon)) = 1$

$\nu(a^n, (a, \varepsilon)) = 1$
$\nu(a^n, (b, b)) = n$
$\iota(\varepsilon) = 4$
$\iota(b) = 3$
be a weighted contextual grammar with choice over semiring $(\mathbb{N}, +, \cdot)$ *in external mode.*

Formal power series realized by this grammar is

$$||G||_{ex} = \sum_{n \in \mathbb{N}} 4a^n + 3a^n b + (4n+3)ba^n b$$

We can see, that there are infinitely many words with weight $4n + 3$*. According to the Dirrichlet theorem the sequence of integers in form* $4n + 3$ *contains infinitely many primes.[9]*

Weight of word is calculated as sum of products. If there was an unabiguous grammar that realizes the same formal power series as G, than this prime weight would be calculated just as product of some numbers. We however know that prime numbers can be calculated only as product of ones and itself. As there are infinitely many such words, they can not be all axioms, which mean, there has to be axiom with weight one. That is contradiction to $||G||_{ex}$ *as all weights of words are greater than one.*

We will now prove, that ambiguity depends on semiring and weighted variants of languages in ECC over some semirings are unambiguous. The sufficient condition on semiring for weighted language in WECC to be unambiguous that we prove here is that there are inverse elements for multiplication. The appropriate algebraic structure is semifield.

Definition 19 (Semifield) *Let* $(S, +, \cdot)$ *be a semiring. We say that it is semifield, if for each* $x, y \in S$ *such that* $y \neq 0$ *there is unique* $z_l, z_r \in S$ *for which* $x = z_l \cdot y$ *and* $x = y \cdot z_r$*.*

Theorem 3 (Unambiguity of WECC over semifield) *For each language in WECC with weights over semifield exists weighted contextual grammar G where* $||G||_{ex} = L$ *and each word have only one derivation in G.*

The proof is existentional based on induction. As we have seen earlier, the weight of some word in external mode depends only on weight of word w' it is derived from and weight of the transition used. If we say, that w' has only one derivation from the induction, the only problem that holds is, that there still can be multiple different combinations of w' and some context that derive the same word. We will solve this by choosing arbitrary derivation and removing all selectors used in the last step of all other derivations. Weight of this derivation step will be than adjusted to generate proper weight of w.

Consider now a weighted contextual grammar with choice $G = (\Sigma, A, C, \varphi, \nu, \iota)$ We construct a sequence $G_i = (\Sigma, A, C, \varphi_i, \nu_i, \iota'), i > 0$, of weighted contextual grammars as follows. The first grammar G_0 will be empty.
$G_0 = (\Sigma, \emptyset, C, \varphi_0, \nu_0, \iota_0),\ \varphi_0(x) = \emptyset,\ \forall x \in V^*, \nu_1(w, c) = 0\ \forall (w, c) \in V^* \times C$

Some axioms can also have multiple derivations in G, so we say, that the only right derivation is axiom itself and we adjust weight to be as in $||G||_{ex}$.

$\iota'(w) = (w, ||G||),\ \forall w \in A$
$G_1 = (\Sigma, A, C, \varphi_1, \nu_1, \iota'),\ \varphi_1(x) = \emptyset,\ \forall x \in V^*, \nu_1(w, c) = 0\ \forall (w, c) \in V^* \times C$
For $i > 0$ we consider
$G_{i+1} = (\Sigma, A, C, \varphi_{i+1}, \nu_{i+1}, \iota'),$
with the mappings φ_{i+1} and nu_{i+1} defined as follows.

φ_{i+1} is always defined only on words w such that $(w, ||G_i||_{ex}) \neq 0$.

