Variants of weighted sticker systems with different weighting spaces

W.H. Fong[a,*], **Y.S. Gan**[b], **N.H. Sarmin**[a], **S. Turaev**[c]

[a] Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia
81310 UTM Johor Bahru, Johor, Malaysia

[b] Xiamen University Malaysia, Jalan Sunsuria, Bandar Sunsuria 43900 Sepang, Selangor Dahrul Ehsan, Malaysia

[c] Department of Computer Science, Kulliyyah of Information and Communication Technology, International Islamic University of Malaysia 50728 Kuala Lumpur, Malaysia

*Corresponding author, e-mail: fwh@utm.my

ABSTRACT: Sticker systems have been introduced as a type of DNA computing model by using the recombination behaviour of DNA molecules. The systems use the Watson-Crick complementary principle of DNA molecules to perform the computation in generating new strings. By associating some weights to the axioms and dominoes of the sticker system, a new version of restricted sticker system, namely, a weighted sticker system has been introduced. In this paper, we study the generative power of weighted sticker systems by considering the languages generated by weighted sticker systems with different weighting spaces and cut points. The relations of the languages generated by weighted sticker systems in the Chomsky hierarchy are also investigated.

KEYWORDS: DNA computing, threshold language, computational power

MSC2010: 68Q15 68Q70 97P20 97M60 93A30

INTRODUCTION

With the modern technology, there is a new development in the field of computer science and bio-molecular sciences. DNA computing thus acts as a bridge between computer science and bio-molecular sciences. In the development of DNA computing, a theoretical model of DNA computing has been proposed, namely, the *splicing system*[1]. Splicing system uses the recombination of DNA molecules (sequences), which are double-stranded structures composed of four nucleotides to perform the computation in splicing systems. According to Watson-Crick complementarity, the nucleotide adenine (A) always bonds with thymine (T) and guanine (G) always bonds with cytosine (C). This feature of Watson-Crick complementarity relation is significant in the development of DNA computing. Using the complementarity relation, it is possible to check the information encoded on a single strand of DNA molecules and to make far-reaching conclusions from another complemented single strand. This is because the information on the other strand can be decoded according to the complementary principle.

Another feature of DNA computing is the *massive parallelism* of DNA strands, which allows the construction of many copies of DNA strands and operations on the encoded information simultaneously. There are many computationally intractable problems such as Hamiltonian path problem[2], the satisfiability problem for arbitrary contact networks[3] and the satisfiability problem for Boolean circuits[4], which are solved by DNA based computers using these two fundamental features of DNA computing.

In 1998, another type of DNA based generating device has been introduced, namely, the sticker system. Kari[5] introduces the sticker system by using the sticker operation which has been used by Adleman[2] in his experiment to solve the Hamiltonian path problem in 1994. In sticker system, the sticker operations are used on DNA molecules: from a given set of incomplete DNA sequences, a complete double-stranded sequence is formed (a DNA language). The axioms and strings generated by a sticker system are considered as encoded models of single and double-stranded DNA molecules. In

sticker systems, the initial sequence starts from the axioms A and then prolongs from the left to the right by using pairs of dominoes in $D(u, v)$ according to the sticker operations μ. This prolongation is continued until there is no blank symbol left to obtain a complete double-stranded sequence[6], denoted by $WK_\rho(V)$. The set of $WK_\rho(V)$ is also denoted by

$$\begin{bmatrix} x \\ y \end{bmatrix}_\rho$$

where

$$\begin{bmatrix} x \\ y \end{bmatrix}_\rho = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \mid a, b \in V, \quad (a, b) \in \rho \right\}.$$

We should notice that by using the form of

$$\begin{pmatrix} x \\ y \end{pmatrix},$$

it is just the pair of (x, y) and there is no relation between x and y; while the form of

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

represents a molecule. However, we shall write

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

$\in WK_\rho(V)$ instead of

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

in order to specify the complementarity relation for the precise bonding between the corresponding symbols from the alphabet V in the upper strand and lower strand of the molecule. Furthermore, the set of incomplete molecules is denoted by $W_\rho(V)$, such that

$$W_\rho(V) = L_\rho(V) \cup R_\rho(V) \cup LR_\rho(V),$$

where

$$L_\rho(V) = \left(\begin{pmatrix} \lambda \\ V^* \end{pmatrix} \cup \begin{pmatrix} V^* \\ \lambda \end{pmatrix} \right) \begin{bmatrix} V \\ V \end{bmatrix}_\rho^*,$$

$$R_\rho(V) = \begin{bmatrix} V \\ V \end{bmatrix}_\rho^* \left(\begin{pmatrix} \lambda \\ V^* \end{pmatrix} \cup \begin{pmatrix} V^* \\ \lambda \end{pmatrix} \right),$$

$$LR_\rho(V) = L_\rho(V) \cup R_\rho(V).$$

Furthermore, any element of $W_\rho(V)$ is called a well-started double-stranded sequence if there is at least a position

$$\begin{bmatrix} a \\ b \end{bmatrix}_\rho^*$$

where $a \neq \lambda$ and $b \neq \lambda$.

