OPTIMIZING OPTIMAL PATH TRACE BACK FOR SMITH-WATERMAN ALGORITHM USING STRUCTURAL MODELLING TECHNIQUE

NUR FARAH AIN BINTI SALIMAN

Faculty of Electrical Engineering

Universiti Teknologi MARA

40450 Shah Alam, Selangor

MALAYSIA

TABLE OF CONTENTS

ABSTRACT

The optimizing of optimal path trace back system for Smith-Waterman algorithm using structural modelling techniques are presented in this paper. The objectives for this paper are to optimize the optimal path trace back scanning and to construct it using simple structural modelling architecture. Thus, this paper is highlighted on theoretical study and experimental design as a method of completion of proposed optimization techniques. Theoretical study emphasis on the mathematical and prove of concept of the proposed techniques. In experimental design, the proposed techniques code is developed, compiled and simulated using Altera Quartus II Version 9.0 EDA Tools targeted to Cyclone II EP2C35 at 50MHz clock oscillation. Both of the proposed technique is compared and analysed. The first design required more Logic Element (LE) as compare to second design which increased gradually in small interval. Therefore, the second design is the best approach in optimizing the optimal path trace back for Smith Waterman algorithm. In addition, the Design 2 is better than Design 1 since the runtime reduce up to 50%.

ACKNOWLEDGEMENT

In the name of Allah, Most Beneficent, Most Merciful. Alhamdulillah, all the praise and thanks are to His willingness and blessings that brought me to completing my Final Year Project Thesis with entitled “Optimizing Optimal Path Trace Back for Smith-Waterman Algorithm using Structural Modelling Technique”. First of all, I would like to express my grateful thanks to my supervisor, Mister Syed Abdul Mutalib Al Junid for all information and guidance from beginning until the end of my project. He was being helpful in giving suggestions, improving in the project design. Without his guidance I won’t think that I was able to complete this Final Year Project. Hopefully the information and the brilliant knowledge that I gained from this report would benefit me in the future. Lastly but not least to my beloved parents Mister Saliman bin Abu Denan and Madame Suriya binti Supardi, my friends and also to those who are involved. Your support and contribution are really appreciated. I wish for them all the best in the future. Without them this thesis would not be as it is today. Thank you and may God bless you all.

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND OF STUDY

Bioinformatics is combination between the development of computer software and biological algorithms to record and analyze the biology related database [1]. In addition, bioinformatics also called as molecular biology information. The central view of molecular biology information contains DNA sequences alignment [2]. This DNA sequences alignments can be classified into local alignment and global alignment. Firstly, global alignments will consider both sequences (search sequence and target sequence) from end to end sequences, whereas in local alignments will identify the most similar pattern sequences and only align at that subsequence. Normally, local alignment is preferred especially in searching the most similar pattern in DNA sequences. Moreover this alignment also only focussed on the subsequence that was preserved during evolution [3].