PAPER *Special Section on Foundations of Computer Science — Foundations of Computer Science Supporting the Information Society —*

Weighted Multiple Context-Free Grammars

Yusuke INOUE[†a)], ***Nonmember***, **Kenji HASHIMOTO**[†b)], ***Member***, ***and*** **Hiroyuki SEKI**[†c)], ***Fellow***

SUMMARY Multiple context-free grammar (MCFG) is an extension of context-free grammar (CFG), which generates tuples of words. The expressive power of MCFG is between CFG and context-sensitive grammar while MCFG inherits good properties of CFG. In this paper, we introduce weighted multiple context-free grammar (WMCFG) as a quantitative extension of MCFG. Then we investigate properties of WMCFG such as polynomial-time computability of basic problems, its closure property and expressive power.

key words: *weighted multiple context-free grammar, weight, formal power series, expressive power, closure property*

1. Introduction

Multiple context-free grammar (MCFG) [8] is an extension of context-free grammar (CFG), which generates tuples of words. The expressive power of MCFG is between CFG and context-sensitive grammar; MCFG can generate some non-context-free languages such as $\{a^n b^n c^n \mid n \geq 0\}$ and $\{a^m b^n c^m d^n \mid m, n \geq 0\}$. MCFG inherits good properties of CFG such as polynomial-time decidability of basic problems and the closure property under language operations. (See [3] for the details.) CFG is also extended to weighted context-free grammar (WCFG). WCFG has been applied to syntax analysis of natural languages and structure prediction of biological sequences. In these applications, we can find the most plausible derivation tree(s) among all derivation trees of a given word by solving the optimization problem (minimum or maximum weight problem) for a given WCFG. On the other hand, there are some structures that cannot be expressed by CFG such as discontinuous construction in natural language syntax and 'pseudoknot' structure in the secondary structure of an RNA. To deal with these problems, probability has been incorporated into MCFG in some analyses [5]–[7]. However, such stochastic extensions of MCFG have been done in rather an *ad hoc* way and properties of MCFG with general weights have not been studied.

In this paper, we propose weighted multiple context-free grammar and investigate its properties. In Sect. 2, we give a formal definition of weighted multiple context-free grammar (WMCFG). WMCFG is a quantitative extension of MCFG as well as an extension of WCFG. A WMCFG defines a formal power series, which is a function that maps a tuple of words to an element of an assumed semiring. Furthermore, we define multiple algebraic system, which is an extension of algebraic system [1], [2]. A multiple algebraic system also defines a formal power series, called a multiple algebraic power series. In Sect. 3, we first discuss a normal form of WMCFG. Then, we show the equivalence of WMCFG and multiple algebraic system, by translating a given multiple algebraic system into a WMCFG that defines the same formal power series and vice versa. In Sect. 4, we give polynomial-time algorithms for the coefficient problem and minimum-weight problem. They are function problems which are natural generalizations of the membership problem for WMCFG and emptiness problem for WMCFG over the tropical semiring, respectively. In Sect. 5, we define four operations on formal power series. These operations are natural extensions of union, intersection, concatenation and Kleene star. We discuss the closure property of the class of multiple algebraic power series under these operations. In Sect. 6, we show pumping lemmas for multiple algebraic power series, and discuss the expressive power of WMCFG.

Manuscript received April 15, 2022.
Manuscript revised August 11, 2022.
Manuscript publicized October 14, 2022.

[†]The authors are with Graduate School of Informatics, Nagoya University, Nagoya-shi, 464–8601 Japan.
a) E-mail: y-inoue@sqlab.jp
b) E-mail: k-hasimt@i.nagoya-u.ac.jp
c) E-mail: seki@sqlab.jp

DOI: 10.1587/transinf.2022FCP0006

2. Preliminaries

Let $\mathbb{N}$ be the set of all positive integers, and let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For $k \in \mathbb{N}_0$, let $[k]$ be the set $\{1, \ldots, k\}$. The cardinality of a set X is denoted by $|X|$.

A *semiring* $(\mathbb{S}, \oplus, \odot, \mathbb{0}, \mathbb{1})$ is an algebraic structure where

- $(\mathbb{S}, \oplus, \mathbb{0})$ is a commutative monoid,
- $(\mathbb{S}, \odot, \mathbb{1})$ is a monoid,
- $\odot$ distributes over $\oplus$,
- $\mathbb{0}$ is the zero element of $\odot$.

A semiring $(\mathbb{S}, \oplus, \odot, \mathbb{0}, \mathbb{1})$ is called a commutative semiring if $(\mathbb{S}, \odot, \mathbb{1})$ is commutative. We abbreviate $(\mathbb{S}, \oplus, \odot, \mathbb{0}, \mathbb{1})$ as $\mathbb{S}$. The following two are examples of semirings: *natural number semiring* $\mathbb{N}_{+,\times} = (\mathbb{N}_0, +, \times, 0, 1)$ and *Boolean semiring* $\mathbb{B} = (\{0, 1\}, \vee, \wedge, 0, 1)$.

A semiring $\mathbb{S}$ is said to be *positive* if the mapping $h_{\mathbb{S}} : \mathbb{S} \to \mathbb{B}$ defined as $h_{\mathbb{S}}(x) = 0 \Leftrightarrow x = \mathbb{0}$ is a homomorphism, i.e, $h_{\mathbb{S}}(\mathbb{0}) = 0$, $h_{\mathbb{S}}(\mathbb{1}) = 1$, $h_{\mathbb{S}}(a \oplus b) = h_{\mathbb{S}}(a) \vee h_{\mathbb{S}}(b)$ and $h_{\mathbb{S}}(a \odot b) = h_{\mathbb{S}}(a) \wedge h_{\mathbb{S}}(b)$ for all $a, b \in \mathbb{S}$. For example, $\mathbb{N}_{+,\times}$ is a positive semiring.

Let Σ be a (finite) alphabet. For a word $w \in \Sigma^*$, the length of w is denoted by $|w|$. The empty word is denoted by

ε, i.e., $|\varepsilon| = 0$.

2.1 Weighted Multiple Context-Free Grammar

Let m be a positive integer and $\mathbb{S}$ be a commutative semiring. An *m-weighted multiple context-free grammar (m-WMCFG) over* $\mathbb{S}$ is a tuple $G = (V, \Sigma, F, P, I, \mathrm{wt})$, where each component is given as follows.

- V is a finite set of nonterminals. With each nonterminal $A \in V$, a positive integer $d(A) \in [m]$ called the *dimension of* A is associated. $I \in V$ is the initial symbol and $d(I) = 1$.
- Σ is a finite set of terminals, disjoint from V.
- F is a finite set of m-mcf functions. We say that f is an m-mcf function if there exist positive integers $d_0, d_1, \dots, d_{a(f)} \in [m]$ where $a(f) \in \mathbb{N}_0$ is the number of arguments of f, and

$$f\left((x_{11}, \dots, x_{1d_1}), \cdots, (x_{a(f)1}, \dots, x_{a(f)d_{a(f)}})\right) = (\alpha_1, \dots, \alpha_{d_0}) \quad (*1)$$

  where $\alpha_1, \dots, \alpha_{d_0}$ are non-empty words over $\Sigma \cup \{x_{ij} \mid i \in [a(f)], j \in [d_i]\}$, provided that each x_{ij} appears exactly once among all of $\alpha_0, \dots, \alpha_{d_0}$. We call $(d_0, d_1, \dots, d_{a(f)})$ the *signature* of f.
- P is a set of rules of the form: $p = (A_0 \to f[A_1, \dots, A_{a(f)}])$ where $A_0 \in V, A_1, \dots, A_{a(f)} \in V \setminus \{I\}, f \in F$ and the signature of f is $(d(A_0), d(A_1), \dots, d(A_{a(f)}))$. If $a(f) = 1$ and any terminal symbol does not appear in the right-hand side of the definition (*1) of f, we call p a *unit rule*. There are no unit rules in P, except for the rules whose left-hand side is I[†]. If $a(f) = 0$, we call p a *terminating rule*. Otherwise, p is called a *nonterminating rule*.
- $\mathrm{wt} : P \to \mathbb{S}$ is a *weight function*.

The set of *derivation trees* of $A \in V$, denoted by $\mathcal{D}_G(A)$ is defined as the smallest set satisfying the following conditions.

- If $p = (A \to f[]) \in P$, then $p \in \mathcal{D}_G(A)$.
- If $p = (A \to f[A_1, \dots, A_{a(f)}]) \in P$ and $t_i \in \mathcal{D}_G(A_i)$ for each $i \in [a(f)]$, then $p(t_1, \dots, t_{a(f)}) \in \mathcal{D}_G(A)$.

We define $\mathcal{D}(G) = \mathcal{D}_G(I)$ for the initial symbol I. For $t \in \mathcal{D}_G(A)$, the tuple of words derived by t, denoted by $\mathrm{yield}(t) \in (\Sigma^*)^{d(A)}$, the height of t, denoted by $\mathrm{height}(t)$, and the weight of t, denoted by $\mathrm{wt}(t)$ are defined as follows.

- If $t = p = (A \to f[])$, then
  - $\mathrm{yield}(t) = f()$,
  - $\mathrm{height}(t) = 1$,
  - $\mathrm{wt}(t) = \mathrm{wt}(p)$.
- If $t = p(t_1, \dots, t_{a(f)})$ and $p = (A \to f[A_1, \dots, A_{a(f)}])$, then
  - $\mathrm{yield}(t) = f\left(\mathrm{yield}(t_1), \dots, \mathrm{yield}(t_{a(f)})\right)$,
  - $\mathrm{height}(t) = \max\{\mathrm{height}(t_1), \dots, \mathrm{height}(t_{a(f)})\} + 1$,
  - $\mathrm{wt}(t) = \mathrm{wt}(p) \odot \bigodot_{i \in [a(f)]} \mathrm{wt}(t_i)$.

