Research Field: Bioinformatics - Sequence Analysis


Note; although the paper is 1994 but no papers till 2008 proposed significant improvement for the Myers algorithm. It is considered the most significant paper in Bioinformatics – sequence alignment.

Key Idea

The main idea of this research paper is sequence alignment. The problem of sequence alignment means trying to identify the most significant sequence in DNA stream that can be converted to a functional protein. Not all the DNA parts is converted to a functional protein. Apparently from evolution theory, what we behold from living things are a result of surviving protein functions. Those that cannot be behelden, corresponds to living things that are exposed to death and are no longer alive.

So, sequence alignment means identifying the main, and important surviving sequence in a DNA sequence. We mean by main surviving sequence; the most functional sequence. If we can represent the DNA sequence as a sequence of strings, we need to align from these strings the sequence that corresponds to a DNA sequence that is to be converted to a protein function.

The paper presents a new algorithm that aligns sequences. The algorithm proposed allows for a rate of error. The algorithm gets a stream of strings that should be aligned on the DNA sequence. The algorithm first, converts the input string to a radix integer, as well as the DNA sequence. Then using indexing, and sorting, it can align each incoming string radix integer to the DNA radix integer by applying integer to integer comparisons, instead of string to string comparisons.

Reason of choosing this paper

The main paradox of sequence alignment is speed. Since DNA streams are actually huge, low speed algorithm is intrinsically a problem. On the other hand, the algorithm speed tends to be linear, and this is quite good for DNA sequence alignment.

What Next?

We need to enhance the algorithm to allow error in alignment process, because this can naturally happen during protein synthesis. But, we need also to develop a control over this error. We need the error to be in a rate that does not potentially affect the correctness of the algorithm.

On the other side, Bioinformatics is the intersection of biology and computer science. So, the main aim of bioinformatics research is to provide some tools that help biologists. We cannot find till now a quite good tool for sequence alignments for biologists. So, implementing this proposed algorithm to develop a sequence alignment tool will be beneficial.