Scientific Annals of Computer Science vol. 29 (1), 2019, pp. 59–80
doi: 10.7561/SACS.2019.1.59

Weighted Context-Free Grammars Over Bimonoids

George Rahonis[1], Faidra Torpari[2]

Abstract

We introduce and investigate weighted context-free grammars over an arbitrary bimonoid K. Thus, we do not assume that the operations of K are commutative or idempotent or they distribute over each other. We prove a Chomsky-Schützenberger type theorem for the series generated by our grammars. Moreover, we show that the class of series generated by weighted right-linear grammars over a linearly ordered alphabet Σ and K coincides with that of recognizable series over Σ and K.

Keywords: weights, context-free grammars, bimonoids.

1 Introduction

Weighted models of computation assign quantitative features to computational processes. For instance finite automata examine whether an input word is accepted or not whereas weighted automata provide information for the cost of the computation, energy consumption, probability of the implementation of the computation, etc. On the other hand, context-free grammars constitute the main generative model with interesting applications in compilers' development (cf. for instance [17]), model checking [15],

[1]Department of Mathematics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, Email: grahonis@math.auth.gr

[2]Department of Mathematics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, Email: ftorpari@yahoo.gr

parameterized verification [5], and runtime verification [23]. Both, weighted automata and weighted context-free grammars have been widely studied over semirings. We refer the reader to [7, 12, 13, 18, 24] for theory and applications of weighted automata over semirings, and to [2, 3, 14, 25] for several results on weighted context-free grammars over semirings.

The last years, it came up that the semiring structure is not sufficient to describe operations needed in modern practical applications, like for instance the average operation. Therefore, several authors built the theory of computational models over more general structures, namely strong bimonoids and valuation monoids [4, 10, 11, 19, 27]. More recently, McCarthy-Kleene logic contributed to the development of an application for runtime verification, within the projects LogicGuard I and LogicGuard II [20, 21]. The extension of that tool, for future applications, required a fuzzy type of McCarthy-Kleene logic as well as weighted computational models. It was proved that the reasonable weight structure is a particular zero-sum free and zero-divisor free bimonoid with only left multiplicative zero. Weighted automata over that bimonoid were investigated in [8, 9][3].

It is the goal of this paper to introduce and investigate a weighted context-free grammar model over an arbitrary bimonoid. Since the commutativity of operations of the weight structure is not required, we equipped the rule sets of our grammars with a linear order. Our main results are as follows.

- We prove a Chomsky-Schützenberger theorem for the class of series generated by weighted context-free grammars. As an intermediate step, we show a folklore result, namely for every weighted context-free grammar we can effectively construct an equivalent one in Chomsky normal form permitting also ε-rules.

- We consider weighted right-linear grammars and show their expressive equivalence to weighted automata models of [8, 9], over an arbitrary bimonoid. For this, we require that the input alphabet is linearly ordered.

The structure of our paper is as follows. Apart from this Introduction the paper contains four sections. In Section 2 we recall notions needed in the sequel and an example of a zero-sum free and zero-divisor free bimonoid

[3] In [8, 9] weighted automata were called MK-fuzzy automata due to the specific weight structure motivated by the fuzzification of McCarthy-Kleene logic.

with left multiplicative zero [8, 9]. In Section 3 we introduce our weighted context-free grammars and show the closure of the class of their series under sum. Furthermore, we show that the class of series generated by unambiguous grammars is closed under multiplication with scalars from the right, provided the bimonoid has left multiplicative zero. Then, in Section 4 we prove the Chomsky-Schützenberger type theorem, and in Section 5 we show the expressive equivalence of weighted right-linear grammars and weighted automata. Finally, in the Conclusion we refer to open problems for future research.

The results of the paper were presented at 9th International Workshop *Weighted Automata: Theory and Applications, WATA 2018* [28].

2 Preliminaries

Let Σ be an alphabet, i.e., a finite nonempty set. We denote by Σ^* the set of all finite words over Σ, i.e., the free monoid generated by Σ and set $\Sigma^+ = \Sigma^* \setminus \{\varepsilon\}$, where ε is the empty word. Assume now that $\leq$ is a linear order on Σ. As usually, we let $a < b$ iff $a \leq b$ and $a \neq b$, and we keep this notation for every order defined in the paper. The lexicographic order $\leq_{\text{lex}}$ on Σ^* is defined as follows. We let $w \leq_{\text{lex}} u$ iff

$$(u = wv \text{ with } v \in \Sigma^*) \quad \text{or}$$

$$(w = vav',\ u = vbv'' \ \text{ with } v, v', v'' \in \Sigma^*,\ a, b \in \Sigma \text{ and } a < b)$$

for every $w, u \in \Sigma^*$.

In the sequel, the lexicographic order is defined, as for Σ^*, on the free monoid generated by any finite linearly ordered set.

A *bimonoid* $(K, +, \cdot, 0, 1)$ (cf. [10]) consists of a set K, two binary operations $+$ and $\cdot$ and two constant elements 0 and 1 such that $(K, +, 0)$ and $(K, \cdot, 1)$ are monoids. The bimonoid is denoted simply by K if the operations and the constant elements are understood. If no confusion arises, we shall denote the $\cdot$ operation also by juxtaposition. The bimonoid K has *left* (resp. *right*) *multiplicative zero* if 0 acts as a left (resp. right) multiplicative zero, i.e., $0 \cdot k = 0$ (resp. $k \cdot 0 = 0$) for every $k \in K$. If the monoid $(K, +, 0)$ is commutative and 0 acts as a left and right multiplicative zero, then the bimonoid is called *strong*. A *semiring* is a strong bimonoid where multiplication distributes over addition.

Next we recall from [8, 9], an example of a bimonoid with left multiplicative zero. The set of the monoid consists of quadruples of the form (t, f, u, e)

where $t, f, u, e \in [0,1]$ and $t+f+u+e=1$. The operations $\sqcup$ and $\sqcap$ of the monoid are motivated by the fuzzification of disjunction and conjunction of McCarthy-Kleene logic [16, 22], respectively. Weighted automata over that bimonoid were introduced and studied in [8, 9] as computational models for the quantitative runtime verification within the LogicGuard projects [20, 21].

Example 1 [8, 9] *We consider the set*

$$K = \{(t,f,u,e) \in [0,1]^4 \mid t+f+u+e=1\}$$

and define on K two binary operations $\sqcup$ and $\sqcap$ as follows. For every $\mathbf{k_1} = (t_1, f_1, u_1, e_1) \in K, \mathbf{k_2} = (t_2, f_2, u_2, e_2) \in K$ we let $\mathbf{k_3} = \mathbf{k_1} \sqcup \mathbf{k_2}$ and $\mathbf{k_4} = \mathbf{k_1} \sqcap \mathbf{k_2}$ where $\mathbf{k_3} = (t_3, f_3, u_3, e_3)$ and $\mathbf{k_4} = (t_4, f_4, u_4, e_4)$ are defined by the relations

$$\begin{array}{ll} t_3 = t_1 + (f_1 + u_1)t_2 & t_4 = t_1 t_2 \\ f_3 = f_1 f_2 & f_4 = f_1 + (t_1 + u_1) f_2 \\ u_3 = f_1 u_2 + u_1(f_2 + u_2) & u_4 = t_1 u_2 + u_1(t_2 + u_2) \\ e_3 = e_1 + (f_1 + u_1)e_2 & e_4 = e_1 + (t_1 + u_1)e_2. \end{array}$$

Then the structure $(K, \sqcup, \sqcap, \mathbf{0}, \mathbf{1})$, where $\mathbf{0} = (0,1,0,0)$ and $\mathbf{1} = (1,0,0,0)$, is a bimonoid with left multiplicative zero. Furthermore, it is zero-sum free and zero-divisor free, i.e., for every $\mathbf{k}, \mathbf{k}' \in K$, $\mathbf{k} \sqcup \mathbf{k}' = \mathbf{0}$ implies $\mathbf{k} = \mathbf{k}' = \mathbf{0}$, and $\mathbf{k} \sqcap \mathbf{k}' = \mathbf{0}$ implies $\mathbf{k} = \mathbf{0}$ or $\mathbf{k}' = \mathbf{0}$, respectively.

We refer the reader to [26] for further examples of bimonoids.

Throughout the paper $(K, +, \cdot, 0, 1)$ denotes an arbitrary bimonoid.

