IGCESH2018
Universiti Teknologi Malaysia, Johor Bahru, Malaysia 13 -15 August 2018

STATIC WATSON-CRICK LINEAR GRAMMAR

Aqilahfarhana Abdul Rahman*[1], Wan Heng Fong[2], Nor Haniza Sarmin[3], Sherzod Turaev[4] and Nurul Liyana Mohamad Zulkufli[5]

[1, 2, 3] Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, Johor Bahru, MALAYSIA.
(E-mail: aqilahfarhana_13@yahoo.com, fwh@utm.my, nhs@utm.my)
[4, 5] Department of Computer Science, Faculty of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur, MALAYSIA.
(E-mail: sherzod@iium.edu.my, liyanazulkufli@iium.edu.my)

ABSTRACT

DNA computing, or more generally, molecular computing, is a recent development on computations using biological molecules, instead of the traditional silicon-chips. Some computational models which are based on different operations of DNA molecules have been developed by using the concept of formal language theory. The operations of DNA molecules inspire various types of formal language tools which include sticker systems, grammars and automata. Recently, the grammar counterparts of Watson-Crick automata known as Watson-Crick grammars which consist of regular, linear and context-free grammars, are defined as grammar models that generate double-stranded strings using the important feature of Watson-Crick complementarity rule. In this research, a new variant of static Watson-Crick linear grammar is introduced as an extension of static Watson-Crick regular grammar. A static Watson-Crick linear grammar is a grammar counterpart of sticker system that generates the double-stranded strings and uses rule as in linear grammar. Here, static Watson-Crick grammar differs from Watson-Crick grammar in term of the way each stranded string generates the double-stranded strings. The main result of the paper is to determine some computational properties of static Watson-Crick linear grammars. Next, the hierarchy between static Watson-Crick languages, Watson-Crick languages, Chomsky languages and families of languages generated by sticker systems are presented.

Key words: Sticker system, Computational power, Watson-Crick grammar, Linear grammar

INTRODUCTION

In DNA computing techniques, there are two fundamental features which are necessary to overcome the limitation of traditional silicon-based computing technologies known as Watson-Crick complementarity and massive parallelism of DNA strands. Historically, some grammar models that were introduced did not use the fundamental feature of Watson-Crick complementarity of DNA molecules [1-3]. Following that, a broad variety of dynamic Watson-Crick grammars that use this fundamental feature have been proposed [4-7]. This model produces each stranded string "independently" and do not fully illustrate the synthesis of DNA molecules. In order to overcome the problem, the aim of this research is to introduce a new variant of static Watson-Crick grammar known as a static Watson-Crick linear grammar, which is an extension of static Watson-Crick regular grammar. A static Watson-

Crick linear grammar is a grammar counterpart of sticker system which uses Watson-Crick complementarity feature of DNA molecules; starting from the incomplete double-stranded sequence and iteratively using sticking operation until complete double-stranded sequence is obtained.

In this research, some computational properties of static Watson-Crick linear grammars are determined. The result shows that the family of static Watson-Crick linear grammar can generate some non context-free languages. Besides, the family of arbitrary sticker languages with no restriction is included in the family of static Watson-Crick linear grammars. On the other hand, the result also shows that the family of Watson-Crick linear languages is included in the family of static Watson-Crick linear languages.

MAIN RESULTS

In this section, the definition of a static Watson-Crick linear grammar, which is an extension of static Watson-Crick regular grammar, is introduced in the following.

Definition 1 (Static Watson-Crick Linear Grammar)
A *static Watson-Crick linear grammar* is a 5-tuple $G=(N,T,\rho,S,P)$ where N,T are disjoint alphabets of nonterminal and terminal, respectively, $\rho \in T \times T$ is a symmetric relation (Watson-Crick complementarity), $S \in N$ is the start symbol (axiom) and P is a finite set of production rules in the form of

(i) $S \rightarrow [u_1/v_1](x_1/y_1)A(x_2/y_2)[u_2/v_2]$ where $A \in N-\{S\}$, $[u_1/v_1](x_1/y_1) \in R_\rho(T)$ and $(x_2/y_2)[u_2/v_2] \in L_\rho(T)$;

(ii) $A \rightarrow (x_1/y_1)B(x_2/y_2)$ where $A,B \in N-\{S\}$ and $(x_1/y_1),(x_2/y_2) \in LR_\rho{}^*(T)$; or

(iii) $A \rightarrow (x_1/y_1)B(x_2/y_2)$ where $A \in N-\{S\}$ and $(x_1/y_1) \in LR_\rho{}^*(T)$.

