AIP Conference Proceedings

The concepts of persistent and permanent in non semi-simple DNA splicing system

Yuhani Yusof, Nor Haniza Sarmin, and Fong Wan Heng

Citation: AIP Conference Proceedings **1605**, 586 (2014); doi: 10.1063/1.4887654
View online: http://dx.doi.org/10.1063/1.4887654
View Table of Contents: http://scitation.aip.org/content/aip/proceeding/aipcp/1605?ver=pdfcov
Published by the AIP Publishing

The Concepts of Persistent and Permanent in Non Semi-Simple DNA Splicing System

Yuhani Yusof[a], Nor Haniza Sarmin[b] and Fong Wan Heng[c]

[a]*Faculty of Industrial Sciences & Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang Darul Makmur.*
[b]*Department of Mathematical Sciences, Faculty of Science,*
[c]*Ibnu Sina Institute for Fundamental Science Studies,*
Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor.

Abstract. The investigation on the behavior of deoxyribonucleic acid (DNA) splicing languages has been of interest of many biologists and mathematicians. Yusof-Goode (Y-G) splicing system has been introduced for the purpose of showing the transparent biological process of DNA splicing systems. In this paper, the approach of Y-G splicing system is applied in presenting the persistency and permanent characteristics of non semi-simple DNA splicing system of Type I and Type II.

Keywords: Yusof-Goode (Y-G) splicing system, persistent, permanent.
PACS: 87.14.gk, 87.14.ej

INTRODUCTION

Splicing system was developed by Head [1] as the generative capacity of systems of restriction enzymes acting on double stranded deoxyribonucleic acid (dsDNA) molecules via Formal Language Theory. In this pioneer paper of 1987, the definition and example of persistent has been introduced while the concept of permanent is first introduced by Gatterdam in [2]. In this paper, a new formulation of splicing system namely Yusof-Goode (Y-G) splicing system presented in [3] is used and the non semi-simple splicing system is introduced.

The continuity of permanent and persistent splicing systems has been studied by Fong, Sarmin and Norddin in [4]. This paper focuses on the equivalence between persistent splicing language and strictly locally testable language. Besides, this paper also presents an example of permanent splicing system in constructing finite state automaton. In [5], Karimi *et al.* present some sufficient conditions of splicing system to be persistent, meanwhile in [6] two new definitions are introduced namely, crossing preserved and self-closed splicing system. A theorem and some examples are given in illustrating the relation of both concepts with persistent splicing system.

Through this paper, the concepts of persistent and permanent splicing systems are applied to non semi-simple splicing system focusing on two rules called of Type I and Type II in expressing their characteristics based on bio molecular operations.

In the next section, the related definitions will be given.

PRELIMINARIES

In this paper, the scope of research is bounded to non semi-simple splicing system restricted to two rules since there does not exist a rule having a single letter as a crossing site. Besides, two rules are chosen in order to optimize the splicing system process. Hence, Y-G and non semi-simple splicing systems are defined presented as Definition 1 and 2 respectively, while Definition 3 and 4 present the persistent and permanent concepts.

Let A be defined as a fixed finite set to be used as an alphabet and A^* as a free monoid that consists of all strings of symbols in A, including the null string.

Definition 1 [3]: **Yusof-Goode (Y-G) Splicing System**
If $r \in R$, where $r = (u, x, v : y, x, z)$ and $s_1 = \alpha uxv\beta$ and $s_2 = \gamma yxz\delta$ are elements of I, then splicing s_1 and s_2 using r produces the initial string I together with $\alpha uxz\delta$ and $\gamma yxv\beta$, presented in either order where

Proceedings of the 21st National Symposium on Mathematical Sciences (SKSM21)
AIP Conf. Proc. 1605, 586-590 (2014); doi: 10.1063/1.4887654

$\alpha, \beta, \gamma, \delta, u, x, v, y$ and $z \in A^*$ are the free monoid generated by A with the concatenation operation and 1 as the identity element. □

Since Y-G approach has been chosen as a medium on presenting the characteristics through this paper, the amended rule of semi-simple splicing system is defined as $R = \{(a,1,1:b,1,1) \mid a,b \in A\}$.