For $t \in \mathcal{D}_G(A)$, we say that A derives $\mathrm{yield}(t)$ with weight $\mathrm{wt}(t)$. Furthermore, the weight of $w \in \Sigma^*$ is defined by

$$[\![G]\!](w) = \bigoplus_{\mathrm{yield}(t) = w} \mathrm{wt}(t)\,.$$

Note that $[\![G]\!](\varepsilon)$ is always $\mathbb{0}$ because there is no tree which derives ε. The function $[\![G]\!] : \Sigma^* \to \mathbb{S}$ is called the *formal power series* defined by m-WMCFG G.

Example 2.1. *Let* $V = \{I, A\}$, $\Sigma = \{a, b, c\}$, $F = \{f_0, f_1, f_2\}$, $P = \{p_0, p_1, p_2\}$, $\mathrm{wt}(p_0) = \mathrm{wt}(p_1) = \mathrm{wt}(p_2) = 1$ *and*

$$\begin{array}{ll} p_0 = I \to f_0[A], & f_0((x_1, x_2, x_3)) = x_1 x_2 x_3, \\ p_1 = A \to f_1[A, A], & f_1((x_1, x_2, x_3), (y_1, y_2, y_3)) \\ & \quad = (x_1 y_1, x_2 y_2, x_3 y_3), \\ p_2 = A \to f_2[], & f_2() = (a, b, c). \end{array}$$

$G_1 = (V, \Sigma, F, P, I, \mathrm{wt})$ *is a* 3*-WMCFG over* $\mathbb{N}_{+,\times}$*, and the formal power series* $[\![G_1]\!]$ *is*

$$[\![G_1]\!](w) = \begin{cases} C_{n-1} & (w = a^n b^n c^n, n \in \mathbb{N})\,, \\ 0 & (\textit{otherwise})\,, \end{cases}$$

where C_n *is Catalan number* $\frac{(2n)!}{(n+1)!n!}$.

Example 2.2. *Let* $V = \{I, B\}$, $\Sigma = \{a, b, c\}$, $F = \{g_0, g_1, g_2\}$, $P = \{p'_0, p'_1, p'_2\}$, $\mathrm{wt}(p'_0) = \mathrm{wt}(p'_1) = \mathrm{wt}(p'_2) = 1$ *and*

$$\begin{array}{ll} p'_0 = I \to g_0[B], & g_0((x_1, x_2)) = x_1 x_2, \\ p'_1 = B \to g_1[B], & g_1((x_1, x_2)) = (a x_1 b, c x_2), \\ p'_2 = B \to g_2[], & g_2() = (ab, c). \end{array}$$

$G_2 = (V, \Sigma, F, P, I, \mathrm{wt})$ *is a* 2*-WMCFG over* $\mathbb{B}$*, and the formal power series* $[\![G_2]\!]$ *is*

$$[\![G_2]\!](w) = \begin{cases} 1 & (w = a^n b^n c^n, n \in \mathbb{N})\,, \\ 0 & (\textit{otherwise})\,. \end{cases}$$

2.2 Multiple Algebraic System

In this subsection, we define *m-multiple algebraic system* as a natural extension of algebraic system [1], [2]. First, we discuss functions from $(\Sigma^*)^m$ to an element of the semiring. These are traditionally formalized as 'power series' (rather than 'functions'), in order to utilize algebraic (symbolical) manipulations such as multiplication over power series (infinite sums).

Let $\mathbb{S}$ be a commutative semiring and Σ be a finite set of terminals. An *m-formal power series* over $\mathbb{S}$ is a function $S : (\Sigma^*)^m \to \mathbb{S}$. We write $S(w)$ as (S, w) and S as

$$S = \bigoplus_{w \in (\Sigma^*)^m} (S, w) w\,.$$

[†]The restriction is along the same line as the definition of LCFRS [9].

(S, w) is called the *coefficient* of w in S. For simplicity, we omit a term $(S, w)w$ if $(S, w) = \mathbb{0}$, and if $(S, w) = \mathbb{0}$ for all $w \in (\Sigma^*)^m$, then we just write $\mathbb{0}$ to express such S. The *support* of S is defined by

$$\mathrm{supp}(S) = \{w \in (\Sigma^*)^m \mid (S, w) \neq \mathbb{0}\} .$$

For example, the support of $[\![G_1]\!]$ in Example 2.1 is equal to the support of $[\![G_2]\!]$ in Example 2.2.

Remark 2.3. *Let $G = (V, \Sigma, F, P, I, \mathrm{wt})$ be a WMCFG over a positive semiring. The support of $[\![G]\!]$ is equal to the language $L(G')$ generated by MCFG $G' = (V, \Sigma, F, P, I)$ in the standard definition.*

If $\mathrm{supp}(S)$ is finite, then S is called an m-polynomial. Let $\mathbb{S}\langle\!\langle\Sigma\rangle\!\rangle^m$ and $\mathbb{S}\langle\Sigma\rangle^m$ be the set of all m-formal power series and all m-polynomials over $\mathbb{S}$, respectively. Furthermore, we define

$$\mathbb{S}\langle\!\langle\Sigma\rangle\!\rangle^{<\omega} = \bigcup_{m\in\mathbb{N}} \mathbb{S}\langle\!\langle\Sigma\rangle\!\rangle^m , \ \mathbb{S}\langle\Sigma\rangle^{<\omega} = \bigcup_{m\in\mathbb{N}} \mathbb{S}\langle\Sigma\rangle^m .$$

Let m be a positive integer and $\mathbb{S}$ be a commutative semiring. An m-multiple algebraic system over $\mathbb{S}$ is a tuple $\mathcal{A} = (V, \Sigma, X, \mathrm{Nt}, \alpha, I)$, where each component is given as follows.

- V is a finite set of nonterminals. With each nonterminal $A \in V$, a positive integer $d(A) \in [m]$ called the *dimension of* A is associated. $I \in V$ is the initial symbol and $d(I) = 1$.
- Σ is a finite set of terminals, disjoint from V.
- X is a finite or countable set of variables. Each variable $\vec{x} \in X$ has dimension $d(\vec{x}) \in \mathbb{N}$. We write $\vec{x} \in X$ as $\vec{x} = (x_1, \ldots, x_{d(\vec{x})})$. For $X' \subseteq X$, let $\mathrm{C}(\vec{x})$ and $\mathrm{C}(X')$ be the sets $\{x_1, \ldots, x_{d(\vec{x})}\}$ and $\bigcup_{\vec{x}\in X'} \mathrm{C}(\vec{x})$, respectively. For $w \in ((\mathrm{C}(X) \cup \Sigma)^*)^k$ with some $k \in [m]$, let $X(w) = \{\vec{x} \mid$ some $x \in \mathrm{C}(\vec{x})$ appears in $w\}$.
- Nt is a mapping $X \to V \setminus \{I\}$.
- α is a system of *simultaneous equations* $\alpha : V \to \mathbb{S}\langle\mathrm{C}(X)\cup\Sigma\rangle^{<\omega}$. For each $A \in V$, $\alpha(A) \in \mathbb{S}\langle\mathrm{C}(X)\cup\Sigma\rangle^{d(A)}$. Furthermore, for every $w \in ((\mathrm{C}(X) \cup \Sigma)^*)^{d(A)}$ such that $(\alpha(A), w) \neq \mathbb{0}$,
  - if $A \neq I$, then $|X(w)| \neq 1$ or some terminal symbol appears in w,
  - ε does not appear in any component of w, and
  - each $x \in \mathrm{C}(X(w))$ appears in w exactly once.

For $A \in V$, we write $A = \alpha(A)$ and call it the *equation* of A in α.

Example 2.4. *Let $V = \{I, A\}$, $\Sigma = \{a, b\}$, $X = \{(x_1, x_2)\}$, $\mathrm{Nt}((x_1, x_2)) = A$ and*

$$\alpha = \begin{cases} I = x_1x_2 , \\ A = (ax_1, ax_2) + 2(bx_1, bx_2) + (a, a) + 2(b, b) . \end{cases}$$

$\mathcal{A} = (V, \Sigma, X, \mathrm{Nt}, \alpha, I)$ is a 2-multiple algebraic system over $\mathbb{N}_{+,\times}$. □

For an m-multiple algebraic system $\mathcal{A} = (V, \Sigma, X, \mathrm{Nt}, \alpha, I)$, let $\mathrm{Sub}(X', \Sigma)$ be the set of substitutions s whose domain is $X' \subseteq X$ such that $s(\vec{x}) \in (\Sigma^*)^{d(\vec{x})}$ for each $\vec{x} \in X'$.