Let Q be a set. A *formal series* (or simply *series*) *over Q and K* is a mapping $s : Q \to K$. The *support of s* is the set $\mathrm{supp}(s) = \{q \in Q \mid s(q) \neq 0\}$. A series with finite support is called also a *polynomial*. The *constant series* $\widetilde{k}$ ($k \in K$) is defined, for every $q \in Q$, by $\widetilde{k}(q) = k$. We denote by $K\langle\langle Q \rangle\rangle$ the class of all series over Q and K. Let $s, r \in K\langle\langle Q \rangle\rangle$ and $k \in K$. The *sum* $s + r$, the *products with scalars* ks and sk, and the *Hadamard product* $s \odot r$ are defined elementwise, respectively by $(s+r)(q) = s(q) + r(q), (ks)(q) = ks(q), (sk)(q) = s(q)k, (s \odot r)(q) = s(q)r(q)$ for every $q \in Q$. Trivially, the structure $(K\langle\langle Q \rangle\rangle, +, \odot, \widetilde{0}, \widetilde{1})$ is a bimonoid. As usual, we write a series $s \in K\langle\langle Q \rangle\rangle$ in the form $s = \sum_{q \in Q} s(q).q$.

3 Weighted Context-Free Grammars

In this section we introduce the concept of weighted context-free grammars over the bimonoid K and state two closure properties of the class of their series. For this, we recall firstly the notion of a context-free grammar. More precisely, a *context-free grammar* is a quadruple $\mathcal{G} = (\Sigma, N, S, R)$ where Σ is the alphabet of *terminals*, N is the alphabet of *variables* (or *nonterminals*) with $\Sigma \cap N = \emptyset$, $S \in N$ is the *initial variable*, and $R \subseteq N \times (\Sigma \cup N)^*$ is the finite set of *production rules* (or simply *rules*). We use capital letters A, B, C, etc. to denote the elements of N. As usual, a rule $(A, v) \in R$ is also written as $A \to v$. A rule $r \in R$ is called a *chain rule* if it is of the form $r = (A \to B)$ with $B \in N$, and it is called an *ε-rule* if it is of the form $r = (A \to \varepsilon)$. The elements of $(\Sigma \cup N)^*$ are called *sentential forms* of $\mathcal{G}$. We define the *direct derivation relation* $\Longrightarrow_{\mathcal{G}} \subseteq (\Sigma \cup N)^* \times (\Sigma \cup N)^*$ as follows. Let $w, u \in (\Sigma \cup N)^*$. Then, we set $w \Longrightarrow_{\mathcal{G}} u$ iff $w = w_1 A w_2$, $u = w_1 v w_2$ with $w_1, w_2 \in (\Sigma \cup N)^*$ and there is a rule $r = (A \to v) \in R$. Sometimes we also write $w \overset{r}{\Longrightarrow}_{\mathcal{G}} u$ if we want to denote that w directly derives u with the application of rule r. The direct derivation $w \Longrightarrow_{\mathcal{G}} u$ is called *leftmost* if $w_1 \in \Sigma^*$. In the sequel, the term derivation refers only to leftmost direct derivations and if no confusion arises we simply write $\Longrightarrow$ instead of $\Longrightarrow_{\mathcal{G}}$. As usual, we denote by $\Longrightarrow^*$ the reflexive and transitive closure of $\Longrightarrow$. Sometimes by a *derivation of* $\mathcal{G}$ we refer to a finite sequence $d = r_0 \ldots r_{n-1}$, $n \geq 1$, of rules $r_i \in R$, $0 \leq i \leq n-1$, such that there are sentential forms $w_i \in (\Sigma \cup N)^*$ with $w_i \overset{r_i}{\Longrightarrow} w_{i+1}$ for every $0 \leq i \leq n-1$. In this case we write $w_0 \overset{d}{\Longrightarrow} w_n$. For every $A \in N$, $w \in \Sigma^*$ an *A-derivation* of w is a derivation d of $\mathcal{G}$ such that $A \overset{d}{\Longrightarrow} w$. We denote by $D(A, w)$ the set of all A-derivations of w and we simply write $D(w)$ for $A = S$. The *language generated by* $\mathcal{G}$ is $L(\mathcal{G}) = \{w \in \Sigma^* \mid D(w) \neq \emptyset\}$. A language $L \subseteq \Sigma^*$ is called *context-free* if there is a context-free grammar $\mathcal{G} = (\Sigma, N, S, R)$ such that $L = L(\mathcal{G})$. A context-free grammar $\mathcal{G} = (\Sigma, N, S, R)$ is called *unambiguous* if $|(D(w)| \leq 1$ for every $w \in \Sigma^*$.

For our weighted context-free grammars we shall need the notion of a loop-free derivation. More precisely, a derivation $w_0 \overset{d}{\Longrightarrow} w_n$ with $d = r_0 \ldots r_{n-1}$ is called *loop-free* if there are no indices $0 \leq j \leq k \leq n-1$ and $A \in N$ such that $A \overset{r_j \ldots r_k}{\Longrightarrow} A$. We shall denote by $lfD(A, w)$ (resp. $lfD(w)$) the set of all loop-free derivations in $D(A, w)$ (resp. in $D(w)$). It should be clear that for every $A \in N$ and $w \in \Sigma^*$ the set $lfD(A, w)$ is finite. Furthermore, if $\mathcal{G}$ is unambiguous, then every derivation of $\mathcal{G}$ is loop-free.

Definition 2 *A* weighted context-free grammar *(wcfg for short)* over Σ and K *is a five-tuple* $\mathcal{G} = (\Sigma, N, S, R, wt)$ *where* (Σ, N, S, R) *is a context-free grammar such that the set* R *of rules is assumed to be linearly ordered and* $wt : R \to K$ *is a mapping assigning* weights *to the rules.*

Let $w \stackrel{d}{\Longrightarrow} u$ with $w, u \in (\Sigma \cup N)^*$ and $d = r_0 \ldots r_{n-1},\ r_i \in R$, $0 \leq i \leq n-1$. The *weight* $weight(d)$ *of* d is determined by $weight(d) = wt(r_0) \cdot \ldots \cdot wt(r_{n-1})$. The *series* $\|\mathcal{G}\| \in K\langle\langle\Sigma^*\rangle\rangle$ *of* $\mathcal{G}$ is defined in the following way. Let $w \in \Sigma^*$ and assume that $lfD(w) = \{d_1, \ldots, d_m\} \neq \emptyset$ with $d_1 <_{\text{lex}} \ldots <_{\text{lex}} d_m$. Then, we set

$$\|\mathcal{G}\|(w) = \sum_{1 \leq i \leq m} weight(d_i)$$

where we sum up in an ascending order according to the usual ordering of natural numbers. If $lfD(w) = \emptyset$, then we let $\|\mathcal{G}\|(w) = 0$. We also say that $\|\mathcal{G}\|$ *is generated by* $\mathcal{G}$. Since $lfD(w)$ is finite for every $w \in \Sigma^*$ the value $\|\mathcal{G}\|(w)$ is well-defined. A series s over Σ and K is called *context-free* if there is a wcfg $\mathcal{G}$ over Σ and K such that $s = \|\mathcal{G}\|$. Two wcfg $\mathcal{G}_1, \mathcal{G}_2$ over Σ and K are called *equivalent* if $\|\mathcal{G}_1\| = \|\mathcal{G}_2\|$.

A wcfg $\mathcal{G} = (\Sigma, N, S, R, wt)$ over Σ and K is called *unambiguous* if the underlying context-free grammar (Σ, N, S, R) is unambiguous.

Proposition 3 *Let* $s_1, s_2 \in K\langle\langle\Sigma^*\rangle\rangle$ *be context-free series. Then, the sum* $s_1 + s_2$ *is also a context-free series.*

Proof: Let $\mathcal{G}_i = (\Sigma, N_i, S_i, R_i, wt_i)$ be wcfg over Σ and K such that $s_i = \|\mathcal{G}_i\|$ for $i = 1, 2$. Without loss of generality, we assume that $N_1 \cap N_2 = \emptyset$ and consider the wcfg $\mathcal{G} = (\Sigma, N, S, R, wt)$ over Σ and K with $N = N_1 \cup N_2 \cup \{S\}$ where S is a new variable, $R = R_1 \cup R_2 \cup \{S \to S_1, S \to S_2\}$, and

$$wr(r) = \begin{cases} wt_1(r) & \text{if } r \in R_1 \\ wt_2(r) & \text{if } r \in R_2 \\ 1 & \text{if } r = (S \to S_1) \text{ or } r = (S \to S_2) \end{cases}$$

for every $r \in R$.