Definition 2 [3]: **Non Semi-Simple Splicing System**
If a Y-G splicing system $S = (A, I, R)$ is not in the form of semi-simple splicing system, that splicing system is called a **non semi-simple splicing system**. □

Definition 3 [1]: **Persistent**
Let $S = (A, I, B, C)$ be a splicing system. Then S is **persistent** if for each pair of strings $ucxdv$ and $pexfq$ in A^* with (c,x,d) and (e,x,f) patterns of the same hand: If y is a sub segment of ucx (respectively xfq) that is the crossing of a site in $ucxdv$ (respectively $pexfq$) –then this same sub segment y of $ucxfq$ contains an occurrence of the crossing of a site in $ucxfq$. □

Next, the permanent concept that has been introduced by Gatterdam is stated.

Definition 4 [2]: **Permanent**
A pair of left and right hand pattern sets B, C is **permanent** if for each pair of strings $uaxbv$, $wcxdz$ in A^* with (a,x,b) and (c,x,d) patterns of the same hand: If y is a sub segment of uax (respectively xdz) that is a crossing of a site in $uaxbv$ (respectively $wcxdz$) then the same sub segment y of $uaxdz$ is a crossing of a site in $uaxdz$ □

In the next section, some characteristics of the non semi-simple deoxyribonucleic acid (DNA) splicing system are given presented as theorem and corollaries.

SOME CHARACTERISTICS OF NON SEMI-SIMPLE DNA SPLICING SYSTEM

In [1], it is stated that all splicing systems that consist of one rule and a null-context splicing system are always persistent. From Definition 4, it is easy to show that both of splicing systems are also permanent since the crossing site is also the crossing of a site in the obtaining string. Focusing on persistent and permanent concepts, the behaviour of the non semi-simple DNA splicing system will be presented as theorems and corollaries. The following two corollaries hold for a non semi-simple DNA splicing system of Type I (Y-G splicing system with one rule) and a null-context of Y-G splicing system due to the element of x in R and it is supported in [3] that there is no apparent change in generative power translated from Goode- Pixton or Head notation to Y-G notation**.**

Corollary 1[3]
A non semi-simple splicing system $S = (A, I, R)$ of Y-G splicing system with one rule (Type I) is always persistent and permanent. ■

Corollary 2 [3]
A null-context of Y-G splicing system $S = (A, I, R)$ is always permanent. ■

In Theorem 1 below shown that the different crossing site of two existing rules in non semi-simple of Y-G splicing system is persistent.

Theorem 1
A non semi-simple splicing system $S = (A, I, R)$ of Y-G splicing system with two rules (Type II) with disjoint crossing site is always persistent. □

Proof
Assume $S = (A, I, R)$ is a non semi-simple splicing system of Type II with different crossing. Hence, the rule $r \in R$ holds $(a_{11}, a_{21}, a_{31} : a_{11}, a_{21}, a_{31})$ and $(a_{12}, a_{22}, a_{32} : a_{12}, a_{22}, a_{32})$ as their form of rule $\forall a_{11}, a_{21}, a_{31}, a_{12}, a_{22}, a_{32} \in A^*$, $\ni a_{21} \neq a_{22}$ and $\forall a_{ij}, i = 1,2,3, j = 1,2$ fulfilled the following conditions:

i. $a_{1j} \notin A^* \setminus \{a,g,c,t\}$, if $a_{2j} = a_{3j} = 1, j = 1,2,$

ii. $a_{2j} \notin A^* \setminus \{a,g,c,t\}$, if $a_{1j} = a_{3j} = 1, j = 1,2,$

iii. $a_{3j} \notin A^* \setminus \{a,g,c,t\}$, if $a_{1j} = a_{2j} = 1, j = 1,2.$

In addition, any rule $r \in R$ is not a combination of the rules above. Since both rules have different crossing site, therefore, two cases are involved:

Case I: sub segment $y = a_{21}$.

Let $ua_{11}a_{21}a_{31}v$ and $pa_{11}a_{21}a_{31}q$ be strings in A^*. By taking a_{21} as a sub segment $ua_{11}a_{21}$ (respectively $a_{21}a_{31}q$), that is crossing of $ua_{11}a_{21}a_{31}v$ (respectively $pa_{11}a_{21}a_{31}q$). Hence this a_{21} also contains an occurrence of the crossing of a site in $ua_{11}a_{21}a_{31}q$.

