CS476: Automata Theory and Formal Languages
Homework 1
Assigned: 1/03/2013
Due: 15/03/2013 17.00

Questions

1. (20pts) State whether the following statements are true or not. You must give a BRIEF explanation or show a counter example to receive full credit.

   (a) (5pts) Any language of finite size is regular.

   (b) (5pts) If L_1 and L_2 are nonregular languages, then so is $L_1 \cup L_2$.

   (c) (5pts) $(R+S)^*S = (R^*S)^*$.

   (d) (5pts) With pumping lemma, we can prove that $L = \{w|\ w$ contains 100 as a substring$\}$ is a regular language.

2. (20pts) Give a DFA or an NFA that accepts the following languages and explain your reasoning.

   (a) (10pts) $L = \{w \in \{0,1\}^*|\ \ n_1(w) \mod 3 > n_0(w) \mod 3\}$.

   (b) (10pts) $L = \{w \in \{0,1,2\}^*|\ w$ when read as a number in base 7, is a multiple of 5$\}$.

3. (20pts) Give a regular expression for each of the following languages and explain your reasoning.

   (a) (5pts) $L = \{w \in \{a,b\}^*|\ 2n_a(w) + 3n_b(w)$ is even$\}$.

   (b) (5pts) The language consisting of all strings where 0s may occur only in the even positions. ($\Sigma = \{0,1\}$)

   (c) (10pts) All strings defined over the alphabet $\{a,b,c\}$ that contain at least one occurrence of all symbols in the alphabet.(You can obtain your expression from NFA of the language.)

4. (20pts) Are the following languages regular? Prove your answer.

   (a) (10pts) $L = \{a^n(a+b)^*b^n|\ n > 1\}$.

   (b) (10pts) $L = \{w\overline{w}|\ w \in \{0,1\}^+$ and $\overline{w}$ is the string obtained by taking bit-wise complement of w $\}$.

5. (20pts) Show that the set of regular languages are closed under the following operations and explain your reasoning.

   (a) (10pts) SWAP(L)$= \{w_1w_2 \in \sum^*|\ w_2w_1 \in L\}$.

   (b) (10pts) SUB(L)$= \{w \in \sum^*|\ \exists\ u,v \in \sum^* : uwv \in L\}$.

6. (20pts) In this question, you will develop Perl scripts for two real-life problems. Use Perl's regular expression find-replace capabilities to develop the following programs.

(a) (10pts) Write a perl scipt named `format_dates.pl` that will change the format of the dates in a given input text file. If the input file's base name is XYZ (i.e., the file name is `XYZ.txt`), the program should be run as `format_dates.pl XYZ.txt` and produce the output file `XYZ-modified.txt`. The program should read the input file line by line. Each line may contain dates formatted as MM/DD/YY. The program will simply change the format of these dates to DD-MM-20YY and write the line to the output text file.

Sample input (`dates.txt`):

```
Today is12/23/02 and ...   03/35/55 ...
This is a fraction 1/2 or something else 3/234/12 ...
```

Sample output (`dates-modified.txt`):

```
Today is23-12-2002 and ...   35-03-2055 ...
This is a fraction 1/2 or something else 3/234/12 ...
```

(b) (10pts) Write a perl scipt named `convert_dir.pl` that will change the names of files and subdirectories in the current working directory of a Unix system. The program will be piped with the `ls` command (it should be run as `ls -F | convert_dir.pl`). The conversion to be done is as follows: Directory names should be completely capitalized. File names should be in lowercase except for the first character, which should be capital. Moreover, both in directory and file names, blanks should replaced with underscores. For instance, let the following be the result of the command `ls -F` in the current directory:

```
Aleksandr  Pushkin/
Anna Karanina
Lev tolsToy/
all poems of Pushkin 2
criMe and pUnishment
fyodor dOsToyevsky/
the brothers karamazov
war and Peace
```

The same command should return the following after the conversion:

```
ALEKSANDR__PUSHKIN/
All_poems_of_pushkin_2
Anna_karanina
Crime_and_punishment
FYODOR_DOSTOYEVSKY/
LEV_TOLSTOY/
The_brothers_karamazov
War_and_peace
```

Caveat: Answers for questions 1-5 should be returned in a stapled hard copy and the answers for question 6 should be an attached to an e-mail that will be sent to your TA (enver@cs.bilkent.edu.tr), with the subject line cs476hw1. Do not send your homework to any other TA or do not use any other subject line. Please use the following procedure to submit your homework. Note that the procedure is strict, if your file/folder names are incorrect we may ignore your submission.

1. Create a folder with name **SurnameName** using English letters, e.g., YildizIbrahim. (This is the only folder you will create, do not put any other folder in it.)

2. Copy your source files **format_dates.pl, convert_dir.pl** into it. Please do not copy any input file to your folder. Your source files will need an input file and the filename must be the first argument your program reads from the command line.

3. Prepare a **readme.txt** which includes the commands and parameters to execute your programs. Copy it into the folder.

4. Compress the folder and create **SurnameName.zip**.

5. Email it to the TA **with the subject line cs476hw1**.