We define a linear order on the set R of rules in the following way. We preserve the orders of R_1 and R_2 and we set $S \to S_1 \leq \min R_1$ and $\max R_1 \leq S \to S_2 \leq \min R_2$.

Let $w \in \Sigma^*$ such that $lfD(w) \neq \emptyset$. By definition of the set R, we trivially get that $lfD(w) = lfD_1(w) \cup lfD_2(w)$ where $lfD_i(w)$ denotes the set of loop-free derivations of the form $S \Longrightarrow_{\mathcal{G}} S_i \Longrightarrow^*_{\mathcal{G}_i} w$, for $i = 1, 2$. Furthermore, the order of R implies $\|\mathcal{G}\|(w) = \|\mathcal{G}_1\|(w) + \|\mathcal{G}_2\|(w)$. If $lfD(w) = \emptyset$, then $lfD_1(w) = lfD_2(w) = \emptyset$, hence again we get $\|\mathcal{G}\|(w) = \|\mathcal{G}_1\|(w) + \|\mathcal{G}_2\|(w)$, and our proof is completed. □

Proposition 4 *Let us assume that the bimonoid* $(K, +, \cdot, 0, 1)$ *has left multiplicative zero. Let also* $s \in K\langle\langle\Sigma^*\rangle\rangle$ *be the series of an unambiguous wcfg over* Σ *and* K*, and* $k \in K$*. Then* sk *is the series of an unambiguous wcfg over* Σ *and* K*.*

Proof: Let $\mathcal{G} = (\Sigma, N, S, R, wt)$ be an unambiguous wcfg over Σ and K such that $s = \|\mathcal{G}\|$. We consider the new variables S', A_k and the wcfg $\mathcal{G}' = (\Sigma, N', S', R', wt')$ over Σ and K with $N' = N \cup \{S', A_k\}$, $R' = R \cup \{S' \to SA_k, A_k \to \varepsilon\}$, $wt'(r) = wt(r)$ for every $r \in R$, $wt'(S' \to SA_k) = 1$, and $wt'(A_k \to \varepsilon) = k$. We extend the order on R to an order on R' by letting $S' \to SA_k \leq \min R$ and $\max R \leq A_k \to \varepsilon$. Let $w \in \Sigma^*$. Since $\mathcal{G}$ is unambiguous we get $|lf(D(w)| \leq 1$. Let us assume that $lfD(w) = \{d\}$ and $S \overset{d}{\Longrightarrow}_{\mathcal{G}} w$. By construction of R', we get that the derivation

$$d' : S' \Longrightarrow_{\mathcal{G}'} SA_k \overset{d}{\Longrightarrow}_{\mathcal{G}'} wA_k \Longrightarrow_{\mathcal{G}'} w$$

of $\mathcal{G}'$ for w is unique. By definition of wt', we get $weight(d') = weight(d)k$, hence $\|\mathcal{G}'\|(w) = (sk)(w)$. If $lfD(w) = \emptyset$, then trivially there is no S'-derivation of $\mathcal{G}'$ for w, and since 0 acts as left multiplicative zero, again we get $(sk)(w) = s(w) \cdot k = 0 \cdot k = \|\mathcal{G}'\|(w)$, and we are done. □

4 Chomsky-Schützenberger Theorem

In this section we show that a Chomsky-Schützenberger type result holds for the class of series generated by wcfg over Σ and K. For this, we prove firstly a folklore result, namely for every wcfg we can effectively construct an equivalent one in Chomsky normal form. Our definition for Chomsky normal form follows the one in [1], hence we permit ε-rules.

Definition 5 *A wcfg* $\mathcal{G} = (\Sigma, N, S, R, wt)$ *over* Σ *and* K *is said to be in* Chomsky normal form *if every rule* $r \in R$ *is of the form* $r = (A \to BC)$ *or* $r = (A \to a)$ *with* $B, C \in N$ *and* $a \in \Sigma \cup \{\varepsilon\}$*.*

By Definition 5, we get that if $\mathcal{G}$ is in Chomsky normal form, then it has no chain rules. We shall need the following lemma.

Lemma 6 *Let $\mathcal{G} = (\Sigma, N, S, R, wt)$ be a wcfg over Σ and K. Then we can effectively construct an equivalent wcfg $\mathcal{G}'$ without chain rules.*

Proof: If $\mathcal{G}$ contains no chain rules, then we set $\mathcal{G}' = \mathcal{G}$. Otherwise, we construct a wcfg $\mathcal{G}' = (\Sigma, N', S, R', wt')$ over Σ and K as follows. We consider a new symbol Y and let $N' = N \cup \{Y\}$. Then, we define R' by adding a new rule $Y \to \varepsilon$ to R, and replacing every chain rule $A \to B \in R$ by a new rule $A \to BY$. The weight mapping wt' coincides with wt on the non-chain rules of R and we set $wt'(A \to BY) = wt(A \to B)$ for every $A \to B \in R$ and $wt'(Y \to \varepsilon) = 1$. Finally, we extend the linear order on R to a linear order on R' by taking the order of R and replacing the rule $A \to B$ by the rule $A \to BY$, and letting $Y \to \varepsilon$ be the maximum element of R'. Then, it is straightforward to show the equivalence of $\mathcal{G}$ and $\mathcal{G}'$. $\square$

Proposition 7 *Let $\mathcal{G} = (\Sigma, N, S, R, wt)$ be a wcfg over Σ and K. Then, we can effectively construct an equivalent one in Chomsky normal form.*

Proof: By Lemma 6 we assume that $\mathcal{G}$ contains no chain rules. For every $a \in \Sigma$, we consider a new variable X_a and a new rule $X_a \to a$. We set $R_l = \{X_a \to a \mid a \in \Sigma\}$ and define an arbitrary linear order on it. Next, let $\tilde{R}$ comprise all rules of R of the form $A \to u_1 a_1 u_2 a_2 \ldots u_k a_k u_{k+1}$ with $k \geq 1$, $a_1, \ldots, a_k \in \Sigma$, and $u_1, \ldots, u_{k+1} \in N^*$ such that $u_1 u_2 a_2 \ldots a_k u_{k+1} \neq \varepsilon$. We replace in R every rule $A \to u_1 a_1 u_2 a_2 \ldots u_k a_k u_{k+1} \in \tilde{R}$ by a new rule $A \to u_1 X_{a_1} u_2 X_{a_2} \ldots u_k X_{a_k} u_{k+1}$, and obtain the set of rules $\overline{R}$. We define a linear order on $\overline{R}$ by taking the order on R and replacing every rule of the form $A \to u_1 a_1 \ldots u_k a_k u_{k+1} \in \tilde{R}$ by its corresponding one $A \to u_1 X_{a_1} \ldots u_k X_{a_k} u_{k+1}$. Moreover, we let $\max \overline{R} \leq \min R_l$. Now, we consider the wcfg $\mathcal{G}' = (\Sigma, N', R', S, wt')$ over Σ and K with $N' = N \cup \{X_a \mid a \in \Sigma\}$ and $R' = \overline{R} \cup R_l$. The weight mapping wt' is determined by

$$wt'(r) = \begin{cases} wt(r) & \text{if } r \in R \setminus \tilde{R} \\ wt(A \to u_1 a_1 \ldots u_k a_k u_{k+1}) & \text{if } r = (A \to u_1 X_{a_1} \ldots u_k X_{a_k} u_{k+1}) \\ 1 & \text{if } r \in R_l \end{cases}$$

for every $r \in R'$.

We aim to show that $\|\mathcal{G}'\| = \|\mathcal{G}\|$. Indeed, let $w \in \Sigma^*$ and assume that $lfD(w) = \{d_1, \ldots, d_n\}$ is the set of all loop-free derivations of $\mathcal{G}$ for w with

$d_1 <_{\text{lex}} \ldots <_{\text{lex}} d_n$. By definition of $\mathcal{G}'$, there is a derivation d_i' of $\mathcal{G}'$ for w, which corresponds to d_i for every $1 \le i \le n$, and vice versa. Moreover, we trivially get

$$weight(d_i') = weight(d_i).$$

It remains to prove that $d_1' <_{\text{lex}} \ldots <_{\text{lex}} d_n'$. For this, we fix an $1 \le i < n$ and let $d_i = r_1 \ldots r_m r u$ and $d_{i+1} = r_1 ... r_m p t$ with $r_1, \ldots, r_m \in R$, $r = (A \to x) \in R$, $p = (A \to y) \in R$, $u, t \in R^*$ and $r < p$ [4]. This implies that $x \neq y$, and

$$S \overset{r_1 \ldots r_m}{\Longrightarrow}_{\mathcal{G}} w_1 A v \overset{r}{\Longrightarrow}_{\mathcal{G}} w_1 x v \Longrightarrow_{\mathcal{G}}^* w$$

for d_i, and

$$S \overset{r_1 \ldots r_m}{\Longrightarrow}_{\mathcal{G}} w_1 A v \overset{p}{\Longrightarrow}_{\mathcal{G}} w_1 y v \Longrightarrow_{\mathcal{G}}^* w$$

for d_{i+1}, where $w_1 \in \Sigma^*$, and $x, y, v \in (\Sigma \cup N)^*$.