Case II: sub segment $y = a_{22}$.

Let $ua_{12}a_{22}a_{32}v$ and $pa_{12}a_{22}a_{32}q$ be strings in A^*. By taking a_{22} as a sub segment $ua_{12}a_{22}$ (respectively $a_{22}a_{32}q$), that is crossing of $ua_{12}a_{22}a_{32}v$ (respectively $pa_{12}a_{22}a_{32}q$). Hence this a_{22} also contains an occurrence of the crossing of a site in $ua_{12}a_{22}a_{32}q$. By both cases, thus S is persistent. ■

Notice that in Theorem 1, the crossing site in $ua_{11}a_{21}a_{31}q$ and $ua_{12}a_{22}a_{32}q$ are equal with the sub segment taken for each both cases, hence produce the following corollary.

Corollary 3

A non semi-simple splicing system $S = (A, I, R)$ of Y-G splicing system with two rules (Type II) with different crossing site is always permanent. ■

In the next theorem, the persistency of non semi-simple splicing system Type II consisting of different pattern of rule is proved.

Theorem 2

A non semi-simple splicing system $S = (A, I, R)$ of Y-G splicing system with two rules (Type II) with different pattern is always persistent. □

Proof

Assume $S = (A, I, R)$ is a non semi-simple splicing system of Type II with different pattern. Thus, the rule $r \in R$ having $(a_{11}; a_{21}, a_{31} : a_{11}; a_{21}, a_{31})$ and $(a_{12}, a_{22}; a_{32} : a_{12}, a_{22}; a_{32})$ as its form of rule for which $a_{11}, a_{21}, a_{31}, a_{12}, a_{22}, a_{32} \in A^*$ and $\forall a_{ij}, i = 1,2,3, j = 1,2$ fulfilled the following conditions:

i. $a_{1j} \notin A^* \setminus \{a,g,c,t\}$, if $a_{2j} = a_{3j} = 1, j = 1,2,$

ii. $a_{2j} \notin A^* \setminus \{a,g,c,t\}$, if $a_{1j} = a_{3j} = 1, j = 1,2,$

iii. $a_{3j} \notin A^* \setminus \{a,g,c,t\}$, if $a_{1j} = a_{2j} = 1, j = 1,2.$

Furthermore, any rule $r \in R$ is not a combination of the rules above. Hence, two cases are considered:

Case I: r in the form of $(a_{11}; a_{21}, a_{31} : a_{11}; a_{21}, a_{31})$

Case II: r in the form of $(a_{12}, a_{22}; a_{32} : a_{12}, a_{22}; a_{32})$

Since both rules are in different pattern, thus Case I and Case II can be independently categorized as non semi-simple splicing system of Type I which lead to the persistency of S. ■

Since in Theorem 2 the splicing system S is associated with non semi-simple splicing system of Type I, thus Corollary 4 is proven.

Corollary 4

A non semi-simple splicing system $S = (A, I, R)$ of Y-G splicing system with two rules (Type II) with different pattern is always permanent. ■

In this last theorem, a non semi-simple splicing system with same crossing site is discussed.

Theorem 3

A non semi-simple splicing system $S=(A,I,R)$ of Y-G splicing system with two rules (Type II) with identical crossing site and one context is always persistent. □

Proof

Suppose $S=(A,I,R)$ is a non semi-simple splicing system of Type II with identical crossing site of one context. Therefore, the rule $r \in R$ is in the form of $(a_{1j}, a_{21}, a_{3j} : a_{1j}, a_{21}, a_{3j}),\ j=1,2$ where $\forall a_{1j}, a_{21}, a_{3j} \in A^*$ either $a_{12}=a_{11}$ or $a_{32}=a_{31}$. In addition, $\forall a_{ij}, i=1,2,3, j=1$, a_{12} and a_{32} fulfilled the following conditions:

i. $a_{1j} \notin A^* \setminus \{a,g,c,t\}$, if $a_{2j}=a_{3j}=1, j=1,2$,

ii. $a_{21} \notin A^* \setminus \{a,g,c,t\}$, if $a_{1j}=a_{3j}=1, j=1,2$,

iii. $a_{3j} \notin A^* \setminus \{a,g,c,t\}$, if $a_{1j}=a_{2j}=1, j=1,2$.

