Science Proceedings Series (SPS)
www.readersinsight.net/SPS

COMPUTATIONAL POWER OF STATIC WATSON-CRICK CONTEXT-FREE GRAMMARS

Wan Heng Fong*

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

Aqilahfarhana Abdul Rahman

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

Nor Haniza Sarmin

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

Sherzod Turaev

Faculty of Engineering and Natural Sciences,
International University of Sarajevo, Hrasnicka cesta 15, 71210, Ilidza, Sarajevo,
Bosnia and Herzegovina

***Corrosponding author's Email:** fwh@utm.my

Peer-review under responsibility of 3rd Asia International Multidisciplanry Conference 2019 editorial board (http://www.utm.my/asia/our-team/)
© 2019 Published by Readers Insight Publisher,
lat 306 Savoy Residencia, Block 3 F11/1,44000 Islamabad. Pakistan,
info@readersinsight.net

Science Proceedings Series (SPS)
www.readersinsight.net/SPS

RESEARCH HIGHLIGHTS

Sticker system is a computer model which is coded with single and double-stranded molecules of DNA; meanwhile Watson-Crick automata is the automata counterpart of sticker system representing the biological properties of DNA. Both are the modelings of DNA molecules in DNA computing which use the feature of Watson-Crick complementarity. Formerly, Watson-Crick grammars which are classified into three classes have been introduced [1]. In this research, a grammar counterpart of sticker systems that uses the rule as in context-free grammar is introduced, known as a static Watson-Crick context-free grammar. The research finding on the computational power of these grammar shows that the family of context-free languages is strictly included in the family of static Watson-Crick context-free languages; the static Watson-Crick context-free grammars can generate non context-free languages; the family of Watson-Crick context-free languages is included in the family of static Watson-Crick context-free languages which are presented in terms of their hierarchy.

Keywords: DNA Computing, Watson-Crick Grammar, Context-Free Grammar, Sticker Systems, Computational Power

RESEARCH OBJECTIVES

DNA computing is a branch of bio-molecular computing that issues the use of DNA as a data carrier. Sticker systems [2] and Watson-Crick automata [3] are DNA computing models based on distinctive standards, however the complementarity relation does exist in a computation or induction step. These computational models have been developed by using the concept of formal language theory. Previous researchers have shown that from the computational viewpoint, formal language can be used to study on molecular processes [4, 5]. However, the grammar models only work with single-stranded strings. In 2012, the usage of Watson-Crick complementarity of DNA molecules has been implemented in Watson-Crick regular grammars [6]. Following that, the research has been extensively studied with Watson-Crick grammars introduced in [1] but this model produces every stranded string independently and a complete double-stranded string can only be checked at the end of the computation. Therefore, this research aims to introduce a new variant of Watson-Crick grammars known as static Watson-Crick context-free grammars where the concept of Watson-Crick grammars and sticker system are implemented; and to analyse their computational powers which are shown in terms of the hierarchy.

MATERIALS AND METHODS

In this research, the static Watson-Crick context-free grammars are introduced by modifying Watson-Crick context-free grammars and by using the concept of sticker systems. In addtition, the computational power related to static Watson-Crick context-free grammars are obtained by classification through the relationship between the Chomsky languages and Watson-Crick languages. The definitions of sticker system, Watson-Crick grammars and Watson-Crick Chomsky normal form are given in the following.

Definition 1 [7] A sticker system is a construct $\gamma = (V, \rho, A, D)$ where V is an alphabet, $\rho \subseteq V \times V$ is a symmetric relation, A is a finite subset of $LR_\rho(V)$ (called *axioms*) and D is a finite subset of $W_\rho(V) \times W_\rho(V)$ (called *dominoes*).

Definition 2 [1] A Watson-Crick (WK) grammar $G = (N, T, \rho, S, P)$ is called

- *regular* if each production has the form $A \rightarrow \langle u/v \rangle$ where $A, B \in N$ and $\langle u/v \rangle \in \langle T^*/T^* \rangle$,

Science Proceedings Series (SPS)
www.readersinsight.net/SPS

- *linear* if each production has the form $A \to \langle u_1/v_1\rangle B\langle u_2/v_2\rangle$ or $A \to \langle u/v\rangle$ where $A, B \in N$ and $\langle u/v\rangle, \langle u_1/v_1\rangle, \langle u_2/v_2\rangle \in \langle T^*/T^*\rangle$,
- *context-free* if each production has the form $A \to \alpha$ where $A \in N$ and $\alpha \in (N \cup \langle T^*/T^*\rangle)^*$.

Definition 3 [8] A Watson-Crick (WK) context-free grammar $G = (N, T, \rho, S, P)$ is said to be in Watson-Crick Chomsky normal form if all productions are of the form

- $A \to BC$,
- $A \to \langle u/v\rangle$, or
- $S \to \langle \lambda/\lambda\rangle$,

where $A \in N$, $B, C \in N - \{S\}$ and $\langle u/v\rangle \in \langle T/\lambda\rangle \cup \langle \lambda/T\rangle$.

RESULTS

The definition of static Watson-Crick context-free grammar, a grammar counterpart of the sticker system that has rules as in context-free grammars, is introduced.