By construction of $\mathcal{G}'$ we get that

$$S \overset{r_{j_1}' \ldots r_{j_k}'}{\Longrightarrow}_{\mathcal{G}'} w_1 A v' \overset{r'}{\Longrightarrow}_{\mathcal{G}'} w_1 x' v' \Longrightarrow_{\mathcal{G}'}^* w$$

for d_i', with $r_{j_1}', \ldots, r_{j_k}' \in R'$, $r' = (A \to x') \in R'$, $w_1 \in \Sigma^*$, $x', v' \in (\Sigma \cup N')^*$ and $k \ge m$, where $x' = x$ if $x \in \Sigma \cup \{\varepsilon\}$, otherwise x' is obtained from x by replacing every letter $a \in \Sigma$ by the variable X_a, and

$$S \overset{r_{j_1}' \ldots r_{j_k}'}{\Longrightarrow}_{\mathcal{G}'} w_1 A v' \overset{p'}{\Longrightarrow}_{\mathcal{G}'} w_1 y' v' \Longrightarrow_{\mathcal{G}'}^* w$$

for d_{i+1}', with $r_{j_1}', \ldots, r_{j_k}' \in R'$, $p' = (A \to y') \in R'$, $w_1 \in \Sigma^*$, $y', v' \in (\Sigma \cup N')^*$, where $y' = y$ if $y \in \Sigma \cup \{\varepsilon\}$, otherwise y' is obtained from y by replacing every letter $a \in \Sigma$ by the variable X_a.
Since $r < p$, taking into account the order of R', we get $r' < p'$ which in turn implies that $d_i' <_{\text{lex}} d_{i+1}'$.

If $lfD(w) = \emptyset$, then obviously there is no derivation of w in $\mathcal{G}'$. We conclude $\|\mathcal{G}'\| = \|\mathcal{G}\|$, as required.

Next for every rule $r' \in R'$ of the form $r' = (A \to B_1 B_2 \ldots B_k)$ with $k \ge 3$, we consider the new variables $Y_1, Y_2, \ldots, Y_{k-2}$, and the new rules

$$p_1^{(r')} = (A \to B_1 Y_1)$$
$$p_2^{(r')} = (Y_1 \to B_2 Y_2)$$

[4]The case $d_i = r_1 \ldots r_m$ and $d_{i+1} = r_1 \ldots r_m v$ with $v \in R^+$ does not occur since $S \overset{d_i}{\Longrightarrow}_{\mathcal{G}} w$ means that the derivation terminates.

$$\vdots$$

$$p_{k-1}^{(r')} = (Y_{k-2} \to B_{k-1}B_k).$$

Then, we replace in R' the rule r' by its corresponding above list of rules and obtain a new set of rules R''. We define a linear order on R'' by taking the order of R' and replacing the rule r' by the linear order

$$p_1^{(r')} \leq p_2^{(r')} \leq \ldots \leq p_{k-1}^{(r')}.$$

We let also N'' be the set N' with the new variables obtained in the above procedure. Now, we consider the wcfg $\mathcal{G}'' = (\Sigma, N'', S, R'', wt'')$ over Σ and K where the weight mapping wt'' is defined as follows. It coincides with wt' on the rules of R'' which belong to R' and, keeping the above notations for r', we let

$$wt''\left(p_1^{(r')}\right) = wt'(r'), \text{ and}$$

$$wt''\left(p_2^{(r')}\right) = \ldots = wt''\left(p_{k-1}^{(r')}\right) = 1.$$

It should be clear that $\mathcal{G}''$ is in Chomsky normal form. Furthermore, let $w \in \Sigma^*$ such that the set of derivations $lfD'(w) = \{d'_1, \ldots, d'_n\}$ of w in $\mathcal{G}'$ is nonempty. Trivially, for every derivation d'_i, $1 \leq i \leq n$, there is a unique derivation d''_i of w in $\mathcal{G}''$ and vice versa. Furthermore, $weight(d''_i) = weight(d'_i)$ and by construction of R'', by standard arguments, we get $d''_1 <_{\text{lex}} \ldots <_{\text{lex}} d''_n$ whenever $d'_1 <_{\text{lex}} \ldots <_{\text{lex}} d'_n$. If $lfD'(w) = \emptyset$, then obviously there is no derivation of w in $\mathcal{G}''$. We conclude that $\|\mathcal{G}''\| = \|\mathcal{G}'\|$, and our proof is completed. $\square$

For our Chomsky-Schützenberger theorem, we still need some preliminary matter. First, we recall the notion of Dyck languages. More precisely, let Y be an alphabet and $\overline{Y} = \{\overline{y} \mid y \in Y\}$ a copy of it. Then, the *Dyck language over* Y, denoted by D_Y, is the context-free language generated by the grammar $\mathcal{G}_Y = (Y \cup \overline{Y}, N, S, R)$ with $N = \{S\}$ and $R = \{S \to yS\overline{y} \mid y \in Y\} \cup \{S \to SS,\ S \to \varepsilon\}$.

A polynomial $s \in K\langle\langle\Sigma^*\rangle\rangle$ is called a *monome* if $|(\text{supp(s)}| \leq 1$. We denote by $K[\Sigma \cup \{\varepsilon\}]$ the set of all monomes whose support is a subset of $\Sigma \cup \{\varepsilon\}$ [11].

Let I be an arbitrary index set and $(s_i)_{i\in I}$ a family of series in $K\langle\langle\Sigma^*\rangle\rangle$. For every $w \in \Sigma^*$ we let $I_w = \{i \in I \mid s_i(w) \neq 0\}$. Then, the family $(s_i)_{i\in I}$ is called *locally finite* if the set I_w is finite for every $w \in \Sigma^*$ (cf. [6]).

Let $h : \Delta \to K[\Sigma \cup \{\varepsilon\}]$ be a mapping. The *alphabetic morphism induced by* h is the mapping $h : \Delta^* \to K\langle\langle\Sigma^*\rangle\rangle$ such that for every $n \geq 1$, $\delta_0, \ldots, \delta_{n-1} \in \Delta$ with $h(\delta_i) = k_i.a_i$, $k_i \in K$, $a_i \in \Sigma \cup \{\varepsilon\}$, we have $h(\delta_0 \ldots \delta_{n-1}) = k_0 \ldots k_{n-1}.a_0 \ldots a_{n-1}$ and $h(\varepsilon) = 1.\varepsilon$. We should note that $h(v)$ is a monome for every $v \in \Delta^*$. If Δ is linearly ordered and $L \subseteq \Delta^*$ such that the family $(h(v))_{v\in L}$ is locally finite, then we let $h(L) = \sum_{v\in L} h(v)$ where we sum up in an ascending order according to the lexicographic order on Δ^* induced by the linear order of Δ. Now we are ready to prove our Chomsky-Schützenberger theorem.

