

CSC 416/565: Compilers
Fall 2023

Written Homework #1

Date Out: Tuesday, September 12

Due Date: Tuesday, October 3 (beginning of class)

Regular Languages

1. Write an English description for the languages generated by the following regular expression:

4 (a) $(0|1|\dots|9|A|B|C|D|E|F)^+(x|X)$

2. Write regular expressions for each of the following.

4 (a) English days of the week: **Monday**, **Tuesday**, ...

4 (b) All strings of digits that contain no leading zeros.

4 (c) All strings of digits that represent even numbers.

3. For each of the following regular expressions determine which of the strings **cc**, **ababb**, **bbcab**, and **ccbbaab** matches it:

4 (a) $(ab)^*c|a^*b^*c^*$

4 (b) $(ab|bc^*)^+$

4 (c) $[ab]^*cc?(ab)^*$

Lexical Specifications

- 3 4. Given the string **abbbaacc** and alphabet $\{a,b,c\}$ what tokenization will the following lexical specification produce?

Token Class	Regex
-----	-----
A	b^+
B	ab^*
C	ac^*

- 3 5. Given the string **babac** and alphabet $\{a,b,c\}$ what tokenization will the following lexical specification produce?

Token Class	Regex
-----	-----
A	$a(ba)^*$
B	$b^*(ab)^*$
C	abc
D	c^+

- 4 6. Given the following lexical specification and alphabet $\{0,1\}$, which of the below strings will be successfully tokenized?

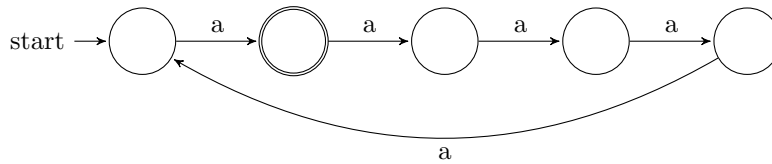
Token Class	Regex
-----	-----
A	$(11)^*$

B	01+
C	10+

1. 1000001
2. 1110010
3. 01100100
4. 10011001

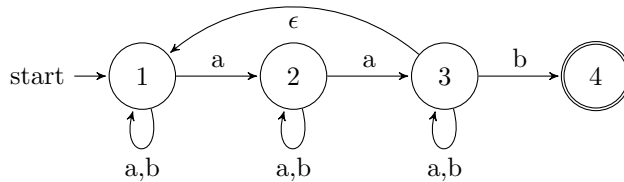
Finite Automata

7. Explain in informal English what this finite-state automata recognizes.

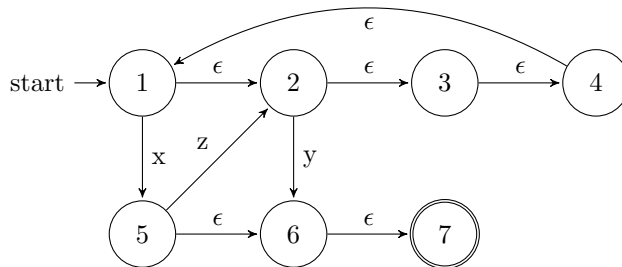


5 (a)

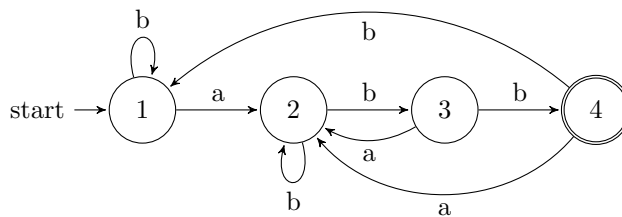
- 6 8. Write a regular expression whose language is equivalent to the following NFA.



9. Convert these NFAs into DFAs.



6 (a)



6 (b)

10. Construct DFA's for each of the following regular expressions. Do it in two steps: construct the NFA using Thompson's construction, then the DFA from the NFA. Let the alphabet be $\{a,b\}$.

6 (a) $a^*|b^*$

7 (b) $a^*(ab)^*a^*$

7 (c) **CSC416 ONLY:**

$[ab]^*abab$

CSC565 ONLY:

$[ab]^*abab[ab]^*$