For a nonterminal $A \in V$, we define the j-th *approximate solution* of A, denoted by $S_j(A)$ as

$$S_0(A) = \mathbb{0}, \text{ and}$$
$$S_j(A) = S_{j-1}(\alpha(A)) \text{ for } j > 0 ,$$

where $S_j(\alpha(A))$ is the *evaluation* of the polynomial $\alpha(A)$ by S_j, and it is defined as follows. First, we define the evaluation of $w \in ((\Sigma \cup \mathrm{C}(X))^*)^k$ with $k \in [m]$ by S_j as

$$S_j(w) = \bigoplus_{s\in\mathrm{Sub}(X(w),\Sigma)} \left(\left(\bigodot_{\vec{x}\in X(w)} (S_j(\mathrm{Nt}(\vec{x})), s(\vec{x})) \right) s(w) \right).$$

Note that $(S_j(\mathrm{Nt}(\vec{x})), s(\vec{x}))$ is the coefficient of $s(\vec{x})$ in the approximate solution $S_j(\mathrm{Nt}(\vec{x}))$. Then, we define $S_j(p)$ for a polynomial $p \in \mathbb{S}\langle\Sigma\rangle^k$ with $k \in [m]$ as

$$S_j(p) = \bigoplus_{w\in((\Sigma\cup\mathrm{C}(X))^*)^k} \left((p, w) \odot S_j(w) \right) .$$

Example 2.4 (*continued*)**.** *We compute approximate solutions of $\mathcal{A}$. By the definition,* 0*-th approximate solutions are*

$$S_0(I) = 0, \ S_0(A) = 0 .$$

Next, $S_1(I) = S_0(\alpha(I)) = S_0(x_1x_2) = 0$. This is because $S_0(x_1x_2)$ is the sum of $(S_0(A), s((x_1, x_2))) \cdot s(x_1x_2)$ for each s and the coefficients of $s((x_1, x_2))$ in $S_0(A)$ are all 0. Furthermore,

$$\begin{aligned} S_1(A) &= S_0(\alpha(A)) \\ &= S_0((ax_1, ax_2) + 2(bx_1, bx_2) + (a, a) + 2(b, b)) \\ &= 1 \cdot S_0((ax_1, ax_2)) + 2 \cdot S_0((bx_1, bx_2)) \\ &\qquad + 1 \cdot S_0((a, a)) + 2 \cdot S_0((b, b)) . \end{aligned}$$

Note that $S_0((ax_1, ax_2))$ is the sum of $(S_0(A), s(x_1, x_2)) \cdot s(ax_1, ax_2)$ for each s and the coefficients of $s((x_1, x_2))$ in $S_0(A)$ are all 0. Hence, $S_0((ax_1, ax_2)) = 0$, and in the same way, $S_0((bx_1, bx_2)) = 0$. Because $X((a, a)) = X((b, b)) = \emptyset$, $1 \cdot S_0((a, a)) = (a, a)$ and $2 \cdot S_0((b, b)) = 2(b, b)$ by the definition of $S_j(w)$. Therefore, we obtain that $S_1(A) = (a, a) + 2(b, b)$. Then, $S_2(I) = S_1(\alpha(I)) = S_1(x_1x_2)$ is the sum of $(S_1(A), s((x_1, x_2))) \cdot s(x_1x_2)$ for each s. Therefore, $S_2(I) = aa + 2bb$. Similarly, each solution can be computed as follows.

$$\begin{aligned} S_2(A) &= (a, a) + 2(b, b) + (aa, aa) \\ &\qquad + 2(ab, ab) + 2(ba, ba) + 4(bb, bb) , \\ S_3(I) &= aa + 2bb + aaaa + 2abab + 2baba + 4bbbb , \\ S_3(A) &= (a, a) + 2(b, b) + (aa, aa) + 2(ab, ab) \\ &\quad + 2(ba, ba) + 4(bb, bb) + (aaa, aaa) \\ &\quad + 2(aab, aab) + 2(aba, aba) + 2(baa, baa) \\ &\quad + 4(abb, abb) + 4(bab, bab) + 4(bba, bba) \\ &\quad + 8(bbb, bbb), \ \cdots \end{aligned}$$
□

The limit of approximate solution $\lim_{j\to\infty} S_j(A)$ is called the *strong solution* of A. For an m-multiple algebraic system $\mathcal{A} = (V, \Sigma, X, \mathrm{Nt}, \alpha, I)$, the strong solution of I, denoted by $[\![\mathcal{A}]\!]$ is called the *formal power series* defined by $\mathcal{A}$. If there exists an m-multiple algebraic system $\mathcal{A}$ such that $S = [\![\mathcal{A}]\!]$, then S is called an *m-multiple algebraic power series*.

Example 2.4 (*continued*)**.** *The strong solutions of $\mathcal{A}$ are*

$$S_k(I) = aa + 2bb + aaaa + 2abab + 2baba + 4bbbb + aaaaaa + 2aabaab + 2abaaba + 2baabaa + \cdots$$

$$S_k(A) = (a,a) + 2(b,b) + (aa,aa) + 2(ab,ab) + 2(ba,ba) + 4(bb,bb) + (aaa,aaa) + 2(aab,aab) + 2(aba,aba) + 2(baa,baa) + \cdots .$$

Therefore, the formal power series defined by $\mathcal{A}$ is

$$[\![\mathcal{A}]\!] = aa + 2bb + aaaa + 2abab + 2baba + 4bbbb + aaaaaa + 2aabaab + 2abaaba + 2baabaa + \cdots .$$

Furthermore, $\mathrm{supp}([\![\mathcal{A}]\!]) = \{ww \mid w \in \Sigma^*\}$ *and the coefficient of each word ww is 2^n where n is the number of occurrences of letter b in w.*

3. Basic Properties

3.1 Normal Forms

We have assumed the following restrictions in the definition of WMCFG.

- There are no unit rules, except for the rules whose left-hand side is the initial symbol.
- For any $f \in F$, ε does not appear in the function value of f. That is, there are no *ε-rules*.
- Each component of each argument of $f \in F$ is used ‘exactly once’, not ‘at most once’. This restriction is called information lossless condition in [8] and non-erasing condition in [9].

It is shown in [8] that these restrictions do not affect the expressive power of (unweighted) MCFG, except for the derivation of ε. However, we do not know whether these restrictions weaken the expressive power of WMCFG. For weighted models, it is not easy to remove these restrictions. The difficulty lies in the fact that the weight of w in a WMCFG G is the sum of the weights of all derivation trees of w and unit rules and ε-rules may admit infinite derivation trees without deriving any terminals. In fact, similar restrictions are assumed in the discussion for WCFG in [2].

On the other hand, there are some restrictions that do not affect the expressive power of WMCFG as stated in the next theorem. We regard WMCFG satisfying these restrictions as a normal form and use it for later discussion.

Theorem 3.1. *For a given m-WMCFG $G = (V, \Sigma, F, P, I, \mathrm{wt})$, we can construct an m-WMCFG $G' = (V', \Sigma, F', P', I', \mathrm{wt}')$ such that $[\![G']\!] = [\![G]\!]$ and the following (A1), (A2) and (A3) hold.*

*(A1) For any nonterminating rule $A \to f[A_1, \ldots, A_{a(f)}] \in P'$, any terminal in Σ does not appear in the right-hand side (i.e., $(\alpha_1, \ldots, \alpha_{d_0})$ in (*1)) of the definition of f.*

(A2) For any terminating rule $A \to f[] \in P'$, $d(A) = 1$ and the function value of f is a terminal symbol, i.e., $f() = a$ for some $a \in \Sigma$.

(A3) For all $p \in P'$, p is not a unit rule, including rules whose left-hand side is I'.

Proof. For a given WMCFG $G = (V, \Sigma, F, P, I, \mathrm{wt})$, we can construct $G' = (V', \Sigma, F', P', I', \mathrm{wt}')$ which satisfies the conditions as follows. We set $V' = V$, $F' = F$, $P' = P$, $I' = I$, and $\mathrm{wt}' = \mathrm{wt}$ as initial values.

(A1) Do the following until this condition is satisfied. Let $p = (A \to f[A_1, \ldots, A_{a(f)}]) \in P$ be a nonterminating rule such that a terminal symbol b appears in the function value of the definition (*1) of f.

- Introduce a new nonterminal B to V', a new function $f_b() = b$ to F', and a new rule $p_b = (B \to f_b[])$ with $\mathrm{wt}'(p_b) = \mathbb{1}$ to P', and
- replace p by $p' = (A \to f'[A_1, \ldots, A_{a(f)}, B]) \in P'$ with $\mathrm{wt}'(p') = \mathrm{wt}'(p)$ where f' is obtained from f by adding x_b as the last argument of f' and replacing b by x_b in the function value of the definition (*1) of f.

(A2) For each terminating rule $p = (A \to f_c[]) \in P'$ where $f_c() = (\alpha_1, \ldots, \alpha_{d(A)})$,

- If $d(A) \geq 2$, replace p by $p' = (A \to \mathrm{id}[A_1, \ldots, A_{d(A)}])$ with new nonterminals $A_1, \ldots, A_{d(A)}$ where $d(A_1), \ldots, d(A_{d(A)}) = 1$, $\mathrm{id}(x_1, \ldots, x_{d(A)}) = (x_1, \ldots, x_{d(A)})$ and $\mathrm{wt}'(p') = \mathrm{wt}'(p)$. Next, for each $i \in [d(A)]$, do the following steps (*2).
  - If $|\alpha_i| = 1$, introduce a new rule $q_i = A_i \to f_i[]$ where $f_i() = \alpha_i$ and $\mathrm{wt}(q_i) = \mathbb{1}$.
  - If $|\alpha_i| \geq 2$, introduce new rules $q_i = (A_i \to f_i[B_1 \ldots, B_{|\alpha_i|}])$ and $r_j = (B_j \to g_j[])$ with new nonterminals $B, B_1, \ldots, B_{|\alpha_i|}$ for each $j \in [|\alpha_i|]$, where $f_i(x_1, \ldots, x_{|\alpha_i|}) = x_1 \cdots x_{|\alpha_i|}$, $g_j() = a_j$ for $\alpha_i = a_1 \cdots a_{|\alpha_i|}$ and $\mathrm{wt}'(q_i) = \mathrm{wt}'(r_j) = \mathbb{1}$.
- If $d(A) = 1$, regarding A as A_i and α_1 as α_i, do the same as (*2).