If $\varphi(w)$ is already defined in previous grammar, we do not modify it. For w such that $(w, ||G_{i-1}||_{ex}) \neq 0$ we put $\varphi_{i+1}(w) = \varphi_i(w)$. The same way we define for those w weight of contexts $\nu_{i+1}(w, c) = \nu_i(w, c), \forall c \in \varphi_i(w)$

For every new word in G_i we will construct φ such way, that no ambiguity can be introduced. Take iteratively each w such that $(w, ||G_i||_{ex} \neq 0$ and $(w, ||G_{i-1}||_{ex}) = 0$ and consider iteratively each context $(u, v) \in \varphi(w)$. If $(uwv||G_i||_{ex}) \neq 0$, then we remove (u, v) from $\varphi(w)$ and set $\nu(w, (u, v)) = 0$. If $(uwv||G_i||_{ex}) = 0$, then we consider the set of all possible derivations of uwv from some word in $||G_i||$

$$M_{uwv} = \{(y, (u', v'))|(y, ||G_i||_{ex} \neq 0, (u', v') \in \varphi(y), u'yv' = uwv\}$$

Assume that there are $n \geq 1$ such derivations

$$M_{uwv} = \{(y_1, (u_1, v_1)), (y_2, (u_2, v_2)), \ldots (y_n, (u_n, v_n))\}$$

with first being the original one $(y_1, (u_1, v_1)) = (w, (u, v))$. If there are more than one such derivations, then all contexts (u_j, v_j) with $j > 1$ are removed from corresponding $\varphi(y_j)$ and weight is set to zero $\nu(y_j, (u_j, v_j)) = 0$.

After iteration over all $(u, v) \in \varphi(w)$ and for each w such that $(w, ||G_i||_{ex} \neq 0$ and $(w, ||G_{i-1}||_{ex}) = 0$, which is also finite number of steps, we can set new $\varphi_{i+i}(w) = \varphi(w)$. The last thing to set is weight of these derivation steps. For each such w and each $(u, v) \in \varphi(w)$ we will set $\nu(w, (u, v) = z$ where $(w, ||G||_{ex}) \cdot z = (uwv, ||G||_{ex})$. From the definition of semifield, we know that such z exists.

In this way, at the limit we obtain a grammar $G' = (\Sigma, A, C, \varphi_{\inf}, \nu_\infty, \iota')$ such that $||G'||_{ex} = ||G||_{ex}$ and for each word there is exactly one derivation in G'. Therefore, G' is unambiguous, so $||G||_{ex}$ is unambiguous.

Chapter 4

Open problems

In this chapter we provide a list of problems that are worth of study. As the fist, we have to mention that most of the classical questions that stayed unanswered like other closure properties of formal power series realized by weighted contextual grammars are in general in the most cases trivially extendable from non-weighted variant as they are counterexamples. There are however a few worth a look.

Other more interesting question is whether $REG \subset ICC$ can be extended to weighted variant and if not, whether it holds for some class of semirings. This is particularly interesting because it means thet regular rewriting can be simulated by context adjoining in some way.

For the ambiguity of WECC we gave one sufficient condition. It might be worth a look to find some other sufficient or necessary condition or even sufficient and necessary.

Contextual grammars with maximal use of selectors, as we defined in the first chapter are mentioned as good candidate to be extended to weighted variant, however they are not in this thesis. There have some interesting properties even thou there is not that much results to prove in weighted variant. We however think that parser for such grammar in weighted variant could be interesting and worth study.

In the whole thesis we avoided the fact that more standard weighted models are automata than grammars, but we work with grammars. The reason is that automata equivalent with contextual grammars seems somewhat complicated and artificial. It might be still worth a look whether they can be extended to weights in some useful way.

The big class of unanswered questions are all other variants of contextual grammars and derived models such as grammars with erased contexts, one-side contexts, deterministic grammars, leftmost derivation, parallel derivation, blocked derivation, depth-first derivation, marked derivation, grammars with infinite number of contexts, two-level contextual grammars, programmed grammars, controlled grammars, insertion grammars and others. Some of them might be interesting and some might not even be worth a look. In fact, most of them are probably not interesting, but weighted variants can sometimes surprise.

Conclusion

The Goal of this thesis was to open the field of weighted contextual grammars with the basic definitions and to support these definitions by proving that basic properties of non-weighted variant hold.

