International Science Postgraduate Conference 2012 (ISPC 2012)

WEIGHTED STICKER SYSTEM

GAN YEE SIANG [1], FONG WAN HENG[2*] NOR HANIZA SARMIN[3]and SHERZOD TURAEV[4]

Abstract. A sticker system is a computational model which usesa sticker operation on DNA molecules. A sticker operation works on the complementary relation of double stranded DNA by using ligation and annealing operation to form a complete double stranded DNA sequence. In this paper, a new variant of sticker systems, called*weighted sticker systems*, is introduced. Some basic properties of language families that are generated by the weighted sticker systems are investigated.This paper also introduces some restricted weighted variants of sticker systems such as weighted one-sided, regular, simple, simple one-sided and simple regular sticker systems. Moreover, the paper shows that the presence of weights increases the generative powers of usual variants of sticker systems.

Keywords weighted sticker systems; one-sided sticker systems; regular languages; simple sticker systems

Abstrak. Sistem pelekat adalah satu model pengiraan yang menggunakan operasi pelekat pada molekul-molekul DNA. Operasi pelekat berfungsi pada hubungan pelengkap DNA bebenang ganda dua dengan menggunakan ligasi dan operasi penyepuh lindapan untuk membentuk jujukan DNA bebenang ganda dua yang lengkap .Dalam kerta kerja ini, satu variasisi stem pelekat yang baru iaitu sistem pelekat berwajaran telah diperkenalkan. Sebahagian ciri-ciri asas daripada keluarga bahasa yang dijana oleh sistem pelekat berwajaran akan disiasat. Kertas kerja ini juga mempekenalkan sebahagian daripada wajaran terhad bagi varian sistem pelekat, iaitu sistem pelekat berwajaran satu sisi, biasa, mudah, mudah satu sisi dan mudahb iasa. Selain itu, kuasa penjanaan untuk varian sistem pelekat boleh dipertingkatkan dengan adanya wajaran.

Kata kunci sistem pelekat berwajaran; satus isi; biasa; mudah; kuasa penjanaan

---

[1, 3] Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor DarulTa'azim, Malaysia
[2] IbnuSinaInstitute for Fundamental Science Studies, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor DarulTa'azim, Malaysia
[4] Department of Computer Science, Kulliyyah of Information and Communication Technology, International Islamic University Malaysia, 50728 Kuala Lumpur, Malaysia
[*]Corresponding author: fwh@utm.my

1. INTRODUCTION

DNA computing has been proven possible after the first DNA computation experiment has been performed by Adleman [1]. DNA computing is based on double stranded structure of DNA molecules. The structure of DNA molecules is composed by four nucleotides: A (adenine), C (cytosine), G (guanine) and T (thymine). By the feature of Watson-Crick complementarity, adenine always bonds with thymine and guanine bonds with cytosine. Because of the complementarity feature, the information can make far-reaching conclusions from the encoded information in one of the single strands to another[2]. Another interesting feature on DNA computing is massive parallelism. This feature allows the construction of many copies of a DNA strand, so that the information can be encoded simultaneously. There are many results on the power of this new computation device; for instance, the application of DNA computing with the features of Watson-Crick complementary and massive parallelism can be found in [1, 3,4].

From the techniques used by Adleman in his experiment, a new computing model, called the sticker system, is introduced in [5]. A sticker system is a computational model which usesa sticker operation on DNA molecules.The sticker operation is used as a generating device to producea "DNA" language from a given set of incomplete DNA sequences to form a complete double stranded sequence. The incomplete DNA stranded sequences stick with the complementary single stranded sequences to form a complete double stranded sequence, or in other word, no "sticky end" exists.

Since without any restriction, sticker systems only generate regular languages[5]. Since several restrictions to sticker systems have been studied.In [6–8], it was shown that the use of some algebra structures such as monoids and groups increases the computational power of sticker systems. It is very interesting from the point of view in DNA computing, since theoretical models of universal programmable DNA based computers can be modeled by using restricted variants of sticker systems [4].