Definition 4. A static Watson-Crick context-free grammar is a 5-tuple $G = (N, T, \rho, S, P)$ where N, T are disjoint nonterminal and terminal alphabets, respectively, $\rho \in T \times T$ is a symmetric relation (Watson-Crick complementarity), $S \in N$ is a start symbol (axiom) and P is a finite set of production rules in the form of

(i) $S \to x_1A_1x_2A_2 \cdots x_kA_kx_{k+1}$ where $A_i \in N - \{S\}$ for $1 \le i \le k$, $x_1 \in R_\rho(T)$, $x_i \in LR_\rho{}^+(T)$ for $2 \le i \le k$ and $x_{k+1} \in L_\rho(T)$;

(ii) $A \to y_1B_1y_2B_2 \cdots y_tB_ty_{t+1}$ where $A, B_i \in N - \{S\}$ for $1 \le i \le t$, $y_i \in LR_\rho{}^+(T)$ for $2 \le i \le t$; or

(iii) $A \to x$ where $A \in N - \{S\}$ and $x \in LR_\rho{}^*(T)$.

The elements $\begin{bmatrix} u \\ v \end{bmatrix}$ in the set of all pairs of strings $T \times T$ can be classified into two cases, whether in the form of $\begin{bmatrix} u \\ v \end{bmatrix} \neq \begin{bmatrix} \lambda \\ \lambda \end{bmatrix}$ or $\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \lambda \\ \lambda \end{bmatrix}$. The reflexive and transitive closure of $\underset{G}{\Rightarrow}$ or $(\Rightarrow)$ is denoted by $\underset{G}{\Rightarrow}^*$ or $(\Rightarrow^*)$. The language generated by a static WK context-free grammar G, denoted by $L(G)$, is defined as $L(G) = \left\{u : \begin{bmatrix} u \\ v \end{bmatrix} \in WK_\rho(T) \text{ and } S \underset{G}{\Rightarrow}^* \begin{bmatrix} u \\ v \end{bmatrix}\right\}$. The family of languages generated by static WK context-free grammar is denoted by **SCF.**

The main results are given in the following.

Lemma 1. **CF ⊆ SCF**.

Lemma 2. **SREG ⊆ SLIN ⊆ SCF**.

Lemma 3. **SCF – CF ≠ ∅**

Lemma 4. **WKCF ⊆ SCF**.

FINDINGS

In this research, a new theoretical model known as the static Watson-Crick context-free grammar is defined along with some of the computational powers. Referring to the results, the computational power of the static Watson-Crick context-free grammar is obtained through the relationship between the families of Chomsky languages and Watson-Crick context-free languages. The findings show that the family of context-free languages is strictly included in the family of static Watson-Crick context-free languages; the static Watson-Crick context-free grammars can generate non context-free languages and the family of Watson-Crick context-free languages is included in the family of static Watson-Crick context-free languages.

Science Proceedings Series (SPS)
www.readersinsight.net/SPS

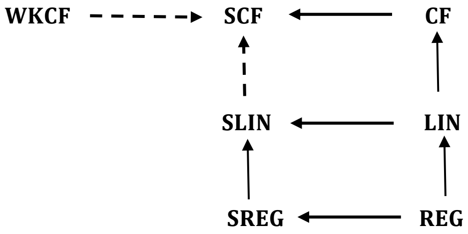

Fig. 1. The hierarchy of static Watson-Crick, Watson-Crick and Chomsky languages

ACKNOWLEDGEMENT

This research work is supported by Ministry of Education (MOE) and Research Management Centre (RMC), Universiti Teknologi Malaysia (UTM) through Research University Grant Vote No. 5F022 and UTM under Zamalah Scholarship.

REFERENCES

1. Zulkufli NLM, Turaev S, Tamrin M.I, Azeddine M. Closure properties of Watson-Crick grammars. In AIP Conference Proceedings 2015 Dec 11 (Vol. 1691, No. 1, p. 040032). AIP Publishing.
2. Kari L, Păun G, Rozenberg G, Salomaa A, Yu S. DNA computing, sticker systems, and universality. Acta Informatica. 1998 May 1; 35(5):401-20.
3. Freund R. Watson-Crick finite automata. In Proc. Third Annual DIMACS Symp. on DNA Based Computers, Philadelphia 1997.
4. Datta S, Mukhopadhyay S. A composite method based on formal grammar and DNA structural features in detecting human polymerase II promoter region. PloS one. 2013 Feb 20;8(2):e54843.
5. Algwaiz A, Ammar R, Rajasekaran S. Framework for Data Mining of Big Data Using Probabilistic Grammars. In 2015 Fifth International Conference on e-Learning (econf) 2015 Oct 18 (pp. 241-246). IEEE.
6. Subramanian KG, Hemalatha S, Venkat I. On Watson-Crick automata. In Proceedings of the Second International Conference on Computational Science, Engineering and Information Technology 2012 Oct 26 (pp. 151-156). ACM.
7. Paun G, Rozenberg G, Salomaa A. DNA computing: new computing paradigms. Springer Science & Business Media; 2005 Feb 4.
8. Mohamad Zulkufli NL, Turaev S, Mohd Tamrin MI, Messikh A. Watson–Crick Context-Free Grammars: Grammar Simplifications and a Parsing Algorithm. The Computer Journal. 2018 Jan 10;61(9):1361-73.