

Assignment 1-HTML to Latex

ANANT JAIN (2019MCS2557)

1. Lex and yacc operations:-

a. In the lexical part of the assignments, all the html tags like `<html>`, `<head>` etc are converted into tokens and are sent for parsing. The code for lex file is written in “`hlat.l`” file.

b. Also, Html comments are ignored and data are converted into ‘WORD’ tokens.

c. Attributes of Html tags like `<a>`, `<table>` etc are extracted in the lexical operations.

d. The bison file named “`makeast.y`” contain grammar rules and productions for making tree.

2. Storing of AST:-

a. The nodes of AST are of form :-

```
typedef struct node
{
    int i;
    vector <node *> children;
    string data;
}node;
```

where integer i – used for type of nodes

vector children – used for storing children of node.

String data – used for storing data.

It make use of functions – `makenode()` and `newstr(string)` to make nodes in the tree.

3. Translation of AST :-

1. The appropriate latex tags are mapped to html tags during parsing and stored in the nodes of the AST.
2. The AST tree is traversed using function named “traverse_main” which give the latex text stored in the tree.
3. Traversal of AST tree is like the preorder traversal of the tree.

4. Programming languages used :-

Language used in this assignment – C, C++.