

A New Relation of Second Order Limit Language in Simple and Semi-Simple Splicing System

Muhammad Azrin Ahmad^a, Nor Haniza Sarmin^{a*}, Wan Heng Fong^b, Yuhani Yusof^c

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

^bIbnu Sina Institute for Fundamental Science Studies, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor Malaysia

^cFaculty of Industrial Science and Technology, Universiti Malaysia Pahang, 26300 UMP Gambang, Pahang

*Corresponding author: nhs@utm.my

Article history

Received : 31 July 2014

Received in revised form :

23 November 2014

Accepted : 1 December 2014

Graphical abstract

$$L_{2, \text{left pattern, right pattern}}(S) = L(S) = \{ab(ab)^k a\beta, \beta^k(ab)^k a\beta\}$$

Abstract

Splicing system, which is an abstraction of operations on DNA molecules, can be modelled mathematically under the framework of formal language theory and informational macromolecules. The recombinant behavior of the set of double-stranded DNA molecules under the influence of restriction enzyme and ligase further lead to the cut and paste phenomenon in splicing system. The theoretical study of splicing language has contributed to a new type of splicing language known as a second order limit language, which is an extension of limit language. Some types of splicing system can produce second order limit language. Y-G splicing system is chosen among other models to model the DNA splicing process as this model preserves the biological traits and presents the transparent behavior of the DNA splicing process. In this paper, the relation between second order limit language with simple splicing and semi-simple splicing system are presented.

Keywords: Y-G splicing system; Y-G splicing language; second order limit language

Abstrak

Sistem hiris-cantum yang merupakan suatu operasi pemujaradan ke atas molekul-molekul DNA boleh di modelkan secara matematik di bawah rangka kerja teori bahasa formal dan makromolekul berinformasi. Tingkah laku rekombinan molekul-molekul DNA bebanang dua di bawah pengaruh enzim pembatas dan ligase telah membawa kepada fenomena potong dan tampal di dalam sistem hiris-cantum. Kajian teori bahasa hiris-cantum telah menyumbang kepada jenis bahasa hiris-cantum yang baharu yang dikenali sebagai bahasa batas berperingkat dua yang juga merupakan lanjutan kepada bahasa batas. Beberapa jenis sistem hiris-cantum boleh menghasilkan bahasa batas berperingkat dua. Sistem hiris-cantum Y-G dipilih daripada model-model lain bagi membentuk proses hiris-cantum DNA kerana model ini dapat mengekalkan ciri-ciri biologi dan mewakili tingkah laku telus proses hiris-cantum DNA. Dalam kertas kerja ini, hubungan di antara bahasa batas berperingkat dua dengan sistem hiris-cantum mudah dan sistem hiris-cantum separa-mudah diperkenalkan.

Kata kunci: Sistem hiris-cantum Y-G; bahasa hiris-cantum Y-G; bahasa batas berperingkat dua

© 2012 Penerbit UTM Press. All rights reserved.

1.0 INTRODUCTION

Deoxyribonucleic acid (DNA) acts as a hereditary factor that passes information from parent cell to offspring cell and also as a coding for protein production in living organisms.¹ DNA is a polymer made up of many monomers called deoxyribonucleotides. These deoxyribonucleotides consist of three different components which are sugar, phosphate and nitrogenous base. Moreover, there are four types of bases namely adenine (A), guanine (G), cytosine (C) and thymine (T). In addition, the bases can be grouped as purines (A and G) and pyrimidines (C and T) accordingly. Watson-Crick complementarity² stated that the only possible pairings are A with T, C with G and vice versa. Basically the sugar component has five carbon atoms that are numbered from 1' to 5'. The

phosphate is attached to the 5' carbon while the base is attached to the 1' carbon. Meanwhile a hydroxyl group (OH) is attached to the 3' carbon.

In 1987, Head³ pioneered a mathematical modelling of splicing system. A splicing system is a study of the recombinant behavior of restriction enzymes on DNA molecules. The recognition process that determines the cutting site is depicted by the restriction enzyme. The restriction enzyme will clamp at the specific nucleotides sequence known as the crossing site on DNA molecules and the cutting process will take place. Besides that, interdisciplinary knowledge between formal language theory and the study of informational molecules is initiated. In splicing system, a splicing model is either based on the generation of language or a splicing model that preserves the biological traits in a splicing system.