

CS406: Compilers

Homework Assignment 1 - Regular Expressions, Automata, Tries, CFGs Due: 1/2/2022

The objective of this assignment is to practice applying the concepts underlying:

1. Regular expressions
2. Automata
3. Efficient lexical analyzer algorithms and data structures.
4. Context Free Grammars

1 Problem Set

1. (1 point) Suppose student roll numbers (unique IDs) in an institute have the following format: `aa0bbccc`, where:
 - `aa` represents the last two digits of the year in which a student enrolls for a program.
 - `bb` represents the two-digit department code of the department where the program is offered.
 - `ccc` represents the unique three-digit number.

write a regular expression that matches *valid* roll numbers.

2. (2 points) Draw a DFA for the regular expression that you answered previously.
3. (2 points) Draw an NFA that matches the following regular expression:
`((((01)+|(11)+)))*`
4. (2 points) Draw the equivalent DFA for NFA drawn in the previous question.
5. (1 point) Find the reduced DFA (if it exists) for the DFA drawn in the previous question.
6. (1 point) Choose any 3 roll numbers of your friends from the 3 departments of ECE, MMAE, and CSE each. For the 9 roll numbers chosen, construct/draw a Trie. You may choose any popularly used representation for a Trie.
7. (1 point) Describe the language defined by the following context-free grammar (*in not more than two sentences. Exceeding this limit will attract a penalty.*)
 - (a) $S \rightarrow X$
 - (b) $X \rightarrow aXb$
 - (c) $X \rightarrow Y$
 - (d) $Y \rightarrow cd$
8. (1 point) Can this language be expressed as a regular expression? Why or why not?
9. (1 point) What are the terminals and non-terminals of this language?
10. (1 point) Show the derivation of the string `aacdbb` starting from `S` (specify which production you used in each step), and give the parse tree according to that derivation.

2 What you need to submit?

- A single document named `CS406HA1_response_<yourrollnumber>.pdf`. E.g., a student with roll number 10002000 would submit the response file `CS406HA1_response_10002000.pdf` containing:
 - *handwritten* responses to all questions.

3 How to submit?

1. Click on the link shared with you on the discussion forum. This will create a repository in your GitHub account. Clone the repository into your local development environment.
2. Write your responses clearly on a sheet of paper and convert it to digital form (by taking a photo or scanning). Name the file (digitized) as mentioned previously.
3. Add the file to your GitHub local repository using `git add` command.
4. Save the changes using `git commit` command
5. Upload the changes to GitHub using `git push` command
6. Release your changes by first tagging your commit on the *local environment* using `git tag -a cs406ha1submission -m "<release description>"`.

Next, push the tag to *GitHub* with the help of the following command: `git push -tags`

If you want to make changes after you have submitted (repeat the above steps from 2 to 5 and apply modified commands shown below in place of step 6):

```
> git tag -a -f cs406ha1submission -m "<release description>"
> git push -f --tags
```

Make sure that the digital versions of hand-written documents are clearly visible. You will loose points for unclear responses.