An Introduction to *n*-th Order Limit Language

Siti Hajar Mohd Khairuddin, Muhammad Azrin Ahmad, and Noraziah Adzhar

***Abstract*—The study of the splicing system swiftly grew after Tom Head investigated the biochemical process modelling involving the DNA in 1987. The process of the splicing system consists of a cut and paste of the DNA molecules. Splicing language produced by the splicing system can be classified into inert, transient, and limit language. Previously, second-order limit language was described as a new set of language from the previous splicing language. In this research, we would like to extend the study to *n*-th order limit language by investigating the effect of the number of rules involved in the splicing system. Following from here, its properties are explored using the formal language theory.**

***Index Terms*—DNA, *n*-th order limit language, splicing system, splicing language,**

I. Introduction

THE deoxyribonucleic acid (DNA) is a natural compound spotted in all prokaryotic and eukaryotic cells with a complex atomic structure [1]. The DNA is a particle in charge of conveying and transmitting the inherited materials or the hereditary guidelines from the guardians to the offspring and for the generation of proteins. Hence, DNA is found in every living life form and is principally in charge of hereditary data's legacy in every living being. The cells in the human body have a similar DNA placed in the core of cells known as the nucleus, which is sometimes located in the mitochondrion [2]. The DNA structure is reflected in Figure 1 as follows.

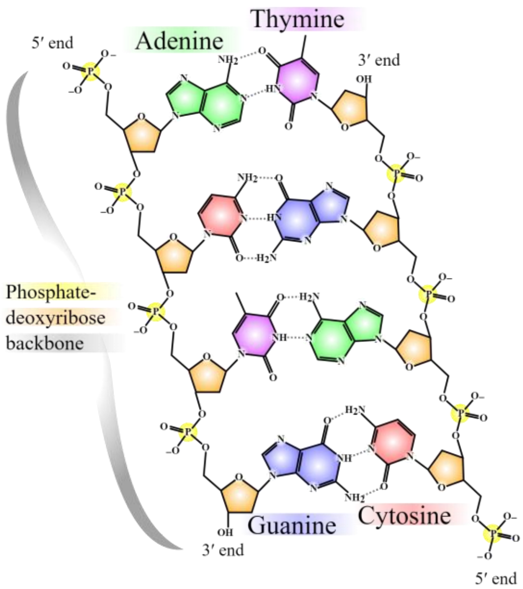

Figure 1. The Structure of the DNA

The DNA structure is a long, twofold helix that takes after a stepping stool, turned at both ends, as portrayed in Figure 1. It is known as nucleic acids, comprising nucleotides. Nucleotides contain three essential substances: a phosphate group, a sugar group, and nitrogenous bases [3], as shown in Figure 2. The sugar element essentially has five carbon atoms numbered between 5' and 3'. The phosphate is attached to the 5' carbon binding the base to the 1' carbon, while the 3' carbon is attached to a hydroxyl group (OH) [4].

The DNA mechanism follows the Watson-Crick complementary, where Adenine (A) is combined with Thymine (T) and Guanine (G) coupled with Cytosine (C), and vice versa (Watson & Crick, 1953). Therefore, it can be paired and represented as *a*, *g*, *c*, *t*, as illustrated below.

A	**G**	**C**	**T**
T	**C**	**G**	**A**

Moreover, the enzyme that divides DNA into fragments at the recognition site is called restriction enzyme or restriction endonuclease or restrictase. Restriction enzymes distinguish a specific sequence, where every restriction enzyme(s) has its targeted site, facilitating it to bind to the molecules and then splice them [5].

Manuscript received January 18, 2021; revised September 10, 2021. This work was supported in part by the Minister of Higher Education (MOHE) Malaysia for the financial funding through Fundamental Research Grant Scheme (FRGS), FGRS/1/2018/STG06/UMP/03/1.

Siti Hajar Mohd Khairuddin is a postgraduate student at the Centre for Mathematical Sciences, Universiti Malaysia Pahang, 26300 Lebuhraya Tun Razak, UMP Gambang, Pahang, Malaysia (corresponding author to provide phone: +601112914899; e-mail: sitihajarmohdkhairuddin96@gmail.com).

Muhammad Azrin Ahmad is a senior lecturer at the Centre for Mathematical Sciences, Universiti Malaysia Pahang, 26300 Lebuhraya Tun Razak, UMP Gambang, Pahang, Malaysia (e-mail: azrin@ump.edu.my).

Noraziah Adzhar is a senior lecturer at the Centre for Mathematical Sciences, Universiti Malaysia Pahang, 26300 Lebuhraya Tun Razak, UMP Gambang, Pahang, Malaysia (e-mail: noraziahadzhar@ump.edu.my).