Cairo University

Faculty of Engineering

Department of Computer Engineering

Languages and Compilers

Programming Assignment

Regex-to-NFA Converter

Remonda Talaat Eskarous
SEC:1, BN:19

Mohamed Shawky Zaky
SEC:2, BN:15

1 Tool Description

This is a `Python` tool that takes an input *regex* and outputs the corresponding *NFA* with its *graph*. It uses **Thompson's** rules for conversion. The input *regex* can contain the following :

- **Meta-characters :** ( ) *for precedence*, | + *for ORing*, * *for repetition*
- **Letters :** from *A* to *Z (uppercase or lowercase)*
- **Digits :** from 1 to 9
- **Special characters :** some special characters can be used as a transition element, however it `must` be preceeded by \, for example \−.

The output follows the same format described in the main assignment document, along with the directed graph visualization.

2 Tool Usage

First, install the requirements in *requirements.txt* :

- pip install -r requirements.txt

After that, run the tool as follows :

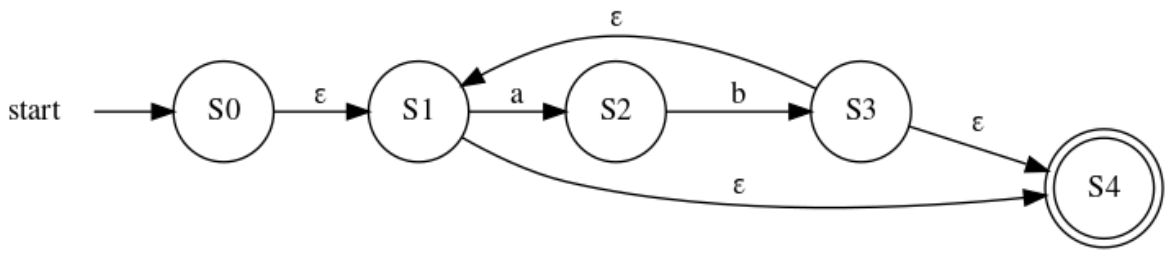
- python convert.py "REGEX"

For example, *python convert.py "(A|B)*(CD)"*
Output *NFA* and its *directed graph* will be exported to `out` folder.
Also, the outputs of given test cases can be found in `test_cases` folder.

3 Test Cases

3.1 Test Case 1

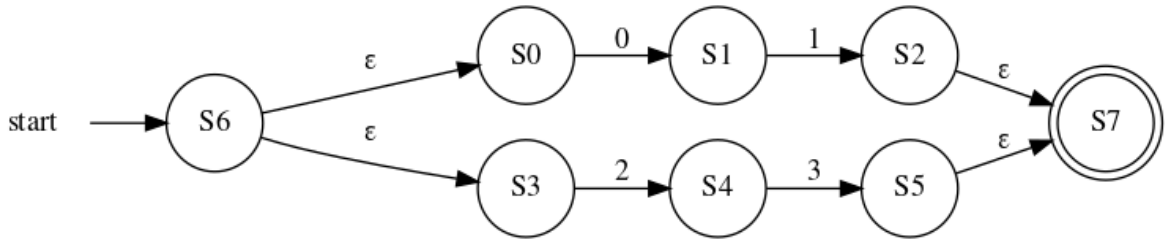
Input : *(ab)**



3.2 Test Case 2

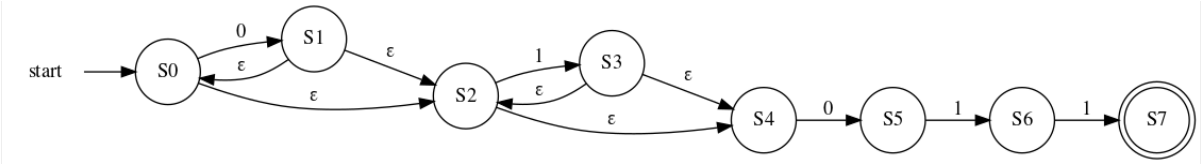
Input : *01|23*

Note : In such case, our converter consider the precedence to be like that *(01)|(23)*, so 01 are considered as a branch and 23 as another branch, then both branches are ORed together.



3.3 Test Case 3

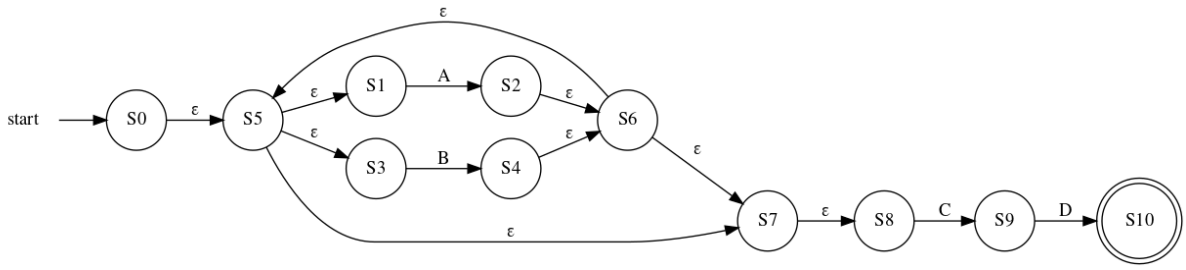
Input : *0*1*011*



4 Other Output Samples

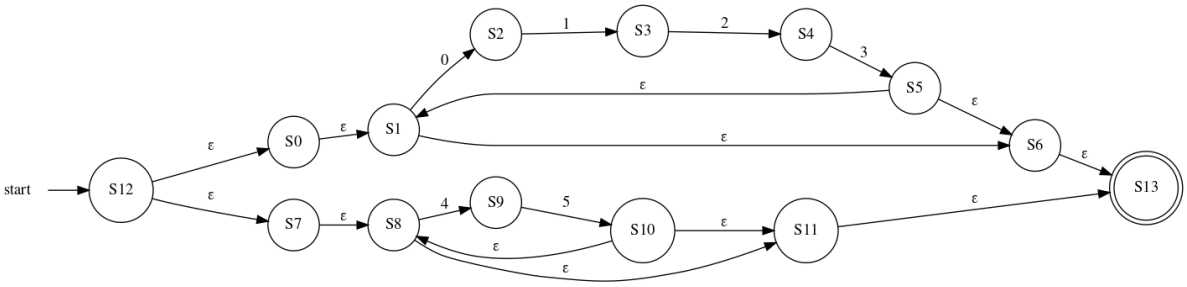
4.1 Example 1

Input : *(A|B)*(CD)*



4.2 Example 2

Input : *(0123)*|(45)**



4.3 Example 3

Input : *A*B*|C*D**