

Assignment 1:

Abstract :

The main reason for development methods in aspects of bioinformatics is the increase of biological. Pattern matching is very important operation in different computational stages. Pattern matching helps to find the places of similar subsequences in a database or in DNA sequence, it checks all database, sequence or text to determined places of the pattern in the text. The paper introduces ways for pattern matching to speed up searching for DNA sequences. The suggested algorithm helps to improve performance by using Keywords during search to save time.

Introduction :

- The paper focus on solving the problem of accurate pattern matching which find all occurrence in the text. the paper introduces three algorithms.
- The operation of suggested algorithms divided into before processing and matching phase.
- in preprocessing phase, possible time intervals for the text are recognized to be matched with the pattern, this intervals called windows.
- And after that in matching phase, windows are scanned carefully to match with the pattern. the fewer windows in the preprocessing phase, the less time it takes to check windows in the match phase.
- The first proposed algorithm aims to determined windows by discovering first and last letter in the text and look at the same time to first and last letter in the pattern.
- The second algorithm makes comparisons depending on words, the words are processed using processor power. the computational length of processors almost 32 or 64bits in each execution cycle. That is, it can process 4 or 8 bytes of data, it can compare 4 or 8 letters (indicated by a word) in same time with 4 or 8 other letters.
- The third algorithm focus on word of the pattern which has the fewest number of iterations in the text and this will reduce the number of windows detected,

which improves performance and reduce cost of time.