# COP-701 Lab assignment 1 (HTML to LATEX)

Submitted By: Om Prakash (2019mcs2567)

September 1, 2019

**a) Details of lex and yacc operations in your assignment.** In the lex file named as P7.l, regular expressions and token definitions are mentioned which are checked from the source file. In the yacc file which is named as P7.y, grammar, types of tokens, types of terminals, and nonterminals are mentioned. In the yacc file, semantic rules are also written to construct the AST and store the respective values in the leaf nodes.

An HTML document is provided as in input to the lex, it scans the source code and when a text from the source file matches with the definition of a token which is mentioned in the P7.l file, it returns a token to the yacc parser. The parser looks into the grammar written in P7.y file and check if it's a part of the correct grammar if not, it throws a syntax error.

**What do you store in the AST and how?** During AST construction, the respective mappings of Latex are stored in the Leaf nodes based on the mentioned grammar. It constructs a binary tree which is evaluated later to convert HTML into Latex. AST contains two types of nodes, the nodes with node type as 'K' are leaf nodes, and the nodes with node type as 'T' are internal nodes. The Tree is traversed in a bottom-up manner; it takes the value from a node and then frees it. Nodes are created dynamically using malloc function and later they are freed to avoid any space wastage.

**c) How did you translate the AST?** Mappings from HTML to Latex are done either in lex file or yacc file. The parser traverses the AST in a bottom-up manner when it finds a value in a leaf node it takes and keeps traversing until the root node is visited. once the root node is visited it writes the collected values in a Latex file and frees the tree. During traversal it is checked if the visited node is an internal node or a leaf node, furthermore it's also checked if any of the children of an internal node is empty.

**d) Which programming language did you use?** I used C programming and gcc parser.

**e) Details about extra things that you did in the assignment (for bonus).** I covered more number of greek letters in my code and all type of text styles like bold, small, em, etc can be used any where in the code. Code is short and simple divided into major four files P7.l P7.y P7.h and P7.c. Standard naming conventions are followed.