

CS461 PA 2: HTML to Latex

Ksenia Burova

September 22nd, 2017

Abstract:

In this programming assignment, the goal was to learn how to use the Lex lexical analyzer generator by creating a program that would convert html file to equivalent Latex file. The were 'comment', 'paragraph', 'ordered list', 'unordered list', and several other html tags matched using regular expressions and turned into Latex ones.

Input / Output:

The input for the program is an **.html** file with a variety of tags, included nested ones. The output is a **.tex** file that gets converted to a **.pdf** aftermath.

Programming approach:

I've started this lab by looking at the provided implementation of 'comment' tag. There was enough information hidden (nesting, reading all chars like they are and etc.) to apply it to the rest of the file scanning.

- **Design**

I tried to keep my solution consistent with provided implementation. I don't have much to say in this part, since the whole design was about identifying the correct set of regular expressions for lexical analysis.

- **Testing:**

I tested my solutions by comparing my resulting pdfs to the ones provided: test, test1, test2.

- **Debugging:**

I've implemented my program by one tag match at the time, so it was easy to debug it. Any error that would occur would correspond to tag I was working on.

- **Issues:**

The only issue with this assignment was was a lack of knowledge in Lex. But things could be guessed and tested easily.