For two symbols $x, y \in LR_\rho(V)$, we have $x \Rightarrow y$ if and only if $y = \mu(u, \mu(x, v))$ for some $(u, v) \in D$, where $\mu(a, b)$ is the sticking operation of dominoes a and b. Since the prolongation to the right is independent with the one to the left, thus $\mu(u, \mu(x, v)) = \mu(\mu(u, x), v)$. When $x_1 \in A$ is used as a starting symbol and x_k is in $WK_\rho(V)$, a sequence $x_1 \Rightarrow x_2 \Rightarrow \cdots \Rightarrow x_k$ is obtained. This sequence is called the a computation in γ with length $k - 1$. If there is no sticky end or no blank symbol present in the last symbol, then the computation is considered as complete. A complete computation will produce a complete string w such that $w \in WK_\rho(V)$. Hence the language of such string is the language generated by γ, or called the sticker language.

Due to the limitation of the generative power in sticker systems, several restrictions in sticker systems have been introduced[7,8], which show that the use of some algebraic structures and probabilities can increase the generative power of sticker systems. Some new variants of sticker systems have been introduced, called the weighted sticker systems[9]. The new variants of sticker systems associate weights to the finite pairs of axioms $LR_\rho(V)$ and dominoes $W_\rho(V) \times W_\rho(V)$. The weight $\omega(z)$ of the strand z produced by strands x and y of $W_\rho(V)$ is calculated from the weights $\omega(x)$ and $\omega(y)$ according to the operation $\odot$ defined as $\omega(z) = \omega(x) \odot \omega(y)$. Hence several types of threshold languages generated by weighted sticker system have been considered with different weighting spaces and cutpoints. The selection of weighting spaces and cutpoints has been shown to affect the generative power of sticker systems.

On the other hand, weighted grammars and automata have been widely investigated in formal language theory since they were introduced in different forms[10–15]. The study of weighted grammars and automata shows that weights can increase the generative power of usual grammars and automata; and just as well construct more accurate models for stochastic phenomena and processes. Hence the use of weights in sticker systems makes possible the development on the application of sticker systems with more efficient parsing and tagging algorithms in processing natural and programming languages.

PRELIMINARIES

In this section, we recall some prerequisites by giving some basic terms, notation and formal definitions on the theories of formal languages and sticker systems. The reader can refer to Refs. 6, 16–19 for further information.

We use the following general notation throughout this paper. The symbols $+$ and $\times$ denote the usual addition and multiplication operations, respectively, while the symbols $\oplus$ and $\otimes$ denote the componentwise addition and multiplication operations, respectively. The sets of integers and positive rational numbers are denoted by $\mathbb{Z}$ and $\mathbb{Q}^+$, while $\mathbb{Z}^n$ denotes the n-dimensional vector space over integers. The set of matrices with integer entities is denoted by $\mathbb{M}$. The null matrix, i.e., the matrix all of whose components are zero, is denoted by $\mathbf{O}$.

The families of recursively enumerable, context-sensitive, context-free, linear, regular and finite languages are denoted by **RE**, **CS**, **CF**, **LIN**, **REG** and **FIN**, respectively. For these language families, the next strict inclusion named *Chomsky hierarchy*, holds:

Theorem 1 (Ref. 16)

$$\mathbf{FIN} \subset \mathbf{REG} \subset \mathbf{LIN} \subset \mathbf{CF} \subset \mathbf{CS} \subset \mathbf{RE}.$$

Next, we list some basic definitions of sticker systems (following Ref. 6) which are needed in the next section.

Definition 1 A sticker system is a construct of 4-tuple

$$\gamma = (V, \rho, A, D),$$

where V is an alphabet, ρ is the symmetric relation in V, A is a finite subset of axioms $LR_\rho(V)$, and D is a finite set of pairs $W_\rho(V) \times W_\rho(V)$. Then the language that is generated by a sticker system γ is defined as

$$L(\gamma) = \left\{ w \in \binom{V}{V}^* \;\middle|\; x \Rightarrow^* w, \quad x \in A \right\}.$$

There are five variants of families of sticker languages [10], namely, one-sided, regular, simple, simple one-sided and simple regular sticker languages as denoted by OSL, RSL, SSL, SOSL, and SRSL, respectively. For an arbitrary sticker system, we denote the language generated as USL. Weighted sticker systems are introduced in Ref. 9, where some weights are considered into the classical model of the sticker system. We state the formal definition of a weighted sticker system as follows.

Definition 2 A weighted sticker system is a 7-tuple

$$\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot)$$

where V is an alphabet, $\rho \subseteq V \times V$ is the symmetric relation, A_ω is a finite set of axioms A where $A \subseteq LR_\rho(V) \times M$, D_ω is a finite subset of $(W_\omega(V) \times W_\omega(V)) \times M$, ω is a weighting function from $LR_\omega(V) \cup (W_\omega(V) \times W_\omega(V))$ to a weighting space M, and $\odot$ is the operation over the weights $\omega(x), x \in LR_\omega(V) \cup (W_\omega(V) \times W_\omega(V))$.

For the weighting space M, different sets of algebraic structures such as integers, rational numbers, real numbers, Cartesian products of the set of numbers and set of matrices with integer entries can be considered. Then the operations over weights are defined with respect to the chosen weighing space.

Next, weighted sticker operation and the language generated by a weighted sticker system are defined.

