

Exercise 1: Lexing

Question 1.1

Circle all of the regular expressions whose language is **disjoint** from the language of $(a|b)^+c^*$:

- A. $a^*b^*c^*$
- B. c^+
- C. $(b|c)^+$
- D. $(a|b)^+c^+a^+$

Question 1.2

Consider the lexical analyzer with token classes defined by following regular expressions:

\langle	IMP	:=	$=>$,
	LESSEQ	:=	$<=$,
	LESS	:=	$<$,
	GREATER	:=	$>$

Circle all the correct statements amongst the ones below, assuming the longest match rule is applied:

- A. $<=>$ will be tokenized to LESS, IMP.
- B. $<=<$ will be tokenized to LESSEQ, LESS.
- C. $<==>$ will be tokenized to LESSEQ, IMP.
- D. $<>$ will be tokenized to GREATER, LESS.

Exercise 1: Regular Languages

Consider the alphabet $A = \{a, b, c\}$.

Question 1.1

Which of the following regular expressions exactly describe the language of all words that contain an **a** immediately followed by **b**, and contain exactly one **c**. Circle letters next to **all** correct options.

- A. $(ab)c(a|b)^*$
- B. $a^*b^*(ca^*b^*ab)^*a^*b^*$
- C. $(a|b|c)^*ab(a|b|c)^*$
- D. $(a|b)^*((ab(a|b)^*c(a|b)^*)|(c(a|b)^*ab(a|b)^*))$
- E. $(a|b)^*(ab|c)(a|b)^*(ab|c)(a|b)^*$
- F. $((a|b)^*ab(a|b)^*c(a|b)^*)|((a|b)^*c(a|b)^*ab(a|b)^*)$