**Project Proposal**

For our project, we will be taking a problem from Online Judge and solving it using parallel programming methodologies. The problem we will be solving is taking multiple strings and processing them. The purpose behind this is to build a predictive text program, the user will enter a single letter, and the program will search for the next word to use based upon the most common word that follows it based upon the strings processed.

Per the problem on Online Judge, Kewl Texting, we will be taking strings and processing them to create a sentence stemming from a single char being entered. Taking sample text

1. What a nice day
2. This is the nice restaurant he talked about
3. we want nice weather and hope it is a nice day

After processing the strings, we make a list of the words that follow each word. In this case, we if we take the word “nice” we see “weather”, “restaurant” and “day” follow, with “day” being the most common with two occurrences. We also take into account the word that follow a single char. Typing a “w” will lead to the words “what”, “we”, “want” and “weather”. When the words are tied for an occurrence, we use the frequency of the word overall. We also want to keep track of the words that start and end sentences. In the event that all factors are equal, number of occurrences, if the word starts or ends a sentence the same number of times, then we will use alphabetical order as a tie breaker. Therefore, using the above data, if we enter the char “T” the sentence expected would be:

“This is a nice day”

To process a large amount of string to get a sizable dictionary of words in a serial program would take a long time. We can use MPI or even Pthreads to break the work load into parts and process chunks of data at the same time, using shared memory to store the data after it has been processed.

Problem: <https://onlinejudge.org/index.php?option=com_onlinejudge&Itemid=8&category=882&page=show_problem&problem=5227>