**COMP90025 Parallel and Multicore Computing**

*Project 2 – Multiple Sequence Alignment with OpenMP and OpenMPI*

FirstName Surname†  
 Department Name  
 Institution/University Name  
 City State Country  
 email@email.com

Hoang Viet Mai - 813361  
School of Computing and Information Systems  
The University of Melbourne  
[*vietm@student.unimelb.edu.au*](mailto:email@email.com)

FirstName Surname  
 Department Name  
 Institution/University Name  
 City State Country  
 email@email.com

Abstract

Introduction

Sequence alignment is the task of arranging two or more sequences of symbols, or more specifically, inserting the necessary gaps to equalise the length of those sequences and maximise their similarities, that is, minimising the number of gaps inserted and symbol mismatches at a given index position. Sequence alignment is most widely utilised in Bioinformatics for DNA sequencing or phylogenetic reconstruction [1], but it also has applications in other fields such as natural language processing [2] and plagiarism detection [3]. Thanks to its variety of practical applications and computing power becomes less and less expensive, many different sequence alignment algorithms and softwares have been developed in recent years. In this paper, I will brief discuss the Smith-Waterman algorithm along with its parallel version by Tristan [4], as well as my implementation of said version using OpenMP.

Background

Related Work

Proposed Approach

Experiment Results

Conclusion

REFERENCES

[1] Chang JM, Di Tommaso P, Notredame C. TCS: a new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction. *Mol Biol Evol*. 2014 Jun;31(6):1625-37. doi: 10.1093/molbev/msu117.

[2] Zeng, Z., Shi, H., Wu, Y., & Hong, Z. (2015). Survey of Natural Language Processing Techniques in Bioinformatics. *Computational and mathematical methods in medicine*, *2015*, 674296. https://doi.org/10.1155/2015/674296

[3] Asad Abdi, Norisma Idris, Rasim M. Alguliyev, and Ramiz M. Aliguliyev. 2015. PDLK: Plagiarism detection using linguistic knowledge. *Expert Syst. Appl*. 42, 22 (2015), 8936--8946. DOI:10.1016/j.eswa.2015.07.048

[4] Tristan C., A pairwise comparison of DNA sequence alignment using an OpenMP implementation of the SWAMP parallel Smith-Waterman algorithm, 2015.

[5] Smith T. F., Waterman M. S. Identification of common molecular subsequences. Journal of Molecular Biology. 1981;147(1):195–197. doi: 10.1016/0022-2836(81)90087-5.

[6] Pearson W. R., Lipman D. J. Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences of the United States of America*. 1988;85(8):2444–2448. doi: 10.1073/pnas.85.8.2444.

[7] Lipman D. J., Pearson W. R. Rapid and sensitive protein similarity searches Science. 1985;227(4693):1435–1441. doi: 10.1126/science.2983426.

[8] Huang, L. T., Wu, C. C., Lai, L. F., & Li, Y. J. (2015). Improving the Mapping of Smith-Waterman Sequence Database Searches onto CUDA-Enabled GPUs. *BioMed research international*, *2015*, 185179.https://doi.org/10.1155/2015/185179

[9] Peiheng Zhang, Guangming Tan, and Guang R. Gao. 2007. Implementation of the Smith-Waterman algorithm on a reconfigurable supercomputing platform. In *Proceedings of the 1st international workshop on High-performance reconfigurable computing technology and applications: held in conjunction with SC07* (*HPRCTA '07*). Association for Computing Machinery, New York, NY, USA, 39–48. DOI:https://doi.org/10.1145/1328554.1328565