COSC561 - M2.3.3 Programming Assignment 1 - html2LaTex

Giuseppe Schintu

University of Tennessee, Knoxville

COSC561 - M2.3.3 Programming Assignment 1 - html2LaTex

As an introduction to the Lex lexical analyzer, I am tasked to complete a parse of partial HTML tags with a one level nesting condition for some predetermined HTML tags. As a disclaimer, I attempted to run the *soln_html2latex* file, but I wouldn't run on my macOS, so I figure that I would skip the quick check, at least initially. I did some research on how LaTex file should look like and got a general idea of what I needed to work on.

My initial approach was to observe how the *<h1>* and COMMENT section Lex rules were implemented. Of course, initially I did not quite understand the reason to use a Lex "start condition" for the COMMENT section. I quickly implemented all the HTML tags just simulating the HTML header *<h1>* given as an example, and I thought I was ready to give it a try with the test.html file. I did not want to copy&paste the test html file every time, so I read the *Makefile* and realized that there was no file passed as an argument, so I went ahead and added the section to the *main()* function. Any first HTML tag level worked fine, but honestly; I spent a few days trying to code to submission the HTML nesting tags like:

*<p>
etc...
<p>*

and

<p><p>.

After a couple days of researching the documentation, the start conditions kept coming in mind, and I realized that there must have been a reason for:

%x COMMENT PARA PRE OL UL.

At this point, I complicated my approach with a new LI start condition *<li>*, and it worked for the most part, but I didn't quite understand why. After some deleting conditions and moving

some combinations around, I realized that I only needed to solve the HTML nesting for cases such as *<p>,<ul>*, and *<ol>*, so I got rid of the LI start condition as it did not need a reset of the condition anyways.

I am not sure if I missed something, or if there is a better way to solve this problem; however, with the use of initial conditions and deriving the concatenation for the nesting tags, I was able to clear the test.html and additional cases such as:

*<p>
etc…
<p>*

and

<p><p>