(A3) By the definition of WMCFG, we only need to consider rules whose left-hand side is I'. For each unit rule $p = (I' \to f[A]) \in P'$, let q be a rule whose left-hand side is A. For each $q = (A \to g[A_1, \ldots, A_{a(g)}]) \in P'$, introduce a new rule $p' = (I' \to f'[A_1, \ldots, A_{a(g)}])$ with $\mathrm{wt}'(p') = \mathrm{wt}'(p) \odot \mathrm{wt}'(q)$ where $f'(\vec{x}_1, \ldots, \vec{x}_{a(g)}) = f(g(\vec{x}_1, \ldots, \vec{x}_{a(g)}))$. Note that p' is not a unit rule because q is not a unit rule. Finally, the condition is satisfied by deleting p. □

3.2 Equivalence with Multiple Algebraic System

As stated in Theorem 3.2 and Theorem 3.4, WMCFG and multiple algebraic system have the same expressive power, that is, a formal power series S is defined by some m-WMCFG iff S is an m-multiple algebraic power series.

Theorem 3.2. *For a given m-WMCFG* $G=(V,\Sigma,F,P,I,\mathrm{wt})$*, we can construct an m-multiple algebraic system* $\mathcal{A}$ *such that* $[\![\mathcal{A}]\!]=[\![G]\!]$.

Proof. Assume that G satisfies the conditions (A1) and (A2) in Theorem 3.1 without loss of generality. For each $A\in V$, define polynomial $\alpha(A)\in S\langle\Sigma\rangle^{d(A)}$ as follows. If there exists a rule $p=(A\to f[A_1,\dots,A_{a(f)}])\in P$ such that $f((x_{11},\dots,x_{1d(A_1)}),\dots,(x_{a(f)1},\dots,x_{a(f)d(A_{a(f)})}))=(w_1,\dots,w_{d(A)})$, then let $(\alpha(A),(w_1,\dots,w_{d(A)}))=\mathrm{wt}(p)$, $(x_{i1},\dots,x_{id(A_i)})\in X$ and $\mathrm{Nt}((x_{i1},\dots,x_{id(A_i)}))=A_i$ for each $i\in[a(f)]$. Otherwise, let $(\alpha(A),(w_1,\dots,w_{d(A)}))=\mathbb{0}$. We construct a multiple algebraic system $\mathcal{A}=(V,\Sigma,X,\mathrm{Nt},\alpha,I)$ for X, Nt and α obtained by the above construction. Then, for every $i\in\mathbb{N}_0$ and $A\in N$, we claim that the following equation holds.

$$S_i(A)=\bigoplus_{t\in\mathcal{D}_G(A),\mathrm{height}(t)\le i}(\mathrm{wt}(t)\,\mathrm{yield}(t))\qquad\text{(A-1)}$$

We prove the Eq. (A-1) holds by induction on i. For $i=0$, (A-1) holds because $S_0(A)=\mathbb{0}$ by the definition and there are no tree t such that $\mathrm{height}(t)=0$. For a given $k\in\mathbb{N}_0$, we assume that

$$S_k(A)=\bigoplus_{t\in\mathcal{D}_G(A),\mathrm{height}(t)\le k}(\mathrm{wt}(t)\,\mathrm{yield}(t))\ ,$$

that is,

$$(S_k(A),w)=\bigoplus_{\substack{t\in\mathcal{D}_G(A),\mathrm{height}(t)\le k\\ \mathrm{yield}(t)=w}}\mathrm{wt}(t)$$

holds as the inductive hypothesis. Then,

$$\begin{aligned}
S_{k+1}(A)&=S_k(\alpha(A))\\
&=\bigoplus_{w\in((\Sigma\cup C(X))^*)^{d(A)}}(\alpha(A),w)\odot S_k(w)\\
&=\bigoplus_{w\in((\Sigma\cup C(X))^*)^{d(A)}}\Big((\alpha(A),w)\odot\\
&\qquad\bigoplus_{s\in\mathrm{Sub}(X(w),\Sigma)}\Big(\bigodot_{\vec{x}\in X(w)}(S_k(\mathrm{Nt}(\vec{x})),s(\vec{x}))\ s(w)\Big)\Big)\\
&=\bigoplus_{\substack{w\in((\Sigma\cup C(X))^*)^{d(A)}\\(\alpha(A),w)\neq\mathbb{0}}}\Big(\bigoplus_{s\in\mathrm{Sub}(X(w),\Sigma)}\\
&\qquad\Big((\alpha(A),w)\odot\bigodot_{\vec{x}\in X(w)}(S_k(\mathrm{Nt}(\vec{x})),s(\vec{x}))\ s(w)\Big)\Big)
\end{aligned}$$

where every word w such that $(\alpha(A),w)\neq\mathbb{0}$ corresponds with every rule whose left-hand side is A, and every substitution s corresponds with ever derivation tree of A in G. Furthermore, by the definition of $\alpha(A)$ and Nt,

$$\begin{aligned}
S_{k+1}(A)&\\
&=\bigoplus_{p=A\to f[A_1,\dots,A_{a(f)}]\in P}\Big(\bigoplus_{t=p(t_1,\dots,t_{a(f)})\in\mathcal{D}_G(A)}\\
&\qquad\Big(\mathrm{wt}(p)\bigodot_{j\in[a(f)]}(S_k(A_j),\mathrm{yield}(t_j))\ \mathrm{yield}(t)\Big)\Big)\ ,
\end{aligned}$$

and by the inductive hypothesis,

$$\begin{aligned}
&=\bigoplus_{p=A\to f[A_1,\dots,A_{a(f)}]\in P}\Big(\bigoplus_{t=p(t_1,\dots,t_{a(f)})\in\mathcal{D}_G(A)}\\
&\qquad\Big(\mathrm{wt}(p)\bigodot_{j\in[a(f)]}\Big(\bigoplus_{\substack{t_j\in\mathcal{D}_G(A_j)\\ \mathrm{height}(t_j)\le k}}\mathrm{wt}(t_j)\Big)\ \mathrm{yield}(t_j)\Big)\Big)\\
&=\bigoplus_{p=A\to f[A_1,\dots,A_{a(f)}]\in P}\Big(\bigoplus_{\substack{t=p(t_1,\dots,t_{a(f)})\in\mathcal{D}_G(A)\\ \mathrm{height}(t_1),\dots,\mathrm{height}(t_{a(f)})\le k}}\\
&\qquad\Big(\mathrm{wt}(p)\bigodot_{j\in[a(f)]}(\mathrm{wt}(t_j)\,\mathrm{yield}(t))\Big)\Big)\\
&=\bigoplus_{p=A\to f[A_1,\dots,A_{a(f)}]\in P}\\
&\qquad\Big(\bigoplus_{\substack{t=p(t_1,\dots,t_{a(f)})\in\mathcal{D}_G(A)\\ \mathrm{height}(t)\le k+1}}(\mathrm{wt}(t)\,\mathrm{yield}(t))\Big)\\
&=\bigoplus_{t\in\mathcal{D}_G(A),\mathrm{height}(t)\le k+1}(\mathrm{wt}(t)\,\mathrm{yield}(t))\ .
\end{aligned}$$

Therefore, (A-1) holds when $i=k+1$. Because the coefficient of w in the strong solution of $\mathcal{A}$ is equal to $[\![G]\!](w)$, $\mathcal{A}$ satisfies $[\![\mathcal{A}]\!]=[\![G]\!]$. □

Example 3.3. *For 3-WMCFG* $G=(V,\Sigma,F,P,I,\mathrm{wt})$ *in Example 2.1, we can construct 3-multiple algebraic system* $\mathcal{A}$ *such that* $[\![\mathcal{A}]\!]=[\![G]\!]$ *as follows:* $\mathcal{A}=(V,\Sigma,X,\mathrm{Nt},\alpha,I)$ *where* $X=\{(x_1,x_2,x_3),(y_1,y_2,y_3)\}$, $\mathrm{Nt}((x_1,x_2,x_3))=\mathrm{Nt}((y_1,y_2,y_3))=A$,

$$\alpha=\begin{cases}I=x_1x_2x_3\\ A=(x_1y_1,x_2y_2,x_3y_3)+(a,b,c)\end{cases}\ .$$

Theorem 3.4. *For a given m-multiple algebraic system* $\mathcal{A}=(V,\Sigma,X,\mathrm{Nt},\alpha,I)$*, we can construct an m-WMCFG G such that* $[\![G]\!]=[\![\mathcal{A}]\!]$.

