DESIGN AND ANALYSIS OF DNA FRAGMENT ASSEMBLY'S CONVERTER AND CHECKER MODULES USING DE BRUIJIN GRAPH

MOHD SHAFIQ B MOHD HELMI

This thesis is presented in partial fulfillment for the award the
Bachelor of Electrical Engineering (Honors)
Universiti Teknologi MARA (UiTM)

JULY, 2012

ACKNOWLEDGEMENT

Praise to Allah the Almighty that has granted me a good health and allows me to complete my thesis successfully. And also, thank to my parents, sister and brothers, grateful acknowledgement because with their support and encouragement I have managed to complete my project. I also would like to express my sincere gratitude to Puan Tuan Norjihan Tuan Yaakub and Encik Abdul Karimi Halim, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding and achieving the goal of this project.

I also wish to record my sincere appreciation to the staff of Electrical Engineering Department, Faculty of Engineering, Universiti Teknologi MARA (UiTM) Malaysia, and also the Department of Electrical Engineering for giving me the opportunities and permissions to use PTAR 3 (Perpustakaan Tunku Abdul Rahman 3) as well as lend few books from it. All the books and journal really help me up to make some references in order for me to complete this thesis. A lot of valuable information and suggestion had been gained from it. Not to forget to my helpful classmates and friends who have been kind by sharing the knowledge and information from the beginning till the end of the thesis. I would like to thank everyone that have contributed directly or indirectly to the completion of this project and their assistance are gratefully acknowledge.

ABSTRACT

This project presents the Design and Analysis of DNA Fragment Assembly's Converter and Checker Modules using De Bruijin Graph. There are several objectives of this project. The first objective is to design and analyze edges and vertices converter. Secondly, this paper will also discuss about designing and analyzing common edge and branch checker. In addition, all the modules will be integrated in one top module. Lastly, the integrated module will be simulated in Xilinx ISE software. Based on the objectives, this project will discuss on constructing a DNA Fragment Assembly module that consist of four submodules which is Edges Converter, Vertices Converter, Common Edge Checker and Branch Checker. The sequencing method used in this paper is Hybridization method. Hybridization method is an option to shotgun sequencing. Hybridization uses the array identifying algorithm to identify the sequence. The algorithm used to construct the module is Bruijn Graph. De Bruijn Graph represents sequences of symbols from an alphabet, and edges that indicate where the sequence may overlap. First, the set of DNA will inserted in to the module, then the module will do all the converting and checking process of the input before produce the output of vertices and the branch of edges. The entire algorithm and theory then been simulated in Xilinx to test the functionality.

TABLE OF CONTENT

CHAPTER 1

INTRODUCTION

1.0 INTRODUCTION

This chapter will discuss on background study of the design and analysis of DNA Fragment Assembly's Converter and Checker Modules using De Bruijin Graph's project, the problem statement, objectives and scope of the project and also summarize for the whole chapter one.

1.1 BACKGROUND OF STUDY

In recent years, scientist and researchers have emphasized DNA sequencing and fragment assembly with the hopes of enhancing their abilities to reconstruct full strands of DNA based on the pieces of data they have. Complication arises with fragment assembly due to long and imperfect data.

Fragment assembly is a typical approach to sequencing long DNA molecules to sample and then sequence fragments from them. The objective of DNA Fragment assembly to obtain the DNA sequence using certain hints, which are (approximate) substrings of the row [1].