Definition 3 Given $(x, \omega(x)), (y, \omega(y)) \in A_\omega$ and $(u, \omega(u)), (v, \omega(v)) \in D_\omega$,

$$[x, \omega(x)] \Rightarrow^* [y, \omega(y)]$$

if and only if
(i) $[y, \omega(y)] = \mu([u, \omega(u)], \mu([x, \omega(x)], [v, \omega(v)]))$ and $\omega(y) = \omega(x) \odot \omega(v)$;
(ii) $[y, \omega(y)] = \mu(\mu[x, \omega(x))], [u, \omega(u)]), [v, \omega(v)]$ and $\omega(y) = \omega(x) \odot \omega(u)$.
The language generated by the weighted sticker system is defined as

$$\omega\mathrm{SL} = \{ y \in WK_\rho(V) \mid [x, \omega(x)] \Rightarrow^* [y, \omega(y)]$$
$$\text{for } [x, \omega(x)] \in A_\omega \}.$$

To increase the generative power of sticker systems, we consider the languages generated by weighted splicing systems with some subsets of the weighting space called the thresholds (cut-points). We consider three types of threshold languages as stated in the next definition.

Definition 4 Let $\omega\mathrm{SL}(\gamma)$ be the language generated by a weighted sticker system $\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot)$. A threshold weighted sticker system language with respect to a threshold (cut-point) $\tau \in M$ is a subset of $\omega\mathrm{SL}(\gamma)$ defined by $\omega\mathrm{SL}(\gamma, \star\tau) = \{ y \in WK_\rho(V) \mid [x, \omega(x)] \Rightarrow^* [y, \omega(y)]$ for $[x, \omega(x)] \in A$ and $\omega(y) \star \tau \}$, where $\star \in \{=, \langle, \rangle\}$ is called the mode of $\omega\mathrm{SL}(\gamma, \star\tau)$.

Furthermore, when we consider some weights to the variants of sticker systems, we have some new variants of sticker systems such as weighted one-sided, weighted regular, weighted simple, weighted simple one-sided and weighted simple regular sticker systems. We state the formal definitions for these variants of weighted sticker systems[9] as follows.

Definition 5 A weighted sticker system $\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot)$ is said to be one-sided if for each pair $((u, \omega(u)), (v, \omega(v))) \in D_\omega$, either $(u, \omega(u)) \Rightarrow (\lambda, e)$ or $(v, \omega(v)) \Rightarrow (\lambda, e)$ with e the identity of the weighting space.

Definition 6 A weighted sticker system $\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot)$ is said to be regular if for each pair $((u, \omega(u)), (v, \omega(v))) \in D_\omega$, $(u, \omega(u)) \Rightarrow (\lambda, e)$ with e the identity of the weighting space.

Definition 7 A weighted sticker system $\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot)$ is said to be simple if for all pairs $((u, \omega(u)), (v, \omega(v))) \in D_\omega$, either

$$(u, \omega(u)), (v, \omega(v)) \in \binom{\lambda}{V^*} \times M$$

or

$$(u, \omega(u)), (v, \omega(v)) \in \binom{V^*}{\lambda} \times M.$$

Definition 8 A weighted sticker system $\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot)$ is said to be simple one-sided if for all pairs $((u, \omega(u)), (v, \omega(v))) \in D_\omega$, either

$$(u, \omega(u)), (v, \omega(v)) \in \binom{\lambda}{V^*} \times M$$

or

$$(u, \omega(u)), (v, \omega(v)) \in \binom{V^*}{\lambda} \times M$$

and for each pair $((u, \omega(u)), (v, \omega(v))) \in D_\omega$, either $(u, \omega(u)) \Rightarrow (\lambda, e)$ or $(v, \omega(v)) \Rightarrow (\lambda, e)$ with e the identity of the weighting space.

Definition 9 A weighted sticker system $\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot)$ is said to be simple regular if for all pairs $((u, \omega(u)), (v, \omega(v))) \in D_\omega$, either

$$(u, \omega(u)), (v, \omega(v)) \in \binom{\lambda}{V^*} \times M$$

or

$$(u, \omega(u)), (v, \omega(v)) \in \binom{V^*}{\lambda} \times M$$

and for each pair $((u, \omega(u)), (v, \omega(v))) \in D_\omega$, $(u, \omega(u)) \Rightarrow (\lambda, e)$ with e the identity of the weighting space.

The languages generated by weighted one-sided, weighted regular, weighted simple, weighted simple one-sided and weighted simple regular sticker systems are denoted by ωOSL, ωRSL, ωSSL, ωSOSL, and ωSRSL, respectively. As with the usual sticker system, the languages generated by arbitrary weighted sticker system is denoted by ωUSL.

THE GENERATIVE POWER FOR VARIANTS OF WEIGHTED STICKER SYSTEMS

To investigate the generative power for variants of sticker systems, we associate weights to each axiom and domino of sticker systems and choose different cut-points for each weighting space. In previous studies, the generative power of sticker systems can be increased when weights are associated with the axioms and dominoes of sticker systems[9]. The generative power of a weighted sticker system increases when some cut-points are chosen for specific weighting operation of different weighting spaces as shown in Proposition 1.

Proposition 1 *For each* $X \in \{U, O, R, S, SO, SR\}$,

$$X\mathrm{SL} \subseteq \omega X\mathrm{SL}(M, \odot)$$

where $(M, \odot) \in \{(\mathbb{Z}, +), (\mathbb{Z}, \times), (\mathbb{Z}^k, +), (\mathbb{Z}^k, \times), (\mathbb{Q}^+, \times), (\mathbb{M}, +)\}$.

