Chi-Hao Tu & Yaoyao Liu

COMP 141: Lexics and CFGs

Instructions: In this exercise, we are going to review

1. regular expressions as a mechanism to specify tokens

2. a toy scanner implementation

1 Regex

1. Define the regex for the following description of tokens:

    (a) Any string that starts with character t.

    Regex: ^t.*

    (b) Any string of at least length 3 that starts with t and ends with u.

    Regex: ^t.{1,}u$

    (c) Any string that specifies the range of numbers between 11 and 23.

    Regex: ^(1[1-9]|2[0-3])$

    (d) Any string that specifies a date in MM:DD:YYYY format.

    Regex: ^(0[1-9]|1[0-2])\:(0[1-9]|1[0-9]|2[0-9]|3[0-1])\:[0-9]{4}$

    In C, an identifier is defined as a string of characters (both upper-case and lower-case), digits, and underscore "_", starting with either a character or underscore. Define the regex for identifiers in C.

    Regex: ^[a-zA-Z_][a-zA-Z_0-9]*$

2. Give five strings that conform with the regex: **[0-9]+((E|e)(\+|\-)?[0-9]+)?**

    1E10, 2e-9, 123E+6, 4e05, 56E-7

2 Toy scanner

1. Write a simple program in C++ that receives a single token as input and checks
   - if the received token is a positive integer number
   - if the received token is a punctuation character: +, *, (,).

   You must use regular expressions in contrast to the example in lecture, and check if the input matches the token definitions above.
   The output of the program may look like the following:

   $ ./main
   716235
   number

   $ ./main
   (punc

   $ ./main
   * punc

~/Desktop/2 Toy scanner.cpp

```
 1  #include <iostream>
 2  #include <regex>
 3  #include <string>
 4  using namespace std;
 5
 6  int main(){
 7      string token;
 8      cout << "Enter a token: ";
 9      cin >> token;
10
11      regex integer_regex("^[+]?[0-9]+$");
12
13      regex punctuation_regex("^[+*(),]$");
14
15      if (regex_match(token, integer_regex)){
16          cout << "number" << endl;
17      }
18      else if (regex_match(token, punctuation_regex)){
19          cout << "punc" << endl;
20      }
21      else {
22          cout << "unknown" << endl;
23      }
24
25      return 0;
26  }
27
28  /* Chi-Hao Tu & Yaoyao Liu*/
```