Proof. Let P and F be the sets of all p and f defined by the following steps: For each $A\in V$ and for each $(\alpha(A),w)$ which is not $\mathbb{0}$, if $X_1,\dots,X_n\in X(w)$, then let $p=(A\to f[\mathrm{Nt}(X_1),\dots,\mathrm{Nt}(X_n)])$, $f(X_1,\dots,X_n)=w$ and $\mathrm{wt}(p)=(\alpha(A),w)$. We construct a WMCFG $G=(V,\Sigma,F,P,I,\mathrm{wt})$. Then for every $i\in\mathbb{N}_0$ and $A\in N$, the following holds:

$$\bigoplus_{t\in\mathcal{D}_G(A),\mathrm{height}(t)\le i}(\mathrm{wt}(t)\,\mathrm{yield}(t))=S_i(A)\ .$$

This can be proved in the same way as Theorem 3.2. Therefore, G satisfies $[\![G]\!]=[\![\mathcal{A}]\!]$. □

Example 3.5. *For 2-multiple algebraic system* $\mathcal{A}=(V,\Sigma,X,\mathrm{Nt},\alpha,I)$ *in Example 2.4, we can construct 2-WMCFG G such that* $[\![G]\!]=[\![\mathcal{A}]\!]$ *as follows:* $G=$

$(V, \Sigma, F, P, I, \mathrm{wt})$ *where* $P = \{p_0, p_1, p_2, p_3, p_4\}, F = \{f_0, f_1, f_2, f_3, f_4\}$ *and*

$$\begin{array}{lll} p_0 = I \to f_0[A] & f_0((x_1, x_2)) = x_1 x_2 & \mathrm{wt}(p_0) = 1 \\ p_1 = A \to f_1[A] & f_1((x_1, x_2)) = (a x_1, a x_2) & \mathrm{wt}(p_1) = 1 \\ p_2 = A \to f_2[A] & f_2((x_1, x_2)) = (b x_1, b x_2) & \mathrm{wt}(p_2) = 2 \\ p_3 = A \to f_3[] & f_3() = (a, a) & \mathrm{wt}(p_3) = 1 \\ p_4 = A \to f_4[] & f_4() = (b, b) & \mathrm{wt}(p_4) = 2\,. \end{array}$$

By Theorems 3.2 and 3.4, we often refer to an m-multiple algebraic power series S to mean that S is a formal power series defined by an m-WMCFG. We define the class *m-WMCF* and *WMCF* of formal power series as

$$m\text{-}\mathit{WMCF} = \{S \mid S \text{ is an } m\text{-multiple algebraic power series.}\}$$

$$\mathit{WMCF} = \bigcup_{m \in \mathbb{N}} m\text{-}\mathit{WMCF}\,.$$

4. Function Problems

In this section, we assume that a single operation of a semiring can be performed in constant time.

4.1 Coefficient Problem

Let $G = (V, \Sigma, F, P, I, \mathrm{wt})$ be a WMCFG. The *coefficient problem* is the function problem of computing $[\![G]\!](w)$ for an input $w \in \Sigma^*$. For a WMCFG over $\mathbb{N}_{+,\times}$ or $\mathbb{B}$, or more generally, over positive semirings, the coefficient problem is a generalization of the membership problem. This is because for a positive semiring, $[\![G]\!](w) \neq \mathbb{0}$ iff there exists at least one derivation tree such that $\mathrm{yield}(t) = w$ and $\mathrm{wt}(t) \neq \mathbb{0}$.

Theorem 4.1. *The coefficient problem for a given word* $w \in \Sigma^*$ *can be computed in* $O(|w|^{e+1})$ *time where* $e =$

$$\max\{\sum_{i \in \{0\} \cup [k]} d(A_i) \mid p = A_0 \to f[A_1, \ldots, A_k] \in P\}\,.$$

Proof. Assume that G satisfies the conditions (A1) and (A2) in Theorem 3.1 without loss of generality. We will search all trees $t \in \mathcal{D}(G)$ such that $\mathrm{yield}(t) = w$, using dynamic programming.

Let $w = a_1 a_2 \cdots a_n$ where $a_1, \ldots, a_n \in \Sigma$. For a tuple $u = (l_1, r_1, \ldots, l_k, r_k) \in [n]^{2k}$ for some $k \in [n]$ such that $l_i \leq r_i$ for all $i \in [k]$ and either $r_i < l_j$ or $r_j < l_i$ for all $i \neq j \in [k]$, we call u a *position vector* of w. For a position vector $u \in [n]^{2k}$, let $w[u] = (a_{l_1} \cdots a_{r_1}, \ldots, a_{l_k} \cdots a_{r_k}) \in (\Sigma^*)^k$ and $\mathrm{len}(u) = \sum_{i \in [k]} (r_i - l_i + 1)$. We say that (A, u, z) is a *derivation triple* if a nonterminal $A \in V$ derives $w[u] \in (\Sigma^*)^k$ with weight $z \in \mathbb{S}$. We will search derivation triples (A, u, z) for each $A \in V$ in the increasing order of $\mathrm{len}(u)$, and add them to the set Z as follows. Let $Z = \emptyset$ as the initial value. For each triple (A, u, z) found, update Z to $Z \cup \{(A, u, z)\}$ if $(A, u, z') \notin Z$ for any z', and update z' to $z' \oplus z$ if $(A, u, z') \in Z$ for some z'. We denote this procedure as $\mathrm{add}(Z, (A, u, z))$. Finally, we can obtain the value $[\![G]\!](w)$ as z' such that $(I, (1, n), z') \in Z$.

Algorithm 1 Coefficient

input : $w = a_1 \cdots a_n \in \Sigma^*$ (where $a_1, \ldots, a_n \in \Sigma$)

$Z := \emptyset$
for each terminating rule $p = (A \to f[]) \in P$ where $f() = a$ **do**
  for each $i \in [n]$ **do**
    if $a_i = a$ **then**
      $\mathrm{add}(Z, (A, (i, i), \mathrm{wt}(p))$
    end if
  end for
end for
for len := 1 to n **do**
  for each nonterminating rule $p = (A \to f[A_1, \ldots, A_k])$ where $f((x_{11}, \ldots, x_{1d(A_1)}), \cdots, (x_{k1}, \ldots, x_{kd(A_k)})) = (y_{11} \cdots y_{1t_1}, y_{21} \cdots y_{2t_2}, \ldots, y_{d1} \cdots y_{dt_d})$ **do**
    for each $(l_{11}, l_{12}, \ldots, l_{1t_1}, r_1, l_{21}, l_{22}, \ldots, l_{2t_2}, r_2, \cdots, l_{d1}, l_{d2}, \ldots, l_{dt_d}, r_d) \in [n]^{e_p}$ where $e_p = \sum_{j \in [d]} (t_j + 1)$ $(= d + \sum_{i \in [k]} d(A_i))$ such that $\sum_{j \in [d]} (r_j - l_{j1}) = \mathrm{len}$ **do**
      let θ be the mapping as: for each $j \in [d], h \in [t_j]$,
      $$\theta(y_{jh}) = \begin{cases} a_{l_{jh}} \cdots a_{r_j} & (h = t_j) \\ a_{l_{jh}} \cdots a_{l_{j(h+1)}-1} & (h < t_j) \end{cases}$$
      if $(A_i, u_i, z_i) \in Z$ where $w[u_i] = (\theta(x_{i1}), \cdots, \theta(x_{id(A_i)}))$ for each $i \in [k]$ **then**
        $\mathrm{add}(Z, (A, u', z'))$ where $z' = \mathrm{wt}(p) \odot \bigodot_{i \in [k]} z_i$ and $u' = (l_{11}, r_1, l_{21}, r_2, \ldots, l_{d1}, r_d)$ is a position vector.
      end if
    end for
  end for
end for
return z' such that $(I, (1, n), z') \in Z$.

(See Algorithm 1.)

First, we search derivation triples (A, u, z) such that $\mathrm{len}(u) = 1$. Because G satisfies the conditions (A1) and (A2), we only need to examine rules of the form $p = (A \to f[])$ where $f() = a$ for some $a \in \Sigma$. For such a rule p, nonterminal A derives a with weight $\mathrm{wt}(p)$. Hence, we add derivation triples $(A, (i, i), \mathrm{wt}(p))$ such that $a_i = a$ to Z. Next, we search derivation triples (A, u, z) such that $\mathrm{len}(u) \geq 2$ by examining nonterminating rules to combine derivation triples already found. Let $p = (A \to f[A_1, \cdots, A_k])$ be a nonterminating rule such that $f((x_{11}, \ldots, x_{1d(A_1)}), \cdots, (x_{k1}, \ldots, x_{kd(A_k)})) = (y_{11} \cdots y_{1t_1}, \ldots, y_{d1} \cdots y_{dt_d})$ with $k = a(f), d = d(A)$ and $y_{i'j'} \in \{x_{ij} \mid i \in [k], j \in [d(A_i)]\}$ for each $i' \in [d], j' \in [t_d]$. For each position vector $u' \in [n]^{2d}$, we check whether or not there is a derivation tree $t = p(t_1, \ldots, t_k) \in \mathcal{D}_G(A)$ for some $t_i \in \mathcal{D}_G(A_i)$ for $i \in [k]$ such that $\mathrm{yield}(t) = w[u']$ as follows. Let $w[u'] = (\alpha_1, \ldots, \alpha_d)$. For each $j \in [d]$, we partition α_j into non-empty subwords $\alpha_{j1}, \ldots, \alpha_{jt_j}$ such that $\alpha_{jh} \neq \varepsilon$ for $h \in [t_j]$ and $\alpha_j = \alpha_{j1} \cdots \alpha_{jt_j}$. Let θ be the mapping from $\{x_{i1}, \ldots, x_{id(A_i)} \mid i \in [k]\}$ to Σ^+ such that $\theta(y_{jh}) = \alpha_{jh}$ for each $j \in [d]$ and $h \in [t_j]$. By the definition of m-mcf function, θ is well-defined, i.e., for each $x_{i\ell}$, there is exactly one $j \in [d]$ and $h \in [t_j]$ such that $y_{jh} = x_{i\ell}$ and vice versa. Then, check whether $(A_i, u_i, z_i) \in Z$ for all $i \in [k]$ where u_i satisfies $w[u_i] = (\theta(x_{i1}), \ldots, \theta(x_{id(A_i)}))$. We can find all derivation trees of w, doing this steps until $\mathrm{len}(u') = n$, i.e., $w[u'] = w$. Note that the number of different u' associated with possible partitions of α_j $(j \in [d])$ is at most n^{e_p} where $e_p = d + \sum_{i \in [k]} d(A_i)$. This is because $w[u']$ and its partition is determined by the left end position of each y_{jh} and the right end position of each y_{jt_j} where $j \in [d], h \in [t_j]$. Therefore, the coefficient problem can be computed in $O(n \times n^e)$ time where $e = \max\{e_p \mid p \in P\}$. □