Proof: For any sticker system γ, we define the weighted sticker system γ' with the associating weights as
(a) $\{0\}$ if $(M, \odot) = (\mathbb{Z}, +)$;
(b) $\{1\}$ if $(M, \odot) = (\mathbb{Z}, \times)$;
(c) $\{(0, 0, \dots, 0)\}$ if $(M, \odot) = (\mathbb{Z}^k, +)$;
(d) $\{(1, 1, \dots, 1)\}$ if $(M, \odot) = (\mathbb{Z}^K, \times)$;
(e) $\{1\}$ if $(M, \odot) = (\mathbb{Q}^+, \times)$;
(f) $\{\mathbf{O}\}$ if $(M, \odot) = (\mathbb{M}, +)$.
Then with each axiom and domino of weighted sticker system γ, we can easily see that
(a) $L_\omega(\gamma', = 0) = L(\gamma)$;
(b) $L_\omega(\gamma', = 1) = L(\gamma)$;
(c) $L_\omega(\gamma', = (0, 0, \dots, 0)) = L(\gamma)$;
(d) $L_\omega(\gamma', = (1, 1, \dots, 1)) = L(\gamma)$;
(e) $L_\omega(\gamma', = 1) = L(\gamma)$;
(f) $L_\omega(\gamma', = \mathbf{O}) = L(\gamma)$.

$\square$

Next, we illustrate the generative power for variants of weighted sticker systems by giving some examples of weighted sticker systems with different weighting spaces, and show that for the same finite sets of axioms and dominoes on a sticker system, the selection of weighting spaces affects the generative power of a weighted sticker system, i.e., the same

sticker system with different weighting spaces can generate regular, context-free and context-sensitive languages.

Example 1 Given a weighted simple one-sided sticker system

$$\gamma_1 = \left(\{a, b, x\}, \{(a, a), (b, b)\}, \left\{ \binom{\lambda}{b} \binom{x}{x} \right. \right.$$

$$\binom{a}{\lambda}, \tau \right\}, \left\{ \left(\binom{\lambda}{\lambda}, \binom{a}{\lambda}, \tau_1 \right), \right.$$

$$\left(\binom{\lambda}{\lambda}, \binom{\lambda}{a}, \tau_1 \right), \left(\binom{\lambda}{b}, \binom{\lambda}{\lambda}, \tau_2 \right),$$

$$\left. \left. \left(\binom{b}{\lambda}, \binom{\lambda}{\lambda}, \tau_2 \right) \right\}, \omega, M, \odot \right).$$

It is clear that the language

$$L(\gamma_1) = \left\{ (b^m x a^n) \mid \left((b^m x a^n), \quad \omega(b^m x a^n) \right) \right.$$
$$\left. \in WK_\rho(V) \times M, n, m \geqslant 1 \right\}.$$

First, let the set of all integers $\mathbb{Z}$ to be the weighting space, and the addition of integers be the operation $\odot$. If $\tau_1 = 1$, $\tau_2 = -1$, it is clear that

$$L(\gamma_1, \star \tau) = \left\{ (b^m x a^n) \mid \left((b^m x a^n), \quad n-m \right) \right.$$
$$\left. \in WK_\rho(V) \times M, n, m \geqslant 1 \right\}.$$

Then with the cut-point $\tau = 0$,

$L(\gamma_1, = 0) = \{b^n x a^n \mid n \geqslant 1\}$, i.e. **CF-REG**,
$L(\gamma_1, > 0) = \{b^m x a^n \mid n > m \geqslant 1\}$, i.e. **CF-REG**,
$L(\gamma_1, < 0) = \{b^m x a^n \mid m > n \geqslant 1\}$, i.e. **CF-REG**.

Second, let the set $\mathbb{Z} \times \mathbb{Z}$ be the weighting space and the componentwise addition of pairs from $\mathbb{Z} \times \mathbb{Z}$ be the operation $\odot$. By letting $\tau_1 = (1, 0)$, $\tau_2 = (-1, 0)$, and cut-point as $(0, 0)$, the same languages as above are obtained.

Third, we consider the set of all 2×2 matrices with integer entries as the weighting space, the componentwise addition as the operation $\odot$. Then for

$$\tau_1 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

and

$$\tau_2 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

with cut-point $\tau = \mathbf{O}$, the same languages as above are also obtained.

Example 2 Given a weighted simple one-sided sticker system

$$\gamma_2 = \left(\{a, b, c, x\}, \{(a, a), (b, b), (c, c)\}, \right.$$

$$\left\{ \binom{x}{x}, \tau \right\}, \left\{ \left(\binom{a}{\lambda}, \binom{\lambda}{\lambda}, \tau_1 \right), \right.$$

$$\left(\binom{\lambda}{\lambda}, \binom{b}{\lambda}, \tau_2 \right), \left(\binom{\lambda}{\lambda}, \binom{c}{\lambda}, \tau_3 \right),$$

$$\left(\binom{\lambda}{a}, \binom{\lambda}{\lambda}, \tau_1 \right), \left(\binom{\lambda}{\lambda}, \binom{\lambda}{b}, \tau_2 \right),$$

$$\left. \left. \left(\binom{\lambda}{\lambda}, \binom{\lambda}{c}, \tau_3 \right) \right\}, \omega, M, \odot \right).$$