In this paper, a new variant of sticker systems, called weighted sticker systemsis introduced. That is, with the finite pairs of axiom and dominoes ($LR_\rho(V)$ and $W_\rho(V) \times W_\rho(V)$) weights are associated. The weight $w(z)$ of the strand $z$produced by strands$x$ and y of $W_\rho(V)$ is calculated from the weights $w(x)$and $w(y)$ according to the operation $\oplus$ defined as $w\ (z) = w(x) \oplus w(y)$. Therefore, several types of threshold languages generated by weighted sticker system have been considered with different weighted spaces and cut-point. The generative power of this variant of sticker systems affected by the selection of the weighting spaces and cut-points is also presented in this paper.

The sticker system starts with initial sequence from A (axioms) and then prolong from the left to the right by using pairs (u,v) of dominoes in D according to the sticker operation μ. The prolongation is continued until there is no blank symbol to obtain a complete double stranded sequence in $WK_{\rho}(V)$ [9]. 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 $(u, v) \in D$. Since the prolongation to the right is independent to the left, thus $\mu(u, \mu(x, v)) = \mu(\mu(u, x), v)$. When $x_1 \in A$ is used as starting symbols and $x_k$is in $WK_{\rho}(V)$, a sequence such that $x_1 \Rightarrow x_2 \Rightarrow \cdots \Rightarrow x_k$is obtained. This sequence is called computation in γ with length of $k - 1$. If there is no sticky end or no blank symbol present in the last symbolthen the computation is considered as complete. A complete computation will produce a complete string w such that$w \in WK_{\rho}(V)$. Therefore, the language of such strings is the language generated by γ or called the sticker language. The concepts of a sticker language a and grammar will be explained in the next section.

2. PRELIMINARIES

In the following subsections, some information on the sticker language and grammar are given as in subsections 2.1 and 2.2 respectively.

2.1 Sticker Language

A sticker language is an abstract model of annealing operations that occurs in DNA computing. It consists of all complete moleculesderived from the axiom with some pair of incomplete molecules.Therefore, the language that is generated by sticker system γ is defined as

$$L(\gamma) = \{w \in \binom{V}{V}^* \mid x \Rightarrow^* w, x \in A\}.$$

There are six families of language generated by sticker systems, which are sticker languages, primitive sticker languages, balanced sticker languages, primitive balanced sticker languages, coherent sticker languages and fair sticker languages denoted by SL, PSL, BSL, $PBSL$, CSL and FSL, respectively [9]. The variants of sticker systems are defined as follows.

Definition 1[2]: **One-Sided Sticker Language (*OSL*)**
The language resulting from a sticker system is called a **one-sided sticker language** if for each pair $(u,v) \in D$, either $u = \lambda$ or $v = \lambda$.

Definition 2 [2]: **Regular Sticker Language (*RSL*)**
The language resulting from a sticker system is called a **regular sticker language** if for each pair $(u, v) \in D$, $u = \lambda$.

Definition 3 [2]: **Simple Sticker Language (*SSL*)**
The language resulting from a sticker system is called a**simple sticker language** if for all pairs$(u, v) \in D$, either $(u, v) \in \binom{V^*}{\lambda}$ or $(u, v) \in \binom{\lambda}{V^*}$.

Definition 4 [2]: **Simple One-Sided Sticker Language (*SOSL*)**
The language resulting from a sticker system is called a**simple one-sided sticker language** if for all pairs $(u, v) \in D$ either $(u, v) \in \binom{V^*}{\lambda}$ or $(u, v) \in \binom{\lambda}{V^*}$ and for each pair $(u, v) \in D$, either $u = \lambda$ or $v = \lambda$.

Definition 5 [2]: **Simple Regular Sticker Language (*SRSL*)**
The language resulting from a sticker system is called a**simple regular sticker language** if for all pairs$(u, v) \in D$,either $v\in \binom{V^*}{\lambda}$ or $v \in \binom{\lambda}{V^*}$ for $u = \lambda$.

In formal language theory, languagesare described by generative devices called *grammars*. The basic definitions of a grammar are discussed in the next subsection.

2.2 Grammar

A grammar G is defined as a quadruple $G = (V, T, S, P)$, where V is a finite set of objects called variables, T is a finite set of objects called terminal symbols, S is the initial variable such that $S \in V$ and P is a finite set of production rules. Note that, the sets V and T are nonempty and disjoint. The set P is finite set of pairs $(u, v)$where $v \in (N \cup T)^*$ and $u$contain at least one symbol from N. Throughout this paper, the production rules are written in the form such that $u \rightarrow v$ where $u \in (N \cup T)^+$ and $v \in (N \cup T)^*$. The families of finite, regular, linear, context-free, context-sensitive and recursively enumerable languages are denoted by FIN, REG, LIN, CF, CS and RE, respectively [10].

