February 8th Milestone Report
Sunny Nahar

Major Changes:
There are changes in the direction of project for this semester. Initially the project was geared towards using frequency prediction as a heuristic to speed up DNA read mapping. Based on the work last semester, we have decided that some work will be dedicated towards improving the accuracy and efficiency of the heuristic that was built. But the focus now will be the improve the efficiency of the read mapping architecture and optimize its performance. This will involve restructuring the code to allow for parallelism and memory efficiency.

What You Have Accomplished Since Your Last Meeting:
Since last week, the goals of this research were still in flux, but now they have been decided. I started some preliminary work with looking at parallel frameworks which I can potentially use. The architecture of the code is very domain specific, so I have begun in researching possible approaches and determining whether they will be worth pursuing.

Meeting Your Milestone:
There are a new set of milestones, as detailed below.

Surprises:
No particular surprises yet.

Looking Ahead:
The focus of the next two weeks developing a viable approach for optimizing the performance the mapper infrastructure. I will look at parallel frameworks, and determine which part of the infrastructure can be effectively improved.

Revisions to Your Future Milestones:
Given the new project, here are a revised set of milestones:

- **February 22nd**: Investigate parallel frameworks. Explore whether CPU computation is the best option, as opposed to GPUs. Decide on a viable approach.

- **March 14th:** Start thinking about how best to parallelize the work. If this involves increasing work, determine how effective the method will be, as writing the code will likely take the rest of the semester. Think about the NUMA (Non-uniform memory) architecture and how to build data structures across different memory nodes. Continue work on heuristic.

- **March 28th**: Continue work on the heuristic. Explore possible approaches to optimize the heuristic by looking at complexity, cache-efficiency, and possible multithreading expansion. Continue work parallelizing rest of the infrastructure.

- **April 11th**: Finish writing the code and make sure everything is in working order. Start work documenting all changes. Obtain datasets for benchmarking.

- **April 25th**: Benchmark and report final performance improvements. Report final mapper efficiency improvements by the heuristic. Develop ideas for future work.

Resources Needed:
All resources are available as of now.