



DEPARTMENT OF COMPUTER SCIENCE

PhD Degree Oral Presentation

PhD Candidate:	Mr Kaiyong ZHAO
Supervisor:	Dr Xiaowen CHU
External Examiner:	Dr Qixin WANG (EEx from Hong Kong Polytechnic University) Dr Kwok Wai CHEUNG (Proxy for Dr Bingsheng HE - EEx from Singapore)
Time:	17 Aug 2016 (Wednesday) 2:30 pm - 4:30 pm (35 mins presentation and 15 mins Q & A)
Venue:	RRS732, Sir Run Run Shaw Bldg., HSH Campus, HSH Campus

“GPU Accelerated Sequence Alignment”

Abstract

DNA sequence alignment is a fundamental task in gene information processing, which is about searching the location of a string (usually based on newly collected DNA data) in the existing huge DNA sequence databases. Given the huge amount of newly generated DNA data and complexity of un-exact string match, sequence alignment becomes a time-consuming process, and hence how to reduce the alignment time becomes a significant research problem. Some algorithms of string alignment based on HASH comparison, suffix array and BWT have been proposed for DNA sequence alignment. Although these algorithms have reached the speed of $O(N)$, they still cannot meet the increasing demand if running on traditional CPUs.

Recently, GPUs have been widely accepted as an efficient accelerator for many scientific and commercial applications. A typical GPU has thousands of processing cores which can speed up repetitive computations significantly as compared to multi-core CPUs. However, sequence alignment is one kind of computation procedure with intensive data access, i.e., it is memory-bounded. The access to GPU memory and IO has more significant influence compared to the computing capabilities of GPU cores. By analyzing GPU memory and IO characteristics, this thesis produces novel parallel algorithms for DNA sequence alignment applications. This thesis consists of six parts. The first two parts explain some basic knowledge of DNA sequence alignment and GPU computing. The third part investigates the performance of data access on different types of GPU memory. The fourth part describes a parallel method to accelerate short-read sequence alignment based on BWT algorithm. The fifth part proposes the parallel algorithm for accelerating BLASTN, one of the most popular sequence alignment software. It shows how multi-threaded control and multiple GPU cards can accelerate the BLASTN algorithm significantly. The six part concludes the whole thesis.

To summarize, this thesis proposes some novel optimization methods for parallel sequence alignment on GPUs, and the outcomes can help practitioners in bioinformatics improve their working efficiency by significantly reducing the sequence alignment time.

***** ALL INTERESTED ARE WELCOME *****