
AUTOMĀTU TEORIJA

Eksāmena materiāli

Author

Automata guy

Riga

2024

Contents

1	Basics	3
1.1	String	3
1.2	Language	3
2	Regex	3
3	Finite state machines (FSM)	3

1 Basics

1.1 String

A string is a finite sequence, possibly empty, of symbols drawn from some alphabet Σ . Given any alphabet Σ , the shortest string that can be formed from Σ is the empty string, which we will write as ϵ . The set of all possible strings over an alphabet Σ is written Σ^* . This notation exploits the Kleene star operator, which we will define more generally below.

1.2 Language

A language is a (finite or infinite) set of strings over a finite alphabet Σ . When we are talking about more than one language, we will use the notation Σ_L to mean the alphabet from which the strings in the language L are formed.

2 Regex

Elements of regex

$*$ - the asterisk indicates zero or more occurrences of the preceding element. For example, ab^*c matches "ac", "abc", "abbc", "abbbc", and so on.

$+$ - the plus sign indicates one or more occurrences of the preceding element. For example, $ab+c$ matches "abc", "abbc", "abbbc", and so on, but not "ac".

3 Finite state machines (FSM)

Formally, a deterministic FSM (or DFSM) M is a quintuple