Then

$$L(\gamma_2) = \left\{ (a^n x b^m c^k), (a^n x c^k b^m) \mid \left((a^n x b^m c^k), \right. \right.$$
$$\omega(a^n x b^m c^k) \big), \big((a^n x c^k b^m), \omega(a^n x c^k b^m) \big)$$
$$\left. \in WK_\rho(V) \times M, n, m, k \geqslant 1 \right\}.$$

First, let $\mathbb{Q}^+$ be the weighting space and the multiplication of rational numbers be the operation $\odot$. Then for $\tau_1 = \frac{1}{3}$, $\tau_2 = \frac{1}{5}$, $\tau_3 = 15$ and cut-point $\tau = 1$, we obtain

$$L(\gamma_2, = 1) = \{a^n x b^n c^n, a^n x c^n b^n \mid n \geqslant 1\},$$
$$\text{i.e. } \textbf{CS-CF},$$
$$L(\gamma_2, > 1) = \{a^n x b^m c^k, a^n x c^k b^m \text{ for } k > m,$$
$$n \geqslant 1\}, \text{ i.e. } \textbf{CF-REG},$$
$$L(\gamma_2, < 1) = \{a^n x b^m c^k, a^n x c^k b^m \text{ for } m,$$
$$n > k \geqslant 1\}, \text{ i.e. } \textbf{CF-REG}.$$

Second, let the set $\mathbb{Z} \times \mathbb{Z}$ be the weighting space, and we define the componentwise addition of pairs from $\mathbb{Z} \times \mathbb{Z}$ be the operation $\odot$. By letting $\tau_1 = (1, 0)$, $\tau_2 = (-1, 1)$, $\tau_3 = (0, -1)$ and cut-point as $(0, 0)$, the following languages are obtained.

$$L(\gamma_2, = (0, 0)) = \{a^n x b^n c^n, a^n x c^n b^n \mid n \geqslant 1\},$$
$$\text{i.e. } \textbf{CS-CF},$$
$$L(\gamma_2, > (0, 0)) = \{a^n x b^m c^k, a^n x c^k b^m \text{ for}$$
$$n > m > k \geqslant 1\}, \text{ i.e. } \textbf{CS-CF},$$
$$L(\gamma_2, < (0, 0)) = \{a^n x b^m c^k, a^n x c^k b^m \text{ for}$$
$$k > m > n \geqslant 1\}, \text{ i.e. } \textbf{CS-CF}.$$

Third, we consider the set of 2×2 matrices with integer entries as the weighting space and the

componentwise addition as its operation. Then for

$$\tau_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\tau_2 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix},$$

and

$$\tau_3 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

with cut-point $\tau = \mathbf{O}$, the same languages as above are obtained.

Example 1 and Example 2 show the generative power of two weighted simple one-sided sticker systems. Next, we consider the cases for weighted simple regular sticker systems, given as Example 3 and Example 4.

Example 3 Given a weighted simple regular sticker system

$$\gamma_3 = \left(\{a, b, x\}, \{(a, a), (b, b)\}, \left\{ \begin{pmatrix} x \\ x \end{pmatrix}, \tau \right\}, \right.$$
$$\left\{ \left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} \lambda \\ a \end{pmatrix}, \tau_1 \right), \left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} \lambda \\ b \end{pmatrix}, \tau_2 \right), \right.$$
$$\left. \left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} a \\ \lambda \end{pmatrix}, \tau_1 \right), \left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} b \\ \lambda \end{pmatrix}, \tau_2 \right) \right\},$$
$$\left. \omega, M, \odot \right).$$

We obtain the language $L(\gamma_3)$, where

$$L(\gamma_3) = \left\{ (xa^n b^m) \mid ((xa^n b^m), \omega(xa^n b^m)) \right.$$
$$\left. \in WK_\rho(V) \times M, \ n, m \geqslant 1 \right\}.$$

First, let the set of all integers $\mathbb{Z}$ be the weighting space and the addition of integers be the operation $\odot$. If $\tau_1 = 1$ and $\tau_2 = -1$, it is clear that

$$L(\gamma_3, \star\tau) = \left\{ (xa^n b^m, xb^m a^n) \mid ((xa^n b^m), n-m), \right.$$
$$((xb^m a^n), n-m) \in WK_\rho(V) \times M,$$
$$\left. n, m \geqslant 1 \right\}.$$

Then with the cut-point $\tau = 0$,

$$L(\gamma_3, = 0) = \{xa^n b^n, xb^m a^n \mid n \geqslant 1\},$$
$$\text{i.e. } \textbf{CF-REG},$$
$$L(\gamma_3, > 0) = \{xa^n b^m, xb^m a^n \mid n > m \geqslant 1\},$$
$$\text{i.e. } \textbf{CF-REG},$$
$$L(\gamma_3, < 0) = \{xa^n b^m, xb^m a^n \mid m > n \geqslant 1\},$$
$$\text{i.e. } \textbf{CF-REG}.$$

Second, let the set $\mathbb{Z} \times \mathbb{Z}$ be the weighting space and the componentwise addition of pairs from $\mathbb{Z} \times \mathbb{Z}$ be the operation $\odot$. By letting $\tau_1 = (1, 0)$, $\tau_2 = (-1, 0)$, and cut-point as $(0, 0)$, the same languages as above are obtained.

