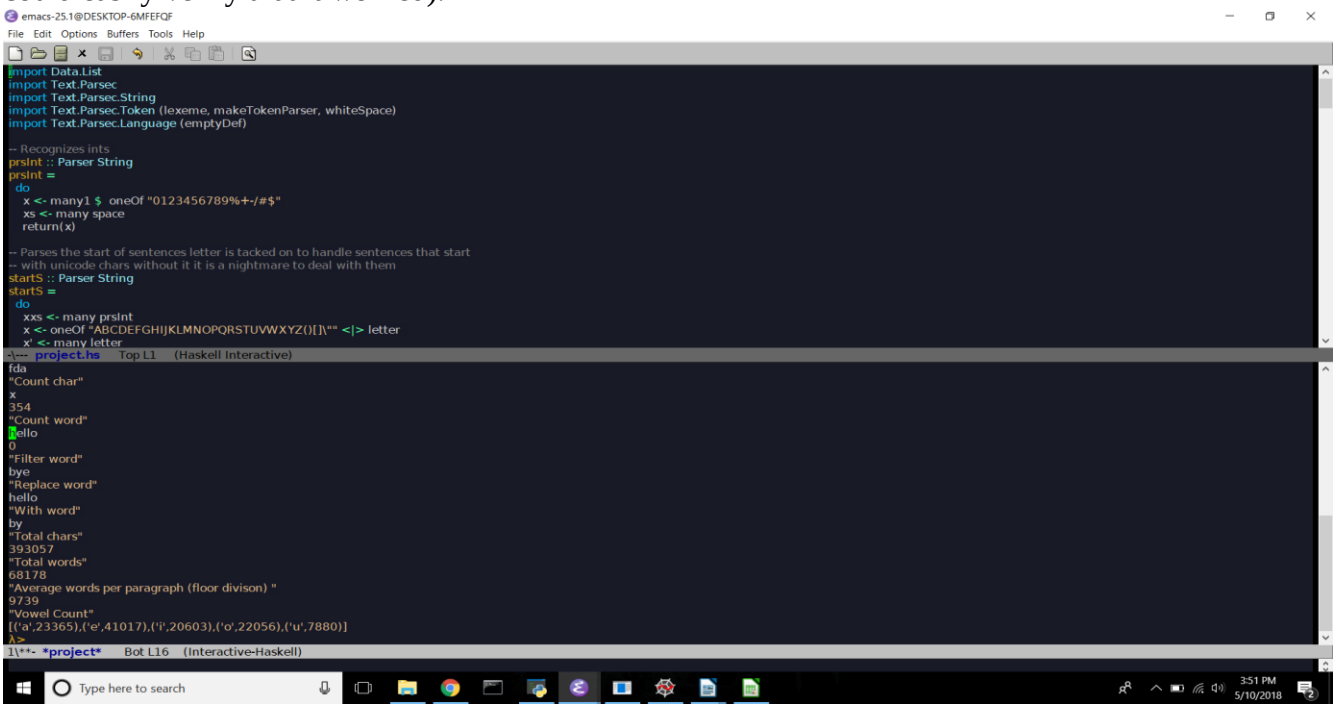
Final Project Information for Users

My parser reads the word the Scarlet Letter by Nathaniel Hawthorne taken from project Gutenberg. In order to run it you can just run main in the parser.hs file and it should work. It works on my machine but you never know for sure. The main.hs file does nothing I created when building the project and its really only there because its in the cabal file. The only changes I made were to remove everything that wasn't the text of the actual story (legal information or tables of contents tend to not look or behave like English sentences). I also had to clean up two instances of bad punctuation from the book. In two cases there were sentences that had this form (some text)!", ← that is grammatically incorrect and my parser is for correct English sentences so punctuation like that isn't allowed. This parser could parse other documents but a surprising number of texts from project Gutenberg have grammatical mistakes that cause the parser to crash and would take some fixing up.

When you run my parser you'll get some prompts once its finished. You can replace substrings in the document. For the substring replacer if you put in the string and it will replace and with whatever you tell it to but it will also replace and where it occurs in other sentences (so land, sand, etc) would also be changed. Next it will prompt you for another string allowing you to  count occurrences of a char. Next it asks for another input and it will count the occurrences of that string as a word (not substring). This action is case sensitive. Then it will prompt you for a word to filter out of the document (this action is also case sensitive). Finally it will prompt you for a full word to replace and the word you will replace it with.

After this is completed it will print out some information from the document such as the number of vowels, the average number of words per paragraph (floor division since you can't have a tenth of a word), and so one. Once everything is concluded it will save your changes to a new document called changes and you can verify what you've done. I uploaded my changes doc for scarlet letter and you can check a few of the things I did. I filtered out all upper case occurrences of Hester (the name of the main character) and I replaced the word the with beans (I choose something odd and a common word so you could easily verify that it worked).




I didn't accomplish everything that I wanted to in the project I was never able to create a grammar corrector as anything that would violate the grammar would cause the parser to fail.