Example 4.2. *Let* $G = (\{I, A, T_a, T_b, T_c\}, \{a, b, c\}, F, P, I, \mathrm{wt})$ *be WMCFG over* $\mathbb{N}_{+,\times}$, *where F and P consist of the following:*

$$\begin{aligned}
p_0 &= I \to f_0[A, A], & f_0((x_1, x_2, x_3), (y_1, y_2, y_3)) &= x_1y_1x_2y_2x_3y_3,\\
p_1 &= A \to f_1[A, A], & f_1((x_1, x_2, x_3), (y_1, y_2, y_3)) &= (x_1y_1, x_2y_2, x_3y_3),\\
p_2 &= A \to f_2[T_a, T_b, T_c], & f_2(x_1, x_2, x_3) &= (x_1, x_2, x_3),\\
p_a &= T_a \to f_a[] & f_a() &= a,\\
p_b &= T_b \to f_b[] & f_b() &= b,\\
p_c &= T_c \to f_c[] & f_c() &= c,
\end{aligned}$$

and $\mathrm{wt}(p) = 1$ *for each* $p \in P$. *Note that G is equivalent to* G_1 *in Example 2.1. Now, we compute* $[\![G]\!](w)$ *for input* $w = a^3b^3c^3$ *using Algorithm 1.*

First, by rules p_a, p_b *and* p_c, *we add derivation triples* $(T_a, (i_a, i_a), 1)$, $(T_b, (i_b, i_b), 1)$ *and* $(T_c, (i_c, i_c), 1)$ *for each* $i_a \in \{1, 2, 3\}$, $i_b \in \{4, 5, 6\}$, $i_c \in \{7, 8, 9\}$.

Next, we add triples $(A, (i_a, i_a, i_b, i_b, i_c, i_c), 1)$ *for each* $i_a \in \{1, 2, 3\}$, $i_b \in \{4, 5, 6\}$, $i_c \in \{7, 8, 9\}$ *by rule* p_2. *For example,* $(A, (1, 1, 4, 4, 8, 8), 1)$ *is added to Z and this means that nonterminal A can derive* $w[(1, 1, 4, 4, 8, 8)] = (a, b, c)$ *with weight* 1. *On the other hand,* $(A, (1, 1, 7, 7, 8, 8), 1)$ *is not added to Z. This is because* $(T_b, (7, 7), 1)$ *is not in Z and the condition in line 13 cannot be satisfied, and in fact A cannot derive* $w[(1, 1, 7, 7, 8, 8)] = (a, c, c)$.

Then, we add triple $(A, (1, 2, 4, 5, 7, 8), 1)$ *to Z combining* $(A, (1, 1, 4, 4, 7, 7), 1)$ *and* $(A, (2, 2, 5, 5, 8, 8), 1)$ *by* p_1. *This triple means that A can derive* $w[(1, 2, 4, 5, 7, 8)] = (aa, bb, cc)$ *with weight* 1. *Similarly, we add* $(A, (i_{aa}, i_{aa} + 1, i_{bb}, i_{bb} + 1, i_{cc}, i_{cc} + 1), 1)$ *for each* $i_{aa} \in \{1, 2\}$, $i_{bb} \in \{4, 5\}$ *and* $i_{cc} \in \{7, 8\}$. *Finally, we add* $(I, (1, 9), 1)$ *twice by* p_0. *One is combined from* $(A, (1, 2, 4, 5, 7, 8), 1)$ *and* $(A, (3, 3, 6, 6, 9, 9), 1)$, *and the other is combined from* $(A, (1, 1, 4, 4, 7, 7), 1)$ *and* $(A, (2, 3, 5, 6, 8, 9), 1)$. *Because no other triples of the form* $(I, (1, 9), z)$ *are added, we can conclude that the weight of w is* 2.

4.2 Minimum-Weight Problem

The *emptiness problem* is the decision problem of determining whether or not $\mathrm{supp}([\![G]\!]) = \emptyset$ for a WMCFG $G = (V, \Sigma, F, P, I, \mathrm{wt})$ such that $\mathrm{wt}(p) \neq \mathbb{0}$ for each $p \in P$. If the semiring is positive, the emptiness problem for a given WMCFG G can be solved in polynomial time of the description length of G as the standard emptiness problem for MCFG (see Remark 2.3). In particular, a stronger problem can be solved in the same complexity for WMCFG over the positive semiring $\mathbb{N}_{\min,+} = (\mathbb{N}_0 \cup \{\infty\}, \min, +, \infty, 0)$, which is known as the tropical semiring.

The *minimum-weight problem* is the function problem of computing the minimum weight of all words, i.e., computing $\min\{[\![G]\!](w) \mid w \in \Sigma^*\}$ for a given WMCFG $G = (V, \Sigma, F, P, I, \mathrm{wt})$ over the tropical semiring. Note that the minimum-weight problem is a generalization of the emptiness problem because the minimum weight is ∞ iff $\mathrm{supp}([\![G]\!]) = \emptyset$.

Theorem 4.3. *The minimum-weight problem for a given WMCFG* $G = (V, \Sigma, F, P, I, \mathrm{wt})$ *over the tropical semiring can be computed in* $O(n^2)$ *time where*

$$n = |V| + \sum_{p=(A \to f[A_1, \ldots, A_{a(f)}]) \in P} (1 + a(f)) .$$

Proof. For each $A \in V$, let $\min A = \min\{\mathrm{wt}(t) \mid t \in \mathcal{D}_G(A)\}$, and call it the *minimum weight* of A. Note that the minimum weight of all words is equal to $\min I$. We can simultaneously compute $\min A$ for each $A \in V$ by computing the greatest fixpoint of the following equations iteratively.

For $k \in \mathbb{N}_0$, $A \in V$ and $p \in P$, let us define $m(A, k)$ and $m(p, k)$ as follows.

$$\begin{aligned}
&m(A, 0) = \infty,\\
&m(A, k) = \min\{m(p, k) \mid p \in P\\
&\qquad\qquad \text{and the left-hand side of } p \text{ is } A\} \quad (k \in \mathbb{N})
\end{aligned}$$

where

$$m(p, k) = \mathrm{wt}(p) + \sum_{i \in [a(f)]} m(A_i, k - 1)$$

for a rule $p = (A \to f[A_1, \ldots, A_{a(f)}]) \in P$. (See Algorithm 2.) These values can be computed in $O(n)$ time for each $k \in \mathbb{N}$. By the definition, $m(A, k') \leq m(A, k)$ and $m(p, k') \leq m(p, k)$ hold for all $A \in V$, $p \in P$ and $k' > k$. At the end of the k-th iteration, let $z_k = \min\{m(A, k) \mid A \in \Delta V_k\}$ and $Z_k = \{A \in \Delta V_k \mid z_k = m(A, k)\}$ where $\Delta V_1 = V$, $\Delta V_{k+1} = \Delta V_k \setminus Z_k$ for each $k \in \mathbb{N}$. Because $z_{k'} \geq z_k$ and $m(A, k') = m(A, k) = \min A$ hold for all

Algorithm 2 Minimum-weight

```
input : WMCFG G = (V, Σ, F, P, I, wt) over N_min,+
for each A ∈ V do
    m(A) := ∞
end for
ΔV := V
while I ∈ ΔV do
    Z := ∅, z := ∞
    for each p = (A → f[A_1, ..., A_d]) ∈ P such that A ∈ ΔV do
        m(p) := wt(p) + Σ_{i∈[d]} m(A_i)
        if m(p) < m(A) then
            m(A) := m(p)
        end if
        if m(p) < z then
            z := m(p), Z := {A}
        else if m(p) = z then
            Z := Z ∪ {A}
        end if
    end for
    ΔV := ΔV \ Z
end while
return m(I)
```

$A \in Z_k$ and $k' > k$, we can determine $\min A$ as z_k for each $A \in Z_k$. Therefore, $k \leq |V|$ when the fixpoint is obtained, and the minimum-weight problem can be computed in $O(n^2)$ time. □

Note that this algorithm cannot be applied to computing the minimum weight of a given WMCFG over other semirings in general. In particular, the algorithm does not compute the minimum weight for WMCFG over the natural number semiring $\mathbb{N}_{+,\times}$. In fact, $\min\{[\![G]\!](w) \mid w \in \Sigma^*\}$ for WMCFG G over $\mathbb{N}_{+,\times}$ is not computable for the following reason. Because for $\mathbb{N}_{+,\times}$, $\min\{[\![G]\!](w) \mid w \in \Sigma^*\} \neq 0$ iff each $w \in \Sigma^*$ has at least one parse tree, we could decide whether $L(G) = \Sigma^*$ or not for an (M)CFG G if we could compute the minimum weight. However, this contradicts the fact that the problem of deciding $L(G) = \Sigma^*$ for an (M)CFG G is undecidable.

