Victor Wu

CS 350

Feb 24, 2015

**Project 2 Design Document**

Objective

To create a program or application that extracts “tweets”, from most likely twitter.

Constraints

* Tweets have at maximum 140 characters (see [dev.twitter.com](https://dev.twitter.com/overview/api/counting-characters) for more info)
* Be able to handle at least 500 KB of data (1 KB each line)
* There is 10 maximum data, or slots, for the circular array

Design

* Create a custom circular array, queue stack would be ideal
* Create a 10x140 buffer size
* Approach this similar to the Producer Consumer model, and use semaphores appropriately
* Use pthreads and block until the successful termination of both, using join

Results

The actual time results were very similar to what we did in lab 1. The recorded time is as follows:

|  |  |  |
| --- | --- | --- |
|  | Wall Clock Time | CPU Time |
| Lab 1 | 0.022 | 0.000 |
| Project 2 | 0.022 | 0.000 |

The reason for such similar timing, despite the ideal theoretical case in which project 2 should be faster is, in my guess, because our input size is not big enough. While we were asked to hold 1KB each line, so that in a 500 line file, there is 500KB – the additional constraint of a tweet being no more than 140 characters contradict this; that is to say c1 contradicts c2.