In the first chapter we introduced concept of context in sense of contextual grammars and defined four basic classes of languages generated by the internal and external contextual grammars with and without choice. These variants are presented with simple examples to provide intuition behind context adjoining. Generative capacity is presented without proofs, as they are either trivial, complicated or not so important for this thesis. Results about closure properties are also presented without proofs as most of proofs are just counterexamples. Two more modes of derivation in contextual grammars are presented next being maximal global and maximal local use of selectors. These modes, however more complicated, model properties of natural languages even better.

The second part of this chapter is dedicated to standard formalism in study of weighted automata. First we defined formal power series with multiple non-commuting variables and coefficients from semiring. This is the best formalism to describe behavior of some weighted model, or in other words to describe set of words with weights from semiring. One of the advantages of this formalism is that we can simply extend union of languages to sum of formal power series. This chapter holds necessary preliminaries for further results in this thesis.

Weighed contextual grammars are defined in the second chapter in the four variants as well. Weighted contextual grammars with choice are defined as extension of contextual grammars with choice and weighted contextual grammars without choice are defined as extension of non-weighted variant without choice, however we also present it as special case of weighed contextual grammars with choice. Behavior of these grammars is defined in both external and internal mode. In both modes we defined formal power series realized by weighted contextual grammar with choice as sum of all derivations in the grammar in corresponding mode. We also show that weight of word in such grammar can be calculated from weights of words it can be derived from in one step. Behavior of weighted contextual grammars without choice can be viewed as special case of variant with choice, so we do not give much attention to it. These definitions are the core of results in this thesis as they open the field of weighted contextual grammars over semiring. All the following results are based on these definitions and we believe that even future results will stand on these definitions.

Basic properties of formal power series realized by grammars as we defined them are studied in the third chapter. In particular we provided a proof of closure of formal power series realized by weighted contextual grammars with choice in external mode over sum. This result can be also interpreted as closure of non-weighted variant over union which is one of a few positive closure results in the field of contextual grammars. The proof we presented is also different from the standard proof in non-weighted variant. This result also supports our definition as the right extension of non-weighted variant. Another property we decided to study is ambiguity. It is particularly interesting in the linguistic view of language and also results into some interesting outcomes. Languages generated by external contextual grammars with choice are in general unambiguous, which mean that there is grammar that can generate each word in only one way. We showed that in weighted variant it depends on semiring. We present example and proof of inherently ambiguous formal power series realized by weighted contextual grammar with choice in external mode over standard semiring of natural numbers as well as sufficient condition on semiring for all power series to be unambiguous with respect to weighted contextual grammars over such semiring in external mode.

Bibliography

[1] Noam Chomsky. On certain formal properties of grammars. *Information and control*, 2(2):137–167, 1959.

[2] Manfred Droste, Werner Kuich, and Heiko Vogler. *Handbook of weighted automata.* Springer Science & Business Media, 2009.

[3] Karin Harbusch. Parsing contextual grammars with linear, regular and context–free selectors. *Grammars and Automata for String Processing. Topics in Computer Mathematics*, 9:45–54, 2004.

[4] Solomon Marcus. Contextual grammars. In *INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS COLING 1969: Preprint No. 48*, 1969.

[5] Solomon Marcus, Gheorghe Păun, and Carlos Martín-Vide. Contextual grammars as generative models of natural languages. *Computational Linguistics*, 24(2):245–274, 1998.

[6] Carlos Martín-Vide, Alexandru Mateescu, Joan Miquel-Vergés, and Gheorghe Păun. Internal contextual grammars: minimal, maximal, and scattered use of selectors. In *Proceedings of The Fourth Bar-Ilan Symposium on Foundations of Artificial Intelligence. Focusing on Natural Languages and Artificial Intelligence-Philosophical and Computational Aspects*, pages 159–168, 1995.

[7] Gheorghe Păun. Marcus contextual grammars, 1997.

[8] Marcel Paul Schützenberger. On the definition of a family of automata. *Information and control*, 4(2-3):245–270, 1961.

[9] Anthony Várilly. Dirichlet's theorem on arithmetic progressions. *Harvard University, Cambridge, MA*, 2138, 2014.