Mason Reali

Assignment Graders

CSE 13

29 November 2020

Write up Assignment six

When toying around with the lab I have discovered a few things, and my eyes have been open to how painful it is to be efficient. The first thing I decided to test, naturally, was what would happen when I pushed my values to the extremes. What I did was set my bloom filter to size one, rendering it useless, and set my hash table to one.

The first thing that happened was an utterly long load time and a massive amount of traversing through a single linked list. I guess that would make sense. By making the bloom filter useless I essentially made every word go straight to the hash table. This meant that every inputted word had to be searched for in the single linked list I had allocated. What this revealed in a big way was just how effective the move to front rule could be. A Lot of the words tested were natural language, and the most common banned word in natural language was “the”. By using the move to front rule “the” was consistently at the top of the linked list, which made every repeat search for it incredibly fast. This, along with every other common

word being moved towards the top, helped to cut my search length down by almost a third.

When the code was working properly, though, with enough space in my bloom filter and hash table, the move to front rule was essentially useless as it was rare to traverse a list with more than two or three values. Overall this showed me how to dynamically deal with large amounts of data and different solutions to overcoming search/sort times.