5. Closure Properties

In this section, we define some operations on formal power series and we discuss closure properties of *WMCF* under the operations. Each of these operations corresponds to union, intersection, concatenation and Kleene star, respectively, and we show that multiple algebraic power series inherit the closure properties from CFL and MCFL in this respect.

For formal power series $S_1, S_2 : \Sigma^* \to \mathbb{S}$, we define the *sum* $S_1 \oplus S_2$, *Cauchy product* $S_1 \odot S_2$, and the *convolution* $S_1 S_2$ as

$$S_1 \oplus S_2 = \bigoplus_{w \in \Sigma^*} (S_1(w) \oplus S_2(w))w\,,$$

$$S_1 \odot S_2 = \bigoplus_{w \in \Sigma^*} (S_1(w) \odot S_2(w))w\,, \text{ and}$$

$$S_1 S_2 = \bigoplus_{w \in \Sigma^*} \bigoplus_{w = w_1 w_2} (S_1(w_1) \odot S_2(w_2))w\,,$$

respectively. For a formal power series $S : \Sigma^* \to \mathbb{S}$, we define the *quasi-inverse* S^+ as

$$S^+ = \bigoplus_{i \in \mathbb{N}} S^i\,,$$

where $S^1 = S$ and $S^{i+1} = S S^i$ for $i \in \mathbb{N}$. Note that S^* can naturally be defined as $S^* = S^+ \oplus \mathbb{1}\varepsilon$.

Theorem 5.1. *For each $m \in \mathbb{N}$, the class m-WMCF is closed under sum and convolution.*

Proof. Let $G_1 = (V_1, \Sigma, F_1, P_1, I_1, \mathrm{wt}_1)$ and $G_2 = (V_2, \Sigma, F_2, P_2, I_2, \mathrm{wt}_2)$ be m-WMCFGs. We can assume that G_1 and G_2 satisfy the condition (A3) in Theorem 3.1, and V_1 is disjoint from V_2, without loss of generality.

Let $G = (V_1 \cup V_2 \cup \{I'\}, \Sigma, F_1 \cup F_2 \cup \{\mathrm{id}\}, P_1 \cup P_2 \cup \{p_1, p_2\}, I', \mathrm{wt}_1 \cup \mathrm{wt}_2 \cup \{p_1 \mapsto \mathbb{1}, p_2 \mapsto \mathbb{1}\}^{\dagger})$ be the m-WMCFG where $p_1 = (I' \to \mathrm{id}[I_1]), p_2 = (I' \to \mathrm{id}[I_2]), d(I') = 1$ and id is defined as $\mathrm{id}(x) = x$. Then, G defines the sum of $[\![G_1]\!]$ and $[\![G_2]\!]$.

Let $G = (V_1 \cup V_2 \cup \{I'\}, \Sigma, F_1 \cup F_2 \cup \{\mathrm{conc}\}, P_1 \cup P_2 \cup \{p'\}, I', \mathrm{wt}_1 \cup \mathrm{wt}_2 \cup \{p' \mapsto \mathbb{1}\})$ be the m-WMCFG where $p' = (I' \to \mathrm{conc}[I_1, I_2]), d(I') = 1$ and conc is defined as $\mathrm{conc}(x, y) = xy$. Then, G defines the convolution of $[\![G_1]\!]$ and $[\![G_2]\!]$. □

Theorem 5.2. *For each $m \in \mathbb{N}$, the class m-WMCF is closed under quasi-inverse.*

Proof. Let $G_1 = (V_1, \Sigma, F_1, P_1, I_1, \mathrm{wt}_1)$ be an m-WMCFG. We can assume that G_1 satisfies the condition (A3) in Theorem 3.1, without loss of generality. First, we introduce new nonterminals $I', I'' \notin V_1$ and new rules $p_1 = (I' \to \mathrm{id}[I''])$ and $p_+ = (I'' \to \mathrm{conc}[I_1, I''])$ with $d(I') = d(I'') = 1$. Next, for each rule $q = (I_1 \to f[A_1, \ldots, A_{a(f)}]) \in P_1$, we introduce new rules $p_q = (I'' \to f[A_1, \ldots, A_{a(f)}])$. Then, $G = (V_1 \cup \{I', I''\}, \Sigma, F_1 \cup \{\mathrm{id}, \mathrm{conc}\}, P_1 \cup \{p_1, p_+\} \cup \{p_q \mid q = (I_1 \to \cdots) \in P_1\}, I', \mathrm{wt}_1 \cup \{p_1 \mapsto \mathbb{1}, p_+ \mapsto \mathbb{1}\} \cup \{p_q \mapsto \mathrm{wt}_1(q) \mid q = (I_1 \to \cdots) \in P_1\})$ defines the quasi-inverse of $[\![G_1]\!]$. □

Theorem 5.3. *There exist WMCFGs G_1 and G_2 such that $[\![G_1]\!] \odot [\![G_2]\!]$ can not be constructible from G_1 and G_2. That is, the class WMCF is not effectively closed under Cauchy product.*

Proof. We suppose that for any WMCFGs G_1 and G_2, we can construct G such that $[\![G]\!] = [\![G_1]\!] \odot [\![G_2]\!]$. For given two CFGs G_1, G_2, we can consider G_1 and G_2 as 1-WMCFGs over $\mathbb{N}_{\min,+}$ such that the weight of every rule is 0. Then, we construct a WMCFG G that defines $[\![G_1]\!] \odot [\![G_2]\!]$. Because $\mathrm{supp}([\![G]\!]) = L(G_1) \cap L(G_2)$, we can decide the emptiness of $L(G_1) \cap L(G_2)$ from Theorem 4.3. This contradicts the fact that the emptiness problem for the intersection of two CFLs is undecidable. □

6. Pumping Lemma for Multiple Algebraic Power Series

Finally, we give a pumping lemma for WMCFG. The grammatical iteration property of WMCFG has no difference from that of MCFG, but we can show more detailed expressive power focusing on the weights.

Let $G = (V, \Sigma, F, P, I, \mathrm{wt})$ be a WMCFG. For nonterminals $A, B \in V$, the set of *contexts* $C_G(A, B)$ is defined as the smallest set such that if $t' \in \mathcal{D}_G(B)$ is a subtree of $t \in \mathcal{D}_G(A)$, then $C[\square] \in C_G(A, B)$ where $C[\square]$ is obtained by replacing t' by $\square$. For a tree $t' \in \mathcal{D}_G(B)$ and a context $C[\square] \in C_G(A, B)$, the tree obtained by substituting t' for $\square$ is denoted by $C[t']$.

---

† $\mathrm{wt}_1 \cup \mathrm{wt}_2 \cup \{p_1 \mapsto \mathbb{1}, p_2 \mapsto \mathbb{1}\}$ denotes the function $\mathrm{wt}' : P_1 \cup P_2 \cup \{p_1, p_2\} \to \mathbb{S}$ defined as:

$$\mathrm{wt}'(p) = \begin{cases} \mathrm{wt}_1(p) & (p \in P_1), \\ \mathrm{wt}_2(p) & (p \in P_2), \\ \mathbb{1} & (p = p_1, p_2). \end{cases}$$

In particular, if $A = B$, the tree $\underbrace{C[C[\cdots C[}_{i} t'] \cdots]]$ is denoted by $C^i[t']$.

Lemma 6.1. *Let $G = (V, \Sigma, F, P, I, \mathrm{wt})$ be an m-WMCFG over a positive semiring $\mathbb{S}$. If $\mathrm{supp}([\![G]\!])$ is infinite, then there exists a derivation tree $t \in \mathcal{D}(G)$ such that*

- $\mathrm{wt}(t) \neq \mathbb{0}$,
- *there exist $A \in V, C_1[\square] \in C_G(I, A), C_2[\square] \in C_G(A, A)$ and $t' \in \mathcal{D}_G(A)$ such that $t = C_1[C_2[t']]$,*
- *there exist words $u_1, u_2, \ldots, u_{2m+1} \in \Sigma^*$ and $v_1, v_2, \ldots, v_{2m} \in \Sigma^*$ such that $\mathrm{yield}(C_1[C_2^i[t']]) = u_1 v_1{}^i u_2 v_2{}^i u_3 \cdots u_{2m} v_{2m}{}^i u_{2m+1}$ for all $i \in \mathbb{N}_0$, and*
- $\sum_{j=1}^{2m} |v_j| > 0$.

Proof. Assume that G satisfies the conditions (A1) and (A2) in Theorem 3.1 without loss of generality. Then, the lemma can be proved in the same way as a pumping lemma for MCFL (Lemma 3.2 in [8]). □

Lemma 6.1 is essentially the same as the pumping lemma for MCFL in [8], and it is known that a stronger pumping lemma as the one for CFL does not always hold for MCFL. (See [4] for detail.)

Next, using the above discussion, we show pumping lemmas for WMCFG over $\mathbb{N}_{+,\times}$ and $\mathbb{N}_{\min,+}$.