Third, we consider the set of all 2×2 matrices with integer entries as the weighting space, and the componentwise addition as the operation $\odot$. Then for

$$\tau_1 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

and

$$\tau_2 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

with cut-point $\tau = \mathbf{O}$, the same languages as above are also obtained.

Example 4 Given a weighted simple regular sticker system

$$\gamma_4 = \left(\{a, b, c, x\}, \{(a, a), (b, b), (c, c)\}, \right.$$
$$\left\{ \begin{pmatrix} x \\ x \end{pmatrix}, \tau \right\}, \left\{ \left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} a \\ \lambda \end{pmatrix}, \tau_1 \right), \right.$$
$$\left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} b \\ \lambda \end{pmatrix}, \tau_2 \right), \left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} c \\ \lambda \end{pmatrix}, \tau_3 \right),$$
$$\left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} \lambda \\ a \end{pmatrix}, \tau_1 \right), \left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} \lambda \\ b \end{pmatrix}, \tau_2 \right),$$
$$\left. \left(\begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} \lambda \\ c \end{pmatrix}, \tau_3 \right) \right\}, \omega, M, \odot \right).$$

Hence it is easy to see that we obtain the language $L(\gamma_4)$ such that

$$L(\gamma_4) = \left\{ (xa^n b^m c^k) \mid ((xa^n b^m c^k), \omega(xa^n b^m c^k)) \right.$$
$$\left. \in WK_\rho(V) \times M, \ n, m, k \geqslant 1 \right\}.$$

Firstly, let $\mathbb{Q}^+$ be the weighting space and the multiplication of rational numbers be the operation $\odot$. Then for $\tau_1 = \frac{1}{3}$, $\tau_2 = \frac{1}{5}$, $\tau_3 = 15$, and cut-point $\tau = 1$. We obtain

$$L(\gamma_4, = 1) = \{xa^n b^n c^n \mid n \geqslant 1\}, \text{ i.e. } \textbf{CS-CF},$$
$$L(\gamma_4, > 1) = \{xa^n b^m c^k \mid k > m, n \geqslant 1\},$$
$$\text{i.e. } \textbf{CF-REG},$$
$$L(\gamma_4, < 1) = \{xa^n b^m c^k \mid m, n > k \geqslant 1\},$$
$$\text{i.e. } \textbf{CF-REG}.$$

Second, let the set $\mathbb{Z} \times \mathbb{Z}$ be the weighting space and the componentwise addition of pairs from $\mathbb{Z} \times \mathbb{Z}$ be the operation $\odot$. By letting $\tau_1 = (1, 0)$, $\tau_2 =$

$(-1, 1)$, $\tau_3 = (0, -1)$, and cut-point as $(0, 0)$, the following languages are obtained.

$$L(\gamma_4, = (0, 0)) = \{xa^n b^m c^k \mid n \geqslant 1\}, \text{ i.e. } \textbf{CS-CF},$$

$$L(\gamma_4, > (0, 0)) = \{xa^n b^m c^k \mid n > m > k \geqslant 1\},$$
$$\text{i.e. } \textbf{CS-CF},$$

$$L(\gamma_4, < (0, 0)) = \{xa^n b^m c^k \mid k > m > n \geqslant 1\},$$
$$\text{i.e. } \textbf{CS-CF}.$$

Third, we consider the set of 2×2 matrices with integer entries as the weighting space and the componentwise addition as its operation, the same languages as above will be obtained if

$$\tau_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\tau_2 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix},$$

and

$$\tau_3 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

with cut-point $\tau = \textbf{O}$.

To further investigate the generative power for variants of weighted sticker system, we use Lemma 1 and Lemma 2 as stated below.

Lemma 1 (Ref. 20) *For each* $X \in \{O, R\}$,

$$X\text{SL} \subset \omega X\text{SL}(M, \odot)$$

where $(M, \odot) \in \{(\mathbb{Z}, +), (\mathbb{Z}, \times), (\mathbb{Z}^k, +), (\mathbb{Z}^k, \times), (\mathbb{Q}^+, \times), (\mathbb{M}, +)\}$.

Lemma 2 (Ref. 20) *For* $X \in \{O, R\}$, $\omega X\text{SL}$ *contains some non-context-free languages.*

Next, we obtain Theorem 2 by using Proposition 1 and Example 1 to Example 4; Theorem 3 by using Example 1 and Example 3; and Theorem 4 by using Example 2 and Example 4.

Theorem 2 *For each* $X \in \{SO, SR\}$,

$$X\text{SL} \subset \omega X\text{SL}(M, \odot)$$

where $(M, \odot) \in \{(\mathbb{Z}, +), (\mathbb{Z}, \times), (\mathbb{Z}^k, +), (\mathbb{Z}^k, \times), (\mathbb{Q}^+, \times), (\mathbb{M}, +)\}$.

Theorem 3 *For* $X \in \{SO, SR\}$,

$$\omega X\text{SL} - \textbf{REG} \neq \varnothing.$$

Theorem 4 *For* $X \in \{SO, SR\}$, $\omega X\text{SL}$ *contains some non-context-free languages.*

From Lemma 1 and Theorem 2, we have Theorem 5.