Moreover, any rule $r \in R$ is not a combination of the rules above. Let $ua_{11}a_{21}a_{31}v$ and $pa_{12}a_{21}a_{32}q$ be strings in A^*. Two cases are considered:

Case 1: equality in left context of the rule, $a_{12}=a_{11}$

By taking a_{21} as a sub segment $pa_{12}a_{21}$ (respectively $a_{21}a_{31}v$), that is crossing of $pa_{12}a_{21}a_{32}q$ (respectively $ua_{11}a_{21}a_{31}v$). Hence this a_{21} also contains an occurrence of the crossing of a site in $pa_{12}a_{21}a_{31}v$ since $a_{12}=a_{11}$.

Case 2: equality in right context of the rule, $a_{32}=a_{31}$

By taking a_{21} as a sub segment $ua_{11}a_{21}$ (respectively $a_{21}a_{32}q$), that is crossing of $ua_{11}a_{21}a_{31}v$ (respectively $pa_{12}a_{21}a_{32}q$). Hence this a_{21} also contains an occurrence of the crossing of a site in $ua_{11}a_{21}a_{32}q$ since $a_{32}=a_{31}$. Thus S is persistent. ■

Note that a_{21} is also a crossing for both $pa_{12}a_{21}a_{31}v$ and $ua_{11}a_{21}a_{32}q$, thus this theorem contributes to the next corollary.

Corollary 5

A non semi-simple splicing system $S=(A,I,R)$ of Y-G splicing system with two rules (Type II) with identical crossing site and one context is always permanent. ■

Not all non semi-simple spicing system of Types II is persistent or permanent. This clause is proved by the following biology-based counterexample.

Example 1

Let $S=(A,I,R)$ be a Y-G splicing system that consists of two restriction enzymes, namely *Alu*I and *Bst*UI which are represented by $r=(ag;1,ct:cg;1,cg)$. Let *aagctt* and *ccgcgc* be two initial strings in I. By taking 1 for the first and second string as a sub segment *aag* (respectively, *cgc*), that is the crossing of *aagctt* (respectively, *ccgcgc*). Thus, this splicing system S is not persistent since the crossing of the yield string *aagcgc* (respectively, *ccgcct*) is not an element of 1. ■

CONCLUSION

As a conclusion, the concepts of persistent and permanent splicing system are theoretically explored, viewing in perspective of crossing site and pattern of the rule. All these characteristics can be summarized as in Table (1).

TABLE (1). Some Characteristics of Non Semi-Simple DNA Splicing System

Type	Rule Condition(s)	Characteristics of DNA Splicing System
Type I	-	Persistent and Permanent
Type II	Disjoint Crossing Site	Persistent and Permanent
Type II	Different Pattern	Persistent and Permanent
Type II	Identical Crossing Site and One Context	Persistent and Permanent

ACKNOWLEDGMENTS

The first author would like to acknowledge Ministry of Education (MOE) of Malaysia and Research and Innovation Department, Universiti Malaysia PAHANG (UMP) for the financial funding through UMP Research Grant Vote No: RDU 130354.

REFERENCES

1. T. Head, *Bulletin of Mathematical Biology* **49**, 737–759 (1987).
2. R. W. Gatterdam, *International Journal of Computer Math.* **31**, 63-67 (1989).
3. Y. Yusof, “DNA Splicing System Inspired by Bio Molecular Operations”, Ph.D. Thesis, Universiti Teknologi Malaysia, 2012.
3. W. H. Fong, N. H. Sarmin and N. I. Norddin, Formal Language Theory and Its Application in Splicing System, in *Proc. International Conference on Research and Education in Mathematics (ICREM2),* (Selangor, Malaysia 2005).
5. F. Karimi, N. H. Sarmin and W. H. Fong, *Australian Journal of Basic and Applied Sciences* **5(1)**, 20-24 (2011).
6. F. Karimi, N. H. Sarmin and W. H. Fong, Crossing Preserved and Persistent Splicing System, in *Proc*. *The Sixth International Conference on Bio-Inspired Computing: (BICTA 2011)*, (Penang, Malaysia 2011).