The family of languages generated by a static Watson-Crick linear grammar is denoted as **SLIN**. The elements $[u/v]$ in the set of all pairs of strings $T \times T$ can be classified into two different cases, either in the form of $[u/v] \neq [\lambda/\lambda]$ or $[u/v]=[\lambda/\lambda]$. Next, the relationship between the families of static Watson-Crick linear grammar with families in Chomsky hierarchy, sticker system and also Watson-Crick grammar are determined. The following lemma follows immediately from the definition above, where **LIN** indicates the set of languages generated by linear grammars.

Lemma 1. **LIN**$\subseteq$**SLIN**.
From Lemma 1, we have Theorem 1.

Theorem 1. **LIN**$\subset$**SLIN**.
Next, Theorem 2 shows that static Watson-Crick linear grammars can generate some non-context free languages.

Theorem 2. **SLIN** $-$ **CF**$\neq \varnothing$.
In [4], the relationship between Watson-Crick linear grammar (**WKLIN)** and the family of arbitrary sticker languages with no restriction **ASL (*n*)** has been established, where **ASL (*n*)** $\subseteq$**WKLIN**. Here, we show that **ASL (*n*)** can be simulated by **SLIN** in the following proposition.

Proposition 1. **ASL (n) $\subseteq$ SLIN**.

Next, the results show that the family of Watson-Crick linear languages is included in the family of static Watson-Crick linear languages.

Lemma 2. **WKLIN$\subseteq$SLIN**.

CONCLUSION

In this paper, we define a static Watson-Crick linear grammar, which is one of the variants of static Watson-Crick grammars and determine its computational power in order to correlate with other family of languages. This research can be further studied by defining other variants of grammar which is useful for DNA based computing devices and algorithmic techniques.

Acknowledgment: The first author would like to thank UTM Zamalah for funding her study at UTM through Zamalah Scholarship. The second and third authors would also like to thank the Ministry of Higher Education (MOHE) and Research Management Centre (RMC), Universiti Teknologi Malaysia (UTM) for the financial funding through Research University Grant Vote No. 13H18.

REFERENCES

1. Varun, R.K. and Gupta, S. Analyzing the ambiguity in RNA structure using probabilistic approach. *International Journal of Information Technology* 5(1) (2012), 107-110.
2. Sutapa, D. and Mukhopadhyay, S. A Composite Method Based on Formal Grammar and DNA Structural Features in Detecting Human Polymerase II Promoter Region. *PLoS One* 8(2) (2013), e54843.
3. Algwaiz, A., Ammar, R., & Rajasekaran, S. Framework for Data Mining of Big Data Using Probabilistic Grammars. In *e-Learning (econf), 2015 Fifth International Conference on*. IEEE, (2015), 241-246.
4. Zulkufli, N.L.M., Turaev, S., Tamrin, M.I.M. and Al-Shaikhli, I.F. Computational Properties of Watson-Crick Context-free Grammars. *In: Proceedings of International Conference on Advanced Computer Science Applications and Technologies – ACSAT* (2015), 186-191.
5. Zulkufli, N.L.M., Turaev, S., Tamrin, M.I.M and Azeddine, M. Closure Properties of Watson-Crick Grammars. In *AIP Conference Proceedings*, AIP Publishing (2015), 040032.
6. Zulkufli, N.L.M., Turaev, S., Tamsin, M.I.M. and Messikh, A. Generative Power and Closure Properties of Watson-Crick Grammars. *Applied Computational Intelligence and Soft Computing* 2016 (2016).
7. Zulkufli, N.L.M., Turaev, S., Tamsin, M.I.M. and Messikh, A. The Computational Power of Watson-Crick Grammars: Revisited. In *Bio-Inspired Computing-Theories and Applications,* Springer, Singapore (2016), 215-225.