In order to investigate different aspects of modeled phenomena different restricted variants of "classical" grammars have introduced. One of such restricted variants, called weighted grammars, has beeninvestigated since 1960s.

The use of weights enables to develop more efficient parsing and tagging algorithms for natural language processing.There are several papers (for instance, see [11-13])related to the study ofweighted grammars.

3. WEIGHTED STICKER SYSTEMS

In the next section, some notations of weighted sticker systems with specified weighting spaces and operations over weights are introduced.

Definition 6: Weighted Sticker System
A **weighted sticker system** is a 7-tuple such as

$$\gamma = (V, \rho, A_w, D_w, w, M, \otimes)$$

where,
V is an alphabet,
ρ is the symmetric relation such as $\rho \subseteq V \times V$,
$A_w$is a finite set of axiom such as $A \subseteq LR_\rho (V) \times M$,
$D_w$is a finite subset of $(W_\rho(V) \times W_\rho(V)) \times M$,
w is a weighting function such that$w$: $(LR_\rho (V) \cup (W\rho(V) \times W\rho(V))) \rightarrow M$,
M is a weighting space,
$\otimes$ is the operation over the weights $w(x)$, $x \in (LR_\rho (V) \cup (W\rho(V) \times W\rho(V)))$.

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

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

Definition 7: (Weighted Sticker Operation)
For $(x, w(x))$, $(y, w(y)) \in A_w$and $((u, w(u)), (v, w(v))) \in D_w$, $[x, w(x)] \Rightarrow^* [y, w(y)]$ if and only if
(i)$[y, w(y)] = \mu([u, w(u)],\mu([x, w(x)], [v, w(v)])$ and $w(y) = w(y) \otimes w(v)$,
(ii) $[y, w(y)] = \mu(\mu([x, w(x)], [u, w(u)]), [v, w(v)])$ and $w(y) = w(y) \otimes w(u)$.

The language generated by the weighted sticker language is defined as

$$wSL(\gamma) = \{y \in WK_\rho(V) \mid [x, w(x)] \Rightarrow^* [y, w(y)] \text{ for } [x, w(x)] \in A\}.$$

To increase the generative power of sticker systems, we can consider languages generated by weighted splicing systems with some subsets of the weighting space called threshold (cut-point). We consider three types of threshold languages.

Definition 8: (Threshold Weighted Sticker SystemLanguage)

Let $wSL(\gamma)$ be the language generated by a weighted sticker system $\gamma = (V, \rho, A, D, w, M, \otimes)$. A **threshold weighted sticker systemlanguage** with respect to a threshold (cut-point) $\tau \in M$ is subset of $wSL(\gamma)$ defined by

$$wSL(\gamma, *\tau) = \{ y \in WK\rho(V) \mid [x, w(x)] \Rightarrow^* [y, w(y)] \text{ for } [x, w(x)] \in A \text{ and } w(y) * \tau \},$$

where $* \in \{=, <, >\}$ is called the mode of $wSL(\gamma, *\tau)$.

From Definition 6 to 8, we obtained Lemma 1, which showsthat the generative power of sticker systems increases by using some weights.

Lemma 1:

$$SL \subseteq w_\alpha SL$$

where$\alpha \in \{a, m\}$ and$w_a$indicates weighted sticker systems with addition operation and $w_m$indicates weighted sticker systems with multiplicative operation.

Proof:

Consider a sticker system $\gamma = (V, \rho, A, D)$. Then the language generated by the sticker system γ is

$$SL(\gamma) = \{z \in \binom{V}{V}^* \mid x \Rightarrow^* z, x \in A\}.$$

Let$\gamma_1 = (V_1, \rho_1, A_w, D_w, w, M, \otimes)$ be a weighted sticker system where$V_1 = V, \rho_1 = \rho$, $A_w = A \times M$, $D_w = D \times M$.The language generated by the weighted sticker systemγ_1:

$$wSL(\gamma_1, *\tau) = \{ z \in \binom{V}{V}^* \mid [x, w(x)] \Rightarrow^* [z, w(z)] \text{ for } [x, w(x)] \in A_w\},$$

where $w(z) = w(x) \otimes w(y_1) \otimes w(y_2) \otimes \ldots \otimes w(y_n)$ for $y_1, y_2, \ldots, y_n \in D_w$.