Theorem 5 *For each* $X \in \{O, R, SO, SR\}$,

$$X\text{SL} \subset \omega X\text{SL}(M, \odot)$$

where $(M, \odot) \in \{(\mathbb{Z}, +), (\mathbb{Z}, \times), (\mathbb{Z}^k, +), (\mathbb{Z}^k, \times), (\mathbb{Q}^+, \times), (\mathbb{M}, +)\}$.

From Lemma 2 and Theorem 4, we have Theorem 6.

Theorem 6 *For* $X \in \{O, R, SO, SR\}$,

$$\omega X\text{SL} - \textbf{CF} \neq \varnothing.$$

Now we consider the cases for different weighting spaces. Surprisingly, weighted sticker systems with different weighting spaces have different levels of generative power.

GENERATIVE POWER OF VARIANTS OF WEIGHTED STICKER SYSTEMS WITH DIFFERENT WEIGHTING SPACES

Firstly, we choose the n-dimensional vector space $\mathbb{Z}^n$, where $n \geqslant 0$, over integers as the weighting space and vector addition as the operation over weights. Then, it is clear that

$$\omega X\text{SL}(\mathbb{Z}^n, +) = \omega X\text{SL},$$

where $X \subseteq \{R, O, SO, SR, S\}$. For $n \geqslant 1$, we have the following results.

Theorem 7 $\omega\text{SL}(\mathbb{Z}^n, +) \subseteq \omega\text{SL}(\mathbb{Z}^{n+1}, +)$.

Proof: Let L be the language generated by $\omega\text{SL}(\mathbb{Z}^n, +)$. Then there is a weighted sticker system $\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot)$, $M \subseteq \mathbb{Z}^n$ such that $L = L_\omega(\gamma, \star\alpha)$. We construct the weighted sticker system $\gamma' = (V, \rho, A'_\omega, D'_\omega, M', \odot)$, $M' \subseteq \mathbb{Z}^{n+1}$ where $A'_\omega = \{x, (a_1, a_2, \ldots, a_n, a_n) \mid (a_1, a_2, \ldots, a_n, a_n) \in M'$ and $(x, (a_1, a_2, \ldots, a_n)) \in A_\omega\}$, $D'_\omega = \{y, (b_1, b_2, \ldots, b_n, b_n) \mid (b_1, b_2, \ldots, b_n, b_n) \in M'$ and $(y, (b_1, b_2, \ldots, b_n)) \in D_\omega\}$. Then it is not difficult to see that for every sticker operation in γ such that

$$[(x, (a_1, a_2, \ldots, a_n)), (y, (b_1, b_2, \ldots, b_n))]$$
$$\Longrightarrow^* [z, (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)],$$

we can define the sticker system in γ' as

$$[(x, (a_1, a_2, \ldots, a_n, a_n)), (y, (a_1, a_2, \ldots, a_n, a_n))]$$
$$\longrightarrow^* [z, (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)].$$

Since the nth and $(n+1)$th components in M' are the same, we conclude that $L_\omega(\gamma', \star\alpha) = L_\omega(\gamma, \star\alpha)$. $\square$

Theorem 8 $\omega X \mathrm{SL}(\mathbb{Z}, +) \subseteq \omega X \mathrm{SL}(\mathbb{Q}^+, \times)$.

Proof: Let γ be a weighted sticker system such that

$$\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot),$$

where $M \subseteq \mathbb{Z}$ and addition is the weight operation. We construct the weighted sticker system $\gamma' = (V, \rho, A'_\omega, D'_\omega, \omega', M', \odot), M' \subseteq \mathbb{Q}$ with multiplication as the weight operation where

$$A'_\omega = \{(x, 2^{\omega(x)}) \mid \omega(x) \in M, \ (x, \omega(x)) \in A_\omega\}.$$

Then for every string x in $L_\omega(\gamma)$ with its weight as $\omega(x)$, it has weights in $L_\omega(\gamma')$ as $2^{\omega(x)}$. Hence, for any cut-point α for weighted sticker system γ, $L_\omega(\gamma', \star\alpha) = L_\omega(\gamma, \star\alpha)$ if the cut-point for weighted sticker system γ' is $2^{\omega(x)}$. □

We now consider the set of $n \times n$ matrices $\mathbb{M}_n$ as the weighting space. We denote $\omega\mathrm{SL}(\mathbb{M}_n)$ to be the families of languages generated by weighted sticker system with the set of $n \times n$ matrices $\mathbb{M}_n$ as the weighting space, and componentwise addition of matrices as the weight operation. Thus it is clear that for $n \geqslant 0$,

$$\omega X \mathrm{SL}(\mathbb{M}_0, +) = \omega X \mathrm{SL}.$$

For $n > 0$, we have the following theorem.

Theorem 9 $\omega X \mathrm{SL}(\mathbb{M}_n, +) \subseteq \omega X \mathrm{SL}(\mathbb{M}_{n+1}, +)$.

Proof: Let L be the language generated by $\omega\mathrm{SL}(\mathbb{M}_n, +)$. Then there is a weighted sticker system $\gamma = (V, \rho, A_\omega, D_\omega, \omega, M, \odot)$, $M \subseteq \mathbb{M}_n$ such that $L = L_\omega(\gamma, \star\alpha)$.

