

Implementation of Multiple Sequence Alignment using Genetic Algorithm

Team Members

Vivek Goyal (UFID: 8388-6402)

Adhiraj Nakhe (UFID: 9939-6580)

Debarshi Mitra (UFID: 3381-3136)

Nishant Varma (UFID: 8593-9695)

Abstract

Multiple Sequence Alignment (MSA) is a problem of alignment of three or more sequences. Generally, these sequences are of the same length. Multiple Sequence Alignment has a lot of applications like construction of phylogenetic tree, prediction of protein structure and is considered to be one of the fundamental problems in bioinformatics. Dynamic programming is the most accurate way to calculate MSA although it is impractical because it is computationally very expensive. A better solution is to use heuristic approach like genetic algorithm to solve the problem. Genetic algorithm is a way to solve optimization problems using techniques that mimic biological phenomena like selection, crossover and mutation.

Plan of Action

- Understanding MSA problem and methods to compute MSA.
- Understanding genetic algorithm.
- Extracting relevant sequences from protein databases.
- Implementation of genetic algorithm to solve MSA problem.
- Comparison of derived results with the expected results.
- Create project report and documentation of the code.

Workload Distribution

Workload will be equally distributed among the four team members.

References

- http://www.tcoffee.org/Publications/Ps_pdf/saga_paper.pdf
- https://deepblue.lib.umich.edu/bitstream/handle/2027.42/46947/10994_2005_Article_422926.pdf
- SAGA: Sequence Alignment by Genetic Algorithm, C. Notredame, D.G. Higgins, Nucleic Acid Research, Vol. 24, 1515-1524, 1996.
- https://www.researchgate.net/publication/220885709_A_Standard_GA_Approach_to_Native_Protein_Conformation_Prediction.
- "IEEE Xplore document - progressive alignment method using genetic algorithm for multiple sequence alignment," 2016. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6151111>.
- "IEEE Xplore document - multiple sequence alignment using genetic algorithm and simulated annealing," 2016. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1307828>.