Hence, we have 2 cases as follows:

Case 1: *weighted sticker system with addition operation.*
If we assign to each axiom and to each element in $D_w$the weight 0, then it is easy to see that $w_aSL(\gamma_1, *0) = SL(\gamma)$. Thus it can be concluded that $SL(\gamma) \subseteq w_aSL$.

Case 2: *weighted sticker system with multiplication operation.*
We assign to each axiom and to each element in $D_w$the weight 1, then it is easy to see that $w_mSL(\gamma_1, *1) = SL(\gamma)$,again$SL(\gamma) \subseteq w_mSL$.

4.0 RESTRICTED VARIANTS OF WEIGHTED STICKER SYSTEMS

In this section, we define weighted variants of some restrictions of sticker systems. We show that the presence of weights increases the generative power of these variants of sticker systems.

Definition 9: Weighted One-Sided Sticker System(*wOSL*)
A weighted sticker system $\gamma = (V, \rho, A, D, w, M, \otimes, A', D')$is said to be a **weighted one-sided sticker system** if for each pair $((u, w(u)), (v, w(v))) \in D'$, either $(u, w(u)) \to (\lambda, e)$ or $(v, w(v)) \to (\lambda, e)$ with e is the identity of the weighting space.

Definition 10: Weighted Regular Sticker System (*wRSL*)
A weighted sticker system $\gamma = (V, \rho, A, D, w, M, \otimes, A', D')$is said to be a **weighted regular sticker system** if for each pair $((u, w(u)), (v, w(v))) \in D'$, $(u, w(u)) \to (\lambda, e)$ with e is the identity of the weighting space.

Definition 11: Weighted Simple Sticker System (*wSSL*)
A weighted sticker system $\gamma = (V, \rho, A, D, w, M, \otimes, A', D')$is said to be a **weighted simple sticker system** if for all pairs $((u, w(u)), (v, w(v))) \in D'$, either $(u, w(u)), (v, w(v)) \in \binom{\lambda}{V^*} \times M$ or $(u, w(u)), (v, w(v)) \in \binom{V^*}{\lambda} \times M$.

Definition 12: Weighted Simple One-Sided Sticker System (*wSOSL*)
A weighted sticker system $\gamma = (V, \rho, A, D, w, M, \otimes, A', D')$is said to be a **weighted simple one-sided sticker system** if for all pairs $((u, w(u)), (v, w(v))) \in D'$, either $(u, w(u)), (v, w(v)) \in \binom{\lambda}{V^*} \times M$ or $(u, w(u)), (v, w(v)) \in \binom{V^*}{\lambda} \times M$ and for each pair $((u, w(u)), (v, w(v))) \in D'$, either $(u, w(u)) \to (\lambda, e)$ or $(v, w(v)) \to (\lambda, e)$ with e is the identity of the weighting space.

Definition 13: Weighted Simple Regular Sticker System (*wSRSL*)
A weighted sticker system γ = (V, ρ, A, D, w, M, $\otimes$, A', D')is said to be a **weighted simple regular sticker system** if for all pairs ((u, $w(u)$), (v, $w(v)$)) $\in D'$, either (u, $w(u)$), (v, $w(v)$) $\in \binom{\lambda}{V^*} \times M$ or (u, $w(u)$), (v, $w(v)$) $\in \binom{V^*}{\lambda} \times M$ and for each pair ((u, $w(u)$), (v, $w(v)$)) $\in D'$, (u, $w(u)$) $\rightarrow$ (λ, e) with e is the identity of the weighting space.

By using the similar argument in the proof of Lemma 1, weobtained the following results which show that the presence of weights can increase the generative power of restricted variants of sticker system.

Theorem 1:
$OSL \subseteq w_\alpha OSL$ such that $\alpha \in \{a, m\}$ where$w_a$stands for a weighted space with addition operation and $w_m$stands for a weighted space with multiplicative operation.