Theorem 6.2. *Let S be an m-multiple algebraic power series over $\mathbb{N}_{+,\times}$. If $\mathrm{supp}(S)$ is infinite, then there exist some words $u_1, u_2, \ldots, u_{2m+1} \in \Sigma^*$, $v_1, v_2, \ldots, v_{2m} \in \Sigma^*$ and constants $c, d \in \mathbb{N}$ such that*

(1) $\sum_{j=1}^{2m} |v_j| > 0$, *and*
(2) $S(w_i) \geq c^i \cdot d$ *for all* $i \in \mathbb{N}_0$
where $w_i = u_1 v_1{}^i u_2 v_2{}^i u_3 \cdots u_{2m} v_{2m}{}^i u_{2m+1}$.

Proof. Let $G = (V, \Sigma, F, P, I, \mathrm{wt})$ be an m-WMCFG over $\mathbb{N}_{+,\times}$ that defines S. By applying Lemma 6.1, we obtain a tree $t = C_1[C_2[t']] \in \mathcal{D}(G)$ such that $\mathrm{wt}(t) > 0$ and words $w_i = \mathrm{yield}(C_1[C_2^i[t']])$ that can be decomposed as required in the claim. We show that the words w_i satisfy the condition for the weight. Let $c_1 = \mathrm{wt}(t)/\mathrm{wt}([C_2[t']])$ and $c_2 = \mathrm{wt}([C_2[t']])/\mathrm{wt}(t')$. Note that c_1 and c_2 are positive integers because the weight of a tree must be the multiples of the weights of their proper subtrees. Then, $\mathrm{wt}(C_1[C_2^i[t']])$ can be represented as $c_2^i \cdot c_1 \cdot \mathrm{wt}(t') > 0$. Therefore, $S(w_i) = \sum_{\mathrm{yield}(T)=w_i} \mathrm{wt}(T) \geq \mathrm{wt}(C_1[C_2^i[t']]) = c_2^i \cdot c_1 \cdot \mathrm{wt}(t') > 0$ for all $i \in \mathbb{N}_0$, and (2) holds by letting $c = c_2$, $d = c_1 \cdot \mathrm{wt}(t')$. □

Theorem 6.3. *Let S be an m-multiple algebraic power series over $\mathbb{N}_{\min,+}$. If $\mathrm{supp}(S)$ is infinite, then there exist some words $u_1, u_2, \ldots, u_{2m+1} \in \Sigma^*$, $v_1, v_2, \ldots, v_{2m} \in \Sigma^*$ and constants $c, d \in \mathbb{N}_0$ such that*

(1) $\sum_{j=1}^{2m} |v_j| > 0$, *and*
(2) $S(w_i) \leq c \cdot i + d$ *for all* $i \in \mathbb{N}_0$
where $w_i = u_1 v_1{}^i u_2 v_2{}^i u_3 \cdots u_{2m} v_{2m}{}^i u_{2m+1}$.

Proof. Let $G = (V, \Sigma, F, P, I, \mathrm{wt})$ be an m-WMCFG over $\mathbb{N}_{\min,+}$ that defines S. By applying Lemma 6.1, we obtain a tree $t = C_1[C_2[t']] \in \mathcal{D}(G)$ such that $\mathrm{wt}(t) \neq \infty$ and words $w_i = \mathrm{yield}(C_1[C_2^i[t']])$ that can be decomposed as required in the claim. We show that the words w_i satisfy the condition for the weight. Let $c_1 = \mathrm{wt}(t) - \mathrm{wt}([C_2[t']])$ and $c_2 = \mathrm{wt}([C_2[t']]) - \mathrm{wt}(t')$. Then, $\mathrm{wt}(C_1[C_2^i[t']])$ can be represented as $(c_2 \cdot i) + c_1 + \mathrm{wt}(t')$. Therefore, $S(w_i) = \min_{\mathrm{yield}(T)=w_i}\{\mathrm{wt}(T)\} \leq \mathrm{wt}(C_1[C_2^i[t']]) = (c_2 \cdot i) + c_1 + \mathrm{wt}(t')$ for all $i \in \mathbb{N}_0$, and (2) holds by letting $c = c_2$, $d = c_1 + \mathrm{wt}(t')$. □

Example 6.4. *Let $\Sigma = \{a_1, a_2, \cdots, a_{2m}\}$ be an alphabet. Formal power series*

$$S(w) = \begin{cases} 2^n & (w = a_1^n a_2^n \cdots a_{2m}^n, n \in \mathbb{N}) \\ \infty & (\textit{otherwise}) \end{cases}$$

is not an m-multiple algebraic power series over $\mathbb{N}_{\min,+}$ for any $m \in \mathbb{N}$.

Proof. Assume that S is an m-multiple algebraic power series over $\mathbb{N}_{\min,+}$. By Theorem 6.3, there exist $u_1, \ldots, u_{2m+1}, v_1, \ldots, v_{2m} \in \Sigma^*$, $c, d \in \mathbb{N}_0$ such that $\sum_{j=1}^{2m} |v_j| > 0$ and $S(w_i) \leq c \cdot i + d$ for all $i \in \mathbb{N}_0$ where $w_i = u_1 v_1^i u_2 v_2^i u_3 \cdots u_{2m} v_{2m}^i u_{2m+1}$. By the definition of S, for each $i \in \mathbb{N}_0$ there is $k_i \in \mathbb{N}_0$ such that $w_i = a_1^{k_i} a_2^{k_i} \cdots a_{2m}^{k_i}$. Because $\sum_{j=1}^{2m} |v_j| > 0$ and $w_i = u_1 v_1^i u_2 v_2^i u_3 \cdots u_{2m} v_{2m}^i u_{2m+1} = a_1^{k_i} a_2^{k_i} \cdots a_{2m}^{k_i}$ for all $i \in \mathbb{N}_0$, there exists $k' \in \mathbb{N}$ such that $v_j = a_j^{k'}$ for all $j \in [2m]$. Then, we have $k_i = k_0 + k' \cdot i$ and hence $S(w_i) = 2^{k_0 + k' \cdot i}$. However, it contradicts the fact that $S(w_i) \leq c \cdot i + d$ for all $i \in \mathbb{N}_0$. Therefore, S is not an m-multiple algebraic power series over $\mathbb{N}_{\min,+}$. □

7. Conclusion

In this paper, we defined WMCFG and multiple algebraic system. We proved that the class of formal power series defined by WMCFG coincides with the class of multiple algebraic power series. We also showed properties of WMCFG such as polynomial-time computability of basic function problems, closure properties and the expressive power.

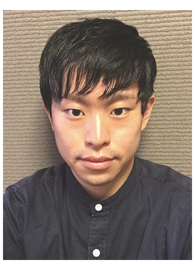
As mentioned in Sect. 3, it is not easy to remove some restrictions such as no unit rules and no ε-rules. It is an open problem whether the restrictions are removable or not, without changing the expressive power. In the discussion of the emptiness problem and pumping lemmas, we assumed that a semiring is positive. Generalizing these observations from positive semirings to general semirings is also left as future work.

Acknowledgments

The authors would like to thank the reviewers for their helpful comments. This work was supported by JSPS KAKENHI Grant Numbers 19H04083 and 22H03568.

References

[1] N. Chomsky and M.P. Schützenberger, The Algebraic Theory of Context-free Languages, Studies in Logic and the Foundations of Mathematics, vol.26, pp.118–161, 1959.
[2] M. Droste, W. Kuich, and H. Vogler, Handbook of Weighted Automata, Springer Science & Business Media, 2009.
[3] L. Kallmeyer, Parsing Beyond Context-Free Grammars, Chapters 6 and 7, pp.109–156, Springer, 2010.
[4] M. Kanazawa, G.M. Kobele, J. Michaelis, S. Salvati, and R. Yoshinaka, The Failure of the Strong Pumping Lemma for Multiple Context-free Languages, Theory of Computing Systems, vol.55, no.1, pp.250–278, 2014.
[5] Y. Kato, T. Akutsu, and H. Seki, A Grammatical Approach to RNA-RNA Interaction Prediction, Pattern Recognition, vol.42, no.4, pp.531–538, 2009.
[6] E. Rivas and S.R. Eddy, A Dynamic Programming Algorithm for RNA Structure Prediction Including Pseudoknots, Journal of Molecular Biology, vol.285, no.5, pp.2053–2068, 1999.
[7] E. Rivas and S.R. Eddy, The language of RNA: A Formal Grammar That Includes Pseudoknots, Bioinformatics, vol.16, no.4, pp.334–340, 2000.
[8] H. Seki, T. Matsumura, M. Fujii, and T. Kasami, On Multiple Context-Free Grammars, Theoretical Computer Science, vol.88, no.2, pp.192–229, 1991.
[9] D.J. Weir, Characterizing Mildly Context-Sensitive Grammar Formalisms, Ph. D. dissertation, University of Pennsylvania, 1988.



Yusuke Inoue received his Bachelor's degree from Nagoya University in 2021. Currently, he is in the Master's program, Graduate School of Informatics, Nagoya University. His research interests include formal language theory.

Kenji Hashimoto received the Ph.D. degree in Information and Computer Sciences from Osaka University in 2009. From 2009 to 2013, he was an Assistant Professor of Nara Institute of Science and Technology. Currently, he is an Assistant Professor of Graduate School of Informatics, Nagoya University. His research interests include formal language, database theory, and information security.

Hiroyuki Seki received his Ph.D. degree from Osaka University in 1987. He was an Assistant Professor, and later, an Associate Professor in Osaka University from 1987 to 1994. In 1994, he joined Nara Institute of Science and Technology, where he was a Professor during 1996 to 2013. Currently, he is a Professor in Nagoya University. His current research interests include formal language theory and formal approach to software development.