Assignment #1

CS347

Date: 19 Jan 2017

Marks: 5 + 5+ 10 = 20

1. Answer the following programs:

    a. Write a Lex program to extract all html tags and data-contents from a given html file. For example,

    ```
    <html>
    <head>
            <title>Hello HTML</title>
    </head>
    <body>
            <p>Hello World!</p>
    </body>
    </html>
    ```
    You need to extract the tags like "html", "head", etc from the above file.

    b. Design your own set of tokens for html code. Given an html file, write a Lex program to extract and display the tokens and the associated attributes (if exist) from the file.

2. Consider a simple C-like language with

    a. Data Type : integer (INT/int), floating point (FLOAT/float)
    b. Condition constructs: if
    c. Loop Constructs: for, while
    d. Relational operators, assignment and arithmetic operators
    e. Functions

Example

main()

```
{
    INT i=0;
    INT sum=0;
    INT count;
    read(count);
    for(i=0;i<10;i++)
    {
        read(x);
        sum+=x;
    }
    print(sum);  // print hello

 /******
 Sample program
 *********/
}
```

Implement a Lexical Analyzer for the above-mentioned simple C-like programming language using 'Lex' tools. Take a file as input (You can test it on the above example file). Output the tokens along with the associated attributes (if any).

Note: This is instructed not to ignore comments. Rather, consider a token for comments also. For the above example, you need to extract "print hello" and "***** Sample program ******** " as tokens for comments.

Submission Deadline- in one lab session

Submission Steps:
- Open 172.16.1.3/~halder
- Click on "Submission System"
- Login: <your roll no.>
- Password: <blank>
- Once Login first time, you must change your password
- Create a directory with name "assignment01" in your computer system
- Put your assignment files in the directory
- ZIP the directory to create "assignment01.zip"
- Upload the zip by clicking on Upload button.