Theorem 2:
$RSL \subseteq w_\alpha RSL$ such that $\alpha \in \{a, m\}$ where$w_a$stands for a weighted space with addition operation and $w_m$stands for a weighted space with multiplicative operation.

Theorem 3:
$SSL \subseteq w_\alpha SSL$ such that $\alpha \in \{a, m\}$ where $w_a$stands for a weighted space with addition operation and $w_m$stands for a weighted space with multiplicative operation.

Theorem 4:
$SOSL \subseteq w_\alpha SOSL$ such that $\alpha \in \{a, m\}$where$w_a$stands for a weighted space with addition operation and $w_m$stands for a weighted space with multiplicative operation.

Theorem 5:
$SRSL \subseteq w_\alpha SRSL$ such that $\alpha \in \{a, m\}$ where$w_a$stands for a weighted space with addition operation and $w_m$stands for a weighted space with multiplicative operation.

5. CONCLUSION

In this paper, the definition of a new restriction of sticker system namely theweighted sticker system has been introduced. In addition, some resultsshowthat the generative power of sticker systems can be increased with the presence of weights. In addition, the definitions of some new restrictions of variants for sticker system, namely, weighted one-sided, regular, simple, simple one-sided and simple regular sticker systemsare also introduced. Moreover, the generative power for each of the variants of restricted sticker system are also shown to increase with the presence of weights as mentioned in Theorem 1 to 5.

ACKNOWLEDGMENTS

The first author would like to thank the Ministry of Higher Education Malaysia for the financial funding through MyBrain15 scholarship. The second and third authors would also like to thank the Ministry of Higher Education (MOHE) and Research Management Centre (RMC), UTM for the financial funding through Research University Fund Vote No. 05J15.

REFERENCES

[1] Adleman, L. *Molecular computation of solutions to combinatorial problems*. Science (266) (Nov. 11, 1994) 1021–1024.

[2] Rozenberg, G., paun, G., Salomaa, A. *DNA computing: A New Computing Paradigms*. New York. Springer-Verlag. 1998.

[3] Lipton, R. *Using DNA to solve NP–complete problems*. Science (268) (Apr. 28, 1995) 542–545.

[4] Boneh, D., Dunworth, C., Lipton, R., Sgall, J.: *On the Computational Power of DNA*. Discrete Applied Mathematics, Special Issue on Computational Molecular Biology (71) (1996) 79–94.

[5] Kari, L., Paun, G. Rozenberg, G. Salomaa, A.and Yu, S. DNA Computing, Sticker Systems and Universality.*ActaInformatica*. 1998. 35:401 – 420.

[6] Hamzah, N. Z. A., Mohdsebry, N. A., Sarmin, N. H., Fong, W. H., Turaev, S. Splicing System over Group.*Regional Annual Fundamental Science Symposium*. 2011. (to appear)

[7] Mohdsebry, N. A., Hamzah, N. Z. A., Sarmin, N. H., Fong, W. H., Turaev, S. Sticker System over Monoid. *Regional Annual Fundamental Science Symposium*. 2011. (to appear)

[8] Hamzah, N. Z. A., MohdSebry, N. A., Fong, W. H., Sarmin, N. H., and Turaev, S., Splicing Systems over Some Permutation Groups. *Regional Annual Fundamental Science Symposium 2011* (RAFSS 2011). 20-21st December 2011: Thistle Hotel, Johor Bahru. (to appear).

[9] Kari, L., Paun, G. Rozenberg, G. Salomaa, A.and Yu, S. DNA Computing, Sticker Systems and Universality.*ActaInformatica*. 1998. 35:401 – 420.

[10] Linz, P. *An Introduction to Formal Languages and Automata*.Jones and Bartlett Publishers. 2006.

[11] Salomaa, A. *Probabilistic and weighted grammars*. Information and Control 15 (1969) 529–544.

[12] Mizumoto, M., Toyoda, J., Tanaka, K. Examples of Formal Grammars With Weights. *Information Processing Letter*.1973. 2: 74 – 78.

[13] Alexandrakis, A., Bozapalidis, S. *Weighted grammars and kleene's theorem*. Information Processing Letters 24 (1987) 1 – 4.