We construct the weighted sticker system

$$\gamma' = (V, \rho, A'_\omega, D'_\omega, M', \odot), \quad M' \subseteq \mathbb{M}_{n+1}$$

where

$$A'_\omega = \left\{ x, \begin{pmatrix} a_{ij} & a_{in} \\ a_{nj} & a_{nn} \end{pmatrix} \middle| \begin{pmatrix} a_{ij} & a_{in} \\ a_{nj} & a_{nn} \end{pmatrix} \in M', \right.$$
$$\left. (x, [a_{ij}]) \in A_\omega, \ [a_{ij}] \in M \right\}$$

and

$$D'_\omega = \left\{ y, \begin{pmatrix} b_{ij} & b_{in} \\ b_{nj} & b_{nn} \end{pmatrix} \middle| \begin{pmatrix} b_{ij} & b_{in} \\ b_{nj} & b_{nn} \end{pmatrix} \in M', \right.$$
$$\left. (x, [b_{ij}]) \in A_\omega, \ [b_{ij}] \in M \right\}.$$

Then it is not difficult to see that for every sticker operation in γ such that

$$[(x, [a_{ij}]), (y, [b_{ij}])] \longrightarrow^* (z, [a_{ij} + b_{ij}]),$$

we can define the sticker operation in γ' as

$$\left[\left(x, \begin{pmatrix} a_{ij} & a_{in} \\ a_{nj} & a_{nn} \end{pmatrix} \right), \left(y, \begin{pmatrix} b_{ij} & b_{in} \\ b_{nj} & b_{nn} \end{pmatrix} \right) \right]$$
$$\longrightarrow^* \left[z, \begin{pmatrix} a_{ij} + b_{ij} & a_{in} + b_{in} \\ a_{nj} + b_{nj} & a_{nn} + b_{nn} \end{pmatrix} \right].$$

Since the nth and $(n+1)$th row and column of matrices in M' are the same, they fulfil the same cut-point requirement. Hence we conclude that $L_\omega(\gamma', \star\alpha) = L_\omega(\gamma, \star\alpha)$. □

By using the same argument as in the proof of Theorem 7 to Theorem 9, we can obtain Theorem 10 as follows.

Theorem 10 *For $n \geqslant 1$,*

$$\omega X \mathrm{SL}(\mathbb{Z}^n, +) \subseteq \omega X \mathrm{SL}(\mathbb{M}_n, +).$$

REFERENCES

1. Head T (1987) Formal language theory and DNA: an analysis of the generative capacity of specific recombination behaviors. *Bull Math Biol* **49**, 737–59.
2. Adleman L (1994) Molecular computation of solutions to combinatorial problems. *Science* **266**, 1021–4.
3. Lipton R (1995) Using DNA to solve NP–complete problems. *Science* **268**, 542–5.
4. Boneh D, Dunworth C, Lipton R, Sgall J (1996) On the computational power of DNA. *Discrete Appl Math* **71**, 79–94.
5. Kari L, Paun G, Rozenberg G, Salomaa A, Yu S (1998) DNA computing, sticker systems and universality. *Acta Inform* **35**, 401–20.
6. Păun G, Rozenberg G, Salomaa A (1998) *DNA Computing: New Computing Paradigms*, Springer-Verlag, Berlin.
7. Mohd Sebry NA, Hamzah NZA, Sarmin NH, Fong WH, Turaev S (2012) Sticker systems over monoids. *Malays J Fund Appl Sci* **8**, 131–6.
8. Selvarajoo M, Fong WH, Sarmin NH, Turaev S (2013) Probabilistic sticker systems. *Malays J Fund Appl Sci* **9**, 150–5.
9. Gan YS, Fong WH, Sarmin NH, Turaev S (2013) Some properties and variants of weighted sticker system. *Int J Appl Math Stat* **45**, 367–75.
10. Salomaa A (1969) Probabilistic and weighted grammars. *Inform Contr* **15**, 529–44.
11. Fu KS, Li T (1969) On stochastic automata and languages. *Inform Sci* **1**, 403–49.

12. Mizumoto M, Toyoda J, Tanaka K (1973) Examples of formal grammars with weights. *Inform Process Lett* **2**, 74–8.

13. Mizumoto M, Toyoda J, Tanaka K (1975) Various kinds of automata with weights. *J Comput Syst Sci* **10**, 219–36.

14. Alexandrakis A, Bozapalidis S (1987) Weighted grammars and Kleene's theorem. *Inform Process Lett* **24**, 1–4.

15. Droste M, Kuich W, Vogler H (2009) *Handbook of Weighted Automata*, Springer-Verlag, Berlin.

16. Rozenberg G, Salomaa A (1997) *Handbook of Formal Languages*, Springer-Verlag, Berlin.

17. Dassow J, Pǎun G (1989) *Regulated Rewriting in Formal Language Theory*, Springer-Verlag, Berlin.

18. Pixton D (1996) Regularity of splicing languages. *Discrete Appl Math* **69**, 101–24.

19. Linz P (2006) *An Introduction to Formal Languages and Automata*, Jones and Bartlett Publishers, USA.

20. Gan YS, Fong WH, Sarmin NH, Turaev S (2014) Generative power of weighted one-sided and regular sticker system. *AIP Conf Proc* **1602**, 855–62.