Theorem 8 (Chomsky-Schützenberger) *Let $s \in K\langle\langle\Sigma^*\rangle\rangle$ be a context-free series. Then, there is a linearly ordered alphabet $Y \cup \overline{Y}$, a recognizable language L over $Y \cup \overline{Y}$, and an alphabetic morphism $h : Y \cup \overline{Y} \to K[\Sigma \cup \{\varepsilon\}]$ such that $s = h(D_Y \cap L)$.*

Proof: Let $\mathcal{G} = (\Sigma, N, S, R, wt)$ be a wcfg over Σ and K such that $s = \|\mathcal{G}\|$. By Proposition 7 we can assume that $\mathcal{G}$ is in Chomsky normal form. Following the proof of Theorem G.1, page 199 in [17], we define for every $r \in R$ the new letters $y_{1,r}, \overline{y_{1,r}}, y_{2,r}, \overline{y_{2,r}}$ and a new rule r' which is determined as follows:

$$r' = \begin{cases} A \to y_{1,r}B\overline{y_{1,r}}y_{2,r}C\overline{y_{2,r}} & \text{if } r = (A \to BC) \text{ with } B, C \in N \\ A \to y_{1,r}\overline{y_{1,r}}y_{2,r}\overline{y_{2,r}} & \text{if } r = (A \to a) \text{ with } a \in \Sigma \cup \{\varepsilon\}. \end{cases}$$

We let $Y = \{y_{1,r}, y_{2,r} \mid r \in R\}$ and $\overline{Y} = \{\overline{y_{1,r}}, \overline{y_{2,r}} \mid r \in R\}$ and consider the context-free grammar $\mathcal{G}' = (Y \cup \overline{Y}, N, S, R')$ with $R' = \{r' \mid r \in R\}$. We define a linear order on R' by letting $r'_1 \leq r'_2$, whenever $r_1 \leq r_2$ for every $r'_1, r'_2 \in R'$. Furthermore, we define a linear order on $Y \cup \overline{Y}$ by setting

$$y_{1,r_1} \leq \overline{y_{1,r_1}} \leq y_{2,r_1} \leq \overline{y_{2,r_1}} \leq y_{1,r_2} \leq \overline{y_{1,r_2}} \leq y_{2,r_2} \leq \overline{y_{2,r_2}}$$

whenever

$$r_1 \leq r_2$$

for every $r_1, r_2 \in R$. Obviously, $L(\mathcal{G}') \subseteq D_Y$. Moreover, by construction $\mathcal{G}'$ is unambiguous, and by the aforementioned proof in [17], we get that there exists a recognizable language L over $Y \cup \overline{Y}$ such that $L(\mathcal{G}') = D_Y \cap L$.

Next we consider the alphabetic morphism h induced by the mapping $h : Y \cup \overline{Y} \to K[\Sigma \cup \{\varepsilon\}]$, where

$$h(\widetilde{y}) = \begin{cases} 1.\varepsilon & \text{if } \widetilde{y} \in \{\overline{y}_{i,r}, y_{2,r} \mid r \in R,\ i = 1, 2\} \\ wt(r).\varepsilon & \text{if } \widetilde{y} = y_{1,r} \text{ and } r = (A \to BC) \text{ with } B, C \in N \\ wt(r).a & \text{if } \widetilde{y} = y_{1,r} \text{ and } r = (A \to a) \text{ with } a \in \Sigma \cup \{\varepsilon\} \end{cases}$$

for every $\widetilde{y} \in Y \cup \overline{Y}$.

Let now $w \in \Sigma^*$ and assume that $lfD(w) = \{d_1, \ldots, d_n\}$ is the set of all loop-free derivations of $\mathcal{G}$ for w with $d_1 <_{\text{lex}} \ldots <_{\text{lex}} d_n$. By construction of $\mathcal{G}'$, there is a derivation d'_i of $\mathcal{G}'$ which corresponds to d_i for every $1 \leq i \leq n$, and vice versa. Moreover, taking into account the linear orders of R and R', we trivially get $d'_1 <_{\text{lex}} \ldots <_{\text{lex}} d'_n$. Let us assume now that $S \stackrel{d'_i}{\Longrightarrow}_{\mathcal{G}'} u_i$ with $u_i \in (Y \cup \overline{Y})^*$ for every $1 \leq i \leq n$. We prove that $u_1 <_{\text{lex}} \ldots <_{\text{lex}} u_n$. For this, we fix an index $1 \leq i \leq n-1$ and show that $u_i <_{\text{lex}} u_{i+1}$. Let $d'_i = r'_1 \ldots r'_m r' v'$ and $d'_{i+1} = r'_1 \ldots r'_m p' z'$ where $r'_1, \ldots, r'_m \in R', r' = (A \to y_{1,r} x) \in R', p' = (A \to y_{1,p} t) \in R'$ with $r' < p'$, and $v', z' \in R'^*$. Then we get

$$S \stackrel{r'_1 \ldots r'_m}{\Longrightarrow}_{\mathcal{G}'} u' A f \stackrel{r'}{\Longrightarrow}_{\mathcal{G}'} u' y_{1,r} x f \Longrightarrow^*_{\mathcal{G}'} u_i$$

for d'_i, and

$$S \stackrel{r'_1 \ldots r'_m}{\Longrightarrow}_{\mathcal{G}'} u' A f \stackrel{p'}{\Longrightarrow}_{\mathcal{G}'} u' y_{1,p} t f \Longrightarrow^*_{\mathcal{G}'} u_{i+1}$$

for d'_{i+1}, with $u' \in (Y \cup \overline{Y})^*$.

Since $r' < p'$ we get $r < p$, and thus $y_{1,r} < y_{1,p}$ which implies that $u_i <_{\text{lex}} u_{i+1}$.

Then, for every $1 \leq i \leq n$, we have $h(u_i) = weight(d_i).w$. Moreover, by construction of $\mathcal{G}'$, it holds $w \notin \text{supp}(h(u))$ for every $u \in (D_Y \cap L) \setminus \{u_1, \ldots, u_n\}$. Therefore, we get

$$\begin{aligned} \|\mathcal{G}\|(w) &= weight(d_1) + \ldots + weight(d_n) \\ &= h(u_1)(w) + \ldots + h(u_n)(w) \\ &= \sum_{u \in D_Y \cap L} h(u)(w) \\ &= \left(\sum_{u \in D_Y \cap L} h(u) \right)(w) \\ &= h(D_Y \cap L)(w). \end{aligned}$$

If $lfD(w) = \emptyset$, then $\{u \in D_Y \cap L \mid w \in \text{supp}(h(u))\} = \emptyset$. We conclude that $s = h(D_Y \cap L)$, and our proof is completed. □

Corollary 9 *Let $s \in K\langle\langle \Sigma^* \rangle\rangle$ be a context-free series. Then, there is a linearly ordered alphabet Δ, an unambiguous context-free grammar $\mathcal{G}$ over Δ, and an alphabetic morphism $h : \Delta \to K[\Sigma \cup \{\varepsilon\}]$ such that $s = h(L(\mathcal{G}))$.*

Proof: We obtain our result by Theorem 8 by letting $\Delta = Y \cup \overline{Y}$ and $\mathcal{G} = \mathcal{G}'$. □

In [11] the authors proved a Chomsky-Schützenberger theorem for wcfg over unital valuation monoids. Since strong bimonoids are particular unital valuation monoids their Chomsky-Schützenberger theorem holds for wcfg over strong bimonoids. On the other hand, strong bimonoids are particular bimonoids, hence Chomsky-Schützenberger theorem for wcfg over strong bimonoids is implied also by our corresponding result. Nevertheless, we require that the alphabet of the involved Dyck language as well as the rules of wcfg to be linearly ordered sets. Therefore, our theory results to a weaker Chomsky-Schützenberger theorem for wcfg over strong bimonoids than that of [11].

5 Weighted Right-Linear Grammars

In this section we show that the well-known expressive equivalence of finite automata and right-linear context-free grammars holds also in the setting of bimonoids. More precisely, we consider weighted right-linear grammars over Σ and K and show that the class of their series coincides with the class of recognizable series over Σ and K. Such recognizable series were investigated in [8, 9] where they were called MK-fuzzy recognizable since the weight structure was the one used for the fuzzification of McCarthy-Kleene logic (cf. Example 1).

Definition 10 *A* weighted right-linear grammar *(wrlg for short) is a wcfg* $\mathcal{G} = (\Sigma, N, S, R, wt)$ *over* Σ *and* K *whose rules are of the form* $A \to aB$ *or* $A \to a$ *or* $A \to \varepsilon$ *with* $a \in \Sigma$ *and* $B \in N$.

By definition of the set of rules of a wrlg $\mathcal{G}$, we get that every derivation of $\mathcal{G}$ is loop-free.

We recall from [8, 9] the concept of weighted automata over bimonoids.

A weighted automaton over Σ and K is a seven-tuple

$$\mathcal{A} = (Q, I, T, F, in, wt, fin), \text{ where}$$

- Q is the finite state set which is assumed to be linearly ordered,

- $I \subseteq Q$ is the set of initial states,

- $T \subseteq Q \times \Sigma \times Q$ is the set of transitions,
- $F \subseteq Q$ is the set of final states,
- $in : I \to K$ is a mapping assigning weights to the initial states,
- $wt : T \to K$ is a mapping assigning weights to the transitions, and
- $fin : F \to K$ is a mapping assigning weights to the final states of the automaton.

