Parallel Programming Optimization: Practical Applications in Bioinformatics

Prerequisites

- Proficient in C++
  - Modern C++

- Data structure and algorithm

- Interest in Bioinformatics

- Strong motivation to learn

Topics

- Quick review of modern C++ features and software development process

- Introduction to **parallel computing**

- Introduction to modern **Bioinformatics**

- Applications in Bioinformatics
  - Algorithms and practical guide for some parallel techniques
    - **Parallel techniques (in modern C++)**
      - **Multithreading**
      - **GPGPU**
      - **SIMD**
      - **MPI**
    - Parallelized sequence analysis algorithms
  - Collaborative Team Project

Bioinformatics

- Basic molecular biology
  - Transcription
  - Translation
- Sequencing technology
  - Next generation sequencing
  - Third generation sequencing
- Applications
  - Genotyping
  - Gene expression
  - Epigenetics

Algorithms

- Biological sequence sorting
- FM-index / FM tree
- Graphical FM-index
- Pairwise Sequence alignment
- Sequence-to-graph alignment
- De Bruijn graph
- …

Parallel programming

- Multithreading
  - C++11 thread
  - C++17 Parallel Algorithms
- GPGPU
  - CUDA
- SIMD
  - Boost::SIMD / simdpp
- MPI
  - OpenMPI

What will you need to do for this class

- 3-4 Coding homework in two months
- Team work project
  - 2-4 people / team (at most 2 members from the same lab)
  - Everyone will need to do the presentation
    - Project proposal (literature summary)
    - Progress report and Panel discussion
  - Github project scheduling and load sharing
    - Commit history
- Demo and final presentation

Project topics (TBD)

- Wavefront algorithm
- sBWT/KISS
- BWA -mem
- Graphical FM index (HISAT2)
- GATK
- Minimap2
- Cache-Oblivious parallel SIMD Viterbi HMM / HH-suite3
- Other topics can be proposed to me before execution

Important changes

- No physical class on Friday
  - **Online prerecorded class between classes (limited access time)**
- Github project scheduling and load sharing
  - Commit history
  - Final demo and presentation slides

Questions?