Let $w = a_0 \ldots a_{n-1}$ be a word over Σ with $a_0, \ldots, a_{n-1} \in \Sigma$. A path P_w of $\mathcal{A}$ over w is a sequence of transitions $P_w := ((q_i, a_i, q_{i+1}))_{0 \le i \le n-1}$, $(q_i, a_i, q_{i+1} \in T$ for every $0 \le i \le n-1$, with $q_0 \in I$ and $q_n \in F$. The *weight of* P_w is defined by

$$weight\,(P_w) = in(q_0) \cdot \prod_{0 \le i \le n-1} wt\,(q_i, a_i, q_{i+1}) \cdot fin(q_n)$$

where in the factor $\prod_{0 \le i \le n-1} wt\,(q_i, a_i, q_{i+1})$ we multiply in an ascending order according to the usual ordering of natural numbers.

The set of paths of $\mathcal{A}$ over w can be linearly ordered in the following way. For two paths $P_w = ((q_i, a_i, q_{i+1}))_{0 \le i \le n-1}$ and $P'_w = \left((q'_i, a_i, q'_{i+1})\right)_{0 \le i \le n-1}$ we let

$$P_w \le P'_w \quad \text{iff} \quad q_0 \ldots q_{n-1} \le_{\text{lex}} q'_0 \ldots q'_{n-1}.$$

The behavior of $\mathcal{A}$ is the series $\|\mathcal{A}\| : \Sigma^* \to K$ and it is defined as follows. Let $w \in \Sigma^+$ and $\{P_{w,1}, \ldots, P_{w,m}\}$ be the set of all paths of $\mathcal{A}$ over w. Furthermore, assume that $P_{w,1} < \ldots < P_{w,m}$. Then, we set

$$\|\mathcal{A}\|(w) = weight(P_{w,1}) + \ldots + weight(P_{w,m}).$$

If there are no paths of $\mathcal{A}$ over w, then we let $\|\mathcal{A}\|(w) = 0$. If $w = \varepsilon$, then

$$\|\mathcal{A}\|(\varepsilon) = (in(q_{i_1}) \cdot fin(q_{i_1})) + \ldots + (in(q_{i_m}) \cdot fin(q_{i_m}))$$

where $I \cap F = \{q_{i_1}, \ldots, q_{i_m}\}$ and $q_{i_1} < \ldots < q_{i_m}$. If $I \cap F = \emptyset$, then we set $\|\mathcal{A}\|(\varepsilon) = 0$. A series $s : \Sigma^* \to K$ is called recognizable if there is a weighted automaton $\mathcal{A}$ over Σ and K such that $s = \|\mathcal{A}\|$.

Theorem 11 *Let Σ be a linearly ordered alphabet and K a bimonoid. Then a series $s \in K\langle\langle \Sigma^* \rangle\rangle$ is generated by a wrlg iff it is recognized by a weighted automaton over Σ and K.*

Proof: Let us assume firstly that s is recognized by a weighted automaton $\mathcal{A} = (Q, I, T, F, in, wt_{\mathcal{A}}, fin)$. We construct the wrlg $\mathcal{G} = (\Sigma, N, S, R, wt_{\mathcal{G}})$ over Σ and K in the following way. We let $N = Q \cup (I \times Q) \cup \{S\}$ where S is a new symbol, and $R = R_1 \cup R_2 \cup R_3 \cup R_4 \cup R_5 \cup R_6$ with

- $R_1 = \{S \to a(q_0, q) \mid q_0 \in I, (q_0, a, q) \in T\}$,
- $R_2 = \{(q_0, q) \to aq' \mid q_0 \in I, (q, a, q') \in T\}$,
- $R_3 = \{q \to aq' \mid (q, a, q') \in T\}$,
- $R_4 = \{q \to \varepsilon \mid q \in F\}$,
- $R_5 = \{(q_0, q) \to \varepsilon \mid q_0 \in I, q \in F\}$, and
- $R_6 = \begin{cases} \{S \to \varepsilon\} & \text{if } I \cap F \neq \emptyset \\ \emptyset & \text{otherwise.} \end{cases}$

We define a linear order on $\Sigma \cup Q$ by taking the orders of Σ and Q, and letting $\max\Sigma \leq \min Q$. This implies a linear order on the set R of rules. More precisely, we let:

$$
\begin{array}{lll}
S \to a(q_0, q) \leq S \to a'(q_0', q') & \text{iff} & aq_0q \leq_{\text{lex}} a'q_0'q' \\
 & & \text{for every } S \to a(q_0, q) \in R_1, \\
 & & \qquad\qquad S \to a'(q_0', q') \in R_1, \\
(q_0, q) \to ap \leq (q_0', q') \to a'p' & \text{iff} & aq_0qp \leq_{\text{lex}} a'q_0'q'p' \\
 & & \text{for every } (q_0, q) \to ap \in R_2, \\
 & & \qquad\qquad (q_0', q') \to a'p' \in R_2, \\
q_1 \to aq_2 \leq q_1' \to a'q_2' & \text{iff} & aq_1q_2 \leq_{\text{lex}} a'q_1'q_2' \\
 & & \text{for every } q_1 \to aq_2 \in R_3 \\
 & & \qquad\qquad q_1' \to a'q_2' \in R_3, \\
q \to \varepsilon \leq q' \to \varepsilon & \text{iff} & q \leq q' \\
 & & \text{for every } q \to \varepsilon \in R_4, \\
 & & \qquad\qquad q' \to \varepsilon \in R_4, \\
(q_0, q) \to \varepsilon \leq (q_0', q') \to \varepsilon & \text{iff} & q_0q \leq_{\text{lex}} q_0'q' \\
 & & \text{for every } (q_0, q) \to \varepsilon \in R_5, \\
 & & \qquad\qquad (q_0', q') \to \varepsilon \in R_5,
\end{array}
$$

and

$$\max R_1 \leq \min R_2,$$

$$\max R_2 \leq \min R_3,$$

$$\max R_3 \le \min R_4,$$

$$\max R_4 \le \min R_5.$$

Moreover, we set $S \to \varepsilon \le \min R_1$ if $R_6 \neq \emptyset$.

The weight mapping $wt_{\mathcal{G}}$ is defined for every rule $r \in R$ by

$$wt_{\mathcal{G}}(r) = \begin{cases} in(q_0) \cdot wt_{\mathcal{A}}(q_0, a, q) & \text{if } r = (S \to a(q_0, q)) \in R_1 \\ wt_{\mathcal{A}}(q, a, q') & \text{if } r = ((q_0, q) \to aq') \in R_2 \\ & \text{or } r = (q \to aq') \in R_3 \\ fin(q) & \text{if } r = (q \to \varepsilon) \in R_4 \\ & \text{or } r = ((q_0, q) \to \varepsilon) \in R_5 \\ \|\mathcal{A}\|(\varepsilon) & \text{if } r = (S \to \varepsilon) \end{cases}.$$

Let $w = a_0 \dots a_{n-1} \in \Sigma^+$ with $a_0, \dots, a_{n-1} \in \Sigma$ and $P_w = ((q_i, a_i, q_{i+1}))_{0 \le i \le n-1}$ be a path of $\mathcal{A}$ over w. By construction of the wrlg $\mathcal{G}$ there is a unique derivation $d = r_0 \dots r_n$ of $\mathcal{G}$ for w, which corresponds to P_w, such that
$S \overset{r_0}{\Longrightarrow} a_0(q_0, q_1) \overset{r_1}{\Longrightarrow} a_0 a_1 q_2 \overset{r_2 \dots r_{n-1}}{\Longrightarrow} a_0 a_1 \dots a_{n-1} q_n \overset{r_n}{\Longrightarrow} a_0 a_1 \dots a_{n-1} = w.$

Conversely, for every derivation d of $\mathcal{G}$ for w, there is a unique path P_w of $\mathcal{A}$ over w which corresponds to d. Furthermore, by a straightforward calculation we get

$$weight(P_w) = weight(d).$$

Hence, there is a one-to-one correspondence among the paths $P_{w,1}, \dots, P_{w,m}$ of $\mathcal{A}$ over w and the derivations $d_1, \dots, d_m$ of $\mathcal{G}$ for w. Next, we show that

$$d_1 <_{\text{lex}} \dots <_{\text{lex}} d_m$$

iff

$$P_{w,1} < \dots < P_{w,m}.$$

Indeed, let us assume that $d_j <_{\text{lex}} d_{j+1}$ for some $1 \le j \le m-1$. This implies that there is an index $0 \le k \le n-1$ such that $d_j = r_0 \dots r_{k-1} r_k \dots r_n$ and $d_{j+1} = r_0 \dots r_{k-1} r'_k \dots r'_n$ with $r_k < r'_k$. Let $r_0 = (S \to a_0(q_0, q_1))$, $r_1 = ((q_0, q_1) \to a_1 q_2)$, $r_l = (q_l \to a_l q_{l+1})$ for every $2 \le l \le n-1$, and $r_n = (q_n \to \varepsilon)$, hence $P_{w,j} = ((q_i, a_i, q_{i+1}))_{0 \le i \le n-1}$. We distinguish the following cases.

- $k = 0$. Then $r'_0 = (S \to a_0(q'_0, q'_1))$, $r'_1 = ((q'_0, q'_1) \to a_1 q'_2)$, $r'_l = (q'_l \to a_l q'_{l+1})$ for every $2 \le l \le n-1$, and $r'_n = (q'_n \to \varepsilon)$.

By definition of order on R, we get $q_0 < q'_0$, or $q_0 = q'_0$ and $q_1 < q'_1$. This implies that $P_{w,j+1} = \big((q'_i, a_i, q'_{i+1})\big)_{0 \leq i \leq n-1}$ or $P_{w,j+1} = (q_0, a_0, q'_1)\big((q'_i, a_i, q'_{i+1})\big)_{1 \leq i \leq n-1}$.

- $k = 1$. Then, by our assumption we get $r'_1 = ((q_0, q_1) \to a_1 q'_2)$ with $q_2 < q'_2$, $r'_l = (q'_l \to a_l q'_{l+1})$ for every $2 \leq l \leq n-1$, and $r'_n = (q'_n \to \varepsilon)$. Hence, $P_{w,j+1} = (q_0, a_0, q_1)(q_1, a_1, q'_2)\big((q'_i, a_i, q'_{i+1})\big)_{2 \leq i \leq n-1}$.

- $1 < k \leq n-1$. Again by our assumption we get $r'_k = (q_k \to a_k q'_{k+1})$ with $q_{k+1} < q'_{k+1}$, $r'_l = (q'_l \to a_l q'_{l+1})$ for every $k+1 \leq l \leq n-1$, and $r'_n = (q'_n \to \varepsilon)$. Thus we get $P_{w,j+1} = ((q_i, a_i, q_{i+1}))_{0 \leq i \leq k-1}$ $(q_k, a_k, q'_{k+1})\big((q'_i, a_i, q'_{i+1})\big)_{k+1 \leq i \leq n-1}$.

We conclude that $P_{w,j} <_{\text{lex}} P_{w,j+1}$ in any case. The converse implication is shown with a similar argument.

Trivially, if there is no path of $\mathcal{A}$ over w, then there is no derivation of $\mathcal{G}$ for w and vice versa. Therefore $\|\mathcal{G}\|(w) = \|\mathcal{A}\|(w)$.

Next let $w = \varepsilon$. If $I \cap F \neq \emptyset$, then $R_6 \neq \emptyset$ and $\|\mathcal{G}\|(\varepsilon) = \|\mathcal{A}\|(\varepsilon)$, whereas if $I \cap F = \emptyset$, then $R_6 = \emptyset$ and both of $\|\mathcal{G}\|(\varepsilon), \|\mathcal{A}\|(\varepsilon)$ equal to 0. We conclude that $\|\mathcal{G}\| = \|\mathcal{A}\|$.

Assume now that s is generated by a wrlg $\mathcal{G} = (\Sigma, N, S, R, wt_{\mathcal{G}})$ over Σ and K. We consider a new symbol $E \notin N$ and construct the weighted automaton $\mathcal{A} = (Q, I, T, F, in, wt_{\mathcal{A}}, fin)$ over Σ and K with

- $Q = \{(A, r) \mid A \in N \text{ and } r = (A \to u) \in R\} \cup \{E\}$,
- $I = \{(S, r) \mid r = (S \to u) \in R\}$,
- $T = \{((A, r), a, (B, p)) \mid r = (A \to aB) \in R, p = (B \to u) \in R\} \cup$ $\{((A, r), a, E) \mid r = (A \to a) \in R\}$,
- $F = \{(A, r) \mid r = (A \to \varepsilon) \in R\} \cup \{E\}$,
- $in(q) = 1$, for every $q \in I$,
- $wt_{\mathcal{A}}((A, r), a, q) = wt_{\mathcal{G}}(r)$ for every $((A, r), a, q) \in T$, and
- $fin(q) = \begin{cases} wt_{\mathcal{G}}(r) & \text{if } q = (A, r) \\ 1 & \text{if } q = E \end{cases}$, for every $q \in F$.

We define a linear order on Q by letting

$$(A, r) \leq (A', r') \quad \text{iff} \quad (r \leq r')$$

for every $(A, r), (A', r') \in Q \setminus \{E\}$ and $\max Q = E$.

Let now $w = a_0 \dots a_{n-1} \in \Sigma^+$ with $a_0, \dots, a_{n-1} \in \Sigma$ and d a derivation of $\mathcal{G}$ for w. Since $\mathcal{G}$ is wrlg, there are rules $r_i = (A_i \to a_i A_{i+1}) \in R$, $0 \le i \le n-2$, with $A_0 = S$ such that

$$S \xRightarrow{r_0} a_0 A_1 \xRightarrow{r_1} a_0 a_1 A_2 \xRightarrow{r_2} \dots \xRightarrow{r_{n-2}} a_0 a_1 \dots a_{n-2} A_{n-1}.$$

We distinguish two cases:

i) there is a rule $r_{n-1} = (A_{n-1} \to a_{n-1}) \in R$,
hence we get $a_0 a_1 \dots a_{n-2} A_{n-1} \xRightarrow{r_{n-1}} a_0 \dots a_{n-1}$, or

ii) there are rules $r_{n-1} = (A_{n-1} \to a_{n-1} A_n) \in R,\ r_n = (A_n \to \varepsilon) \in R$,
hence we get $a_0 a_1 \dots a_{n-2} A_{n-1} \xRightarrow{r_{n-1}} a_0 \dots a_{n-1} A_n \xRightarrow{r_n} a_0 \dots a_{n-1}$.

By construction of the weighted automaton $\mathcal{A}$, we get respectively the paths

i′) $((S, r_0), a_0, (A_1, r_1))((A_1, r_1), a_1, (A_2, r_2)) \dots$
$((A_{n-1}, r_{n-1}), a_{n-1}, E)$, or

ii′) $((S, r_0), a_0, (A_1, r_1))((A_1, r_1), a_1, (A_2, r_2)) \dots$
$((A_{n-1}, r_{n-1}), a_{n-1}, (A_n, r_n))$.

Furthermore, by a straightforward calculation, we get respectively

$$\begin{aligned}
\text{i}'')\ weight(d) &= wt_{\mathcal{G}}(r_0) \cdot \ldots \cdot wt_{\mathcal{G}}(r_{n-1}) \\
&= in(S, r_0) \cdot wt_{\mathcal{A}}((S, r_0), a_0, (A_1, r_1)) \cdot \ldots \cdot \\
&\qquad wt_{\mathcal{A}}((A_{n-1}, r_{n-1}), a_{n-1}, E) \cdot fin(E) \\
&= weight(P_w), \text{ or} \\
\text{ii}'')\ weight(d) &= wt_{\mathcal{G}}(r_0) \cdot \ldots \cdot wt_{\mathcal{G}}(r_{n-1}) \cdot wt_{\mathcal{G}}(r_n) \\
&= in(S, r_0) \cdot wt_{\mathcal{A}}((S, r_0), a_0, (A_1, r_1)) \cdot \ldots \cdot \\
&\qquad wt_{\mathcal{A}}((A_{n-1}, r_{n-1}), a_{n-1}, (A_n, r_n)) \cdot fin(A_n, r_n) \\
&= weight(P_w).
\end{aligned}$$

By similar arguments we can show that for every path P_w of $\mathcal{A}$ over w there is a unique derivation d in $\mathcal{G}$ for w with the same weight. Furthermore, if $d_1, \dots, d_m$ are all the derivations of $\mathcal{G}$ for w and $P_{w,1}, \dots, P_{w,m}$ are the corresponding paths of $\mathcal{A}$ over w, then by standard calculations we get

$$d_1 <_{\text{lex}} \dots <_{\text{lex}} d_m$$

iff

$$P_{w,1} < \ldots < P_{w,m}.$$

Hence, we derive $\|\mathcal{A}\|(w) = \|\mathcal{G}\|(w)$. If $D(w) = \emptyset$, then obviously there are no paths of $\mathcal{A}$ over w, thus both $\|\mathcal{A}\|(w)$ and $\|\mathcal{G}\|(w)$ equal to 0. Finally, assume that there is a rule $r = (S \to \varepsilon) \in R$. Then, by definition of I and F, we get $I \cap F = \{(S, r)\}$ hence, $\|\mathcal{A}\|(\varepsilon) = in(S, r) \cdot fin(S, r) = 1 \cdot wt_{\mathcal{G}}(r) = \|\mathcal{G}\|(\varepsilon)$. We conclude that $\|\mathcal{A}\| = \|\mathcal{G}\|$, and our proof is completed. □

Conclusion

We introduced and studied wcfg over an alphabet Σ and an arbitrary bimonoid K. For our work, we were motivated by recent needs, of weighted computational models over bimonoids, for practical applications [20, 21]. As our main results, we showed that for every wcfg we can effectively construct an equivalent one in Chomsky normal form, and proved a Chomsky-Schützenberger type result for the class of series generated by our grammars. Furthermore, we proved in our setting a well-known result relating the notions of recognizability and context-freeness, namely the class of series generated by weighted right-linear grammars coincides with the class of recognizable series over Σ and K. For this, we required the input alphabet Σ to be linearly ordered. Several problems remain open. More precisely, the closure of the class of series of wcfg with scalars from the left, closure under Cauchy product and Kleene star. Especially, the last two operations are not defined in a unique way for series over bimonoids (cf. [8, 9, 26]) due to the lack of commutativity of the operations. It is an open question also, whether the Hadamard product of a context-free series with a recognizable series over Σ and K is still a context-free series. As our future work, we state the investigation of weighted pushdown automata with weights in K, which turns to an interesting problem for practical applications.

Acknowledgement We are grateful to three anonymous referees who helped us a lot to improve the presentation of the paper.

References

[1] J.-M. Autebert, J. Berstel, L. Boasson. Context-Free Languages and Pushdown Automata. In G. Rozenberg, A. Salomaa (Eds.) *Handbook*

of Formal Languages vol. 1, 111–174, Springer, 1997. doi:10.1007/978-3-642-59136-5.

[2] V. Bhattiprolu, S. Gordon, M. Viswanathan. Extending Parikhs Theorem to Weighted and Probabilistic Context-Free Grammars. In N. Bertrand, L. Bortolussi (Eds.) Quantitative Evaluation of Systems (QEST 2017), *Lecture Notes in Computer Science* 10503, 3–19, 2017. doi:10.1007/978-3-319-66335-7_1.

[3] M. M. Bonsangue, J. Rutten, J. Winter. Defining Context-Free Power Series Coalgebraically. In D. Pattinson, L. Schrder (Eds.) Coalgebraic Methods in Computer Science (CMCS 2012), *Lecture Notes in Computer Science* 7399, 20–39, 2012. doi:10.1007/978-3-642-32784-1_2.

[4] M. Ćirić, M. Droste, J. Ignjatović, H. Vogler. Determinization of Weighted Finite Automata Over Strong Bimonoids. *Information Sciences* 180(18), 3497–3520, 2010. doi:10.1016/j.ins.2010.05.020.

[5] E. M. Clarke, O. Grumberg, S. Jha. Verifying Parameterized Networks Using Abstraction and Regular Languages. In I. Lee, S. A. Smolka (Eds.) Concurrency Theory (CONCUR 1995), *Lecture Notes in Computer Science* 962, 395–407, 1995. doi:10.1007/3-540-60218-6_30.

[6] M. Droste, W. Kuich. Semirings and Formal Power Series. Chapter 1 in [7]. doi:10.1007/978-3-642-01492-5_1.

[7] M. Droste, W. Kuich, H. Vogler (Eds.). *Handbook of Weighted Automata.* EATCS Monographs in Theoretical Computer Science, Springer-Verlag, Berlin Heidelberg, 2009. doi:10.1007/978-3-642-01492-5.

[8] M. Droste, T. Kutsia, G. Rahonis, W. Schreiner. MK-Fuzzy Automata and MSO Logics. In P. Bouyer, A. Orlandini, P. San Pietro (Eds.) Proceedings Eighth International Symposium on Games, Automata, Logics and Formal Verification (GandALF 2017), *Electronic Proceedings in Theoretical Computer Science* 256, 106–120, 2917. doi:10.4204/EPTCS.256.8.

[9] M. Droste, T. Kutsia, G. Rahonis, W. Schreiner. MK-Fuzzy Automata and MSO Logics. *Information and Computation*, to appear.

[10] M. Droste, T. Stüber, H. Vogler, Weighted Finite Automata Over Strong Bimonoids. *Information Sciences* 180(1), 156–166, 2010. doi:10.1016/j.ins.2009.09.003.

[11] M. Droste, H. Vogler. The Chomsky-Schützenberger Theorem for Quantitative Context-Free Languages. *International Journal of Foundations of Computer Science* 25(8), 955–969, 2014. doi:10.1142/S0129054114400176.

[12] S. Eilenberg. *Automata, Languages and Machines*, vol. A. Academic Press, 1974.

[13] Z. Ésik, W. Kuich. *Modern Automata Theory*. Available at: http://www.dmg.tuwien.ac.at/kuich/mat.pdf.

[14] T. Hanneforth. A Practical Algorithm for Intersecting Weighted Context-Free Arammars. In *Proceedings of 9th International Workshop on Finite State Methods and Natural Language Processing*, 57–64, 2011, Association for Computational Linguistics.

[15] G. Hughes, T. Bultan. Interface Grammars for Modular Software Model Checking. In *Proceedings of the 2007 International Symposium on Software Testing and Analysis (ISSTA 2007), ACM*, 39–49, 2017. doi:10.1145/1273463.1273471.

[16] S. Kleene. *Introduction to Metamathematics*. North-Holland, 1952.

[17] D. C. Kozen. *Automata and Computability*. Spinger-Verlag, New York, 1997. doi:10.1007/978-1-4612-1844-9.

[18] W. Kuich, A. Salomaa. *Semirings, Automata, Languages*. EATCS Monographs in Theoretical Computer Science, Springer-Verlag, Berlin Heidelberg, 1986. doi:10.1007/978-3-642-69959-7.

[19] P. Li, Y. Li, S. Geng. The Realization Problems Related to Weighted Transducers Over Strong Bimonoids. In: *Proceedings of 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 1686–1690, 2014. doi:10.1109/FUZZ-IEEE.2014.6891580.

[20] LogicGuard I. http://www.risc.jku.at/projects/LogicGuard/.

[21] LogicGuard II. http://www.risc.jku.at/projects/LogicGuard2/.

[22] J. McCarthy. A Basis for a Mathematical Theory of Computation, Computer Programming and Formal Systems. *Studies in Logic and the Foundations of Mathematics* 35, 33–70, 1963. doi:10.1016/S0049-237X(08)72018-4.

[23] P. O. Meredith, D. Jin, F. Chen, G. Roşu. Efficient Monitoring of Parametric Context-Free Patterns. *Automated Software Engineering* 17(2), 149–180, 2010. doi:10.1007/s10515-010-0063-y.

[24] J. Sakarovitch. *Elements of Automata Theory.* Cambridge University Press, 2009. doi:10.1017/CBO9781139195218.

[25] N. A. Smith, M. Johnson. Weighted and Probabilistic Context-Free Grammars are Equally Expressive. *Computational Linguistics* 33(4), 477–491, 2007. doi:10.1162/coli.2007.33.4.477.

[26] F. Torpari. *Context-Free Series with Values in Bimonoids.* Master Thesis, Aristotle University of Thessaloniki, Thessaloniki 2018. Available at: https://ikee.lib.auth.gr/record/299349/files/FaidraTorpari.pdf.

[27] C. Yang, Y. Li. Bisimulation Relations for Weighted Automata Over Valuation Monoids. In T. H. Fan, S. L. Chen, S. M. Wang, Y. M. Li (Eds.) *Proceedings of 4th International Conference on Quantitative Logic and Soft Computing (QLSC2016). Advances in Intelligent Systems and Computing* 510, 181-191, 2016. doi:10.1007/978-3-319-46206-6_19.

[28] http://www.informatik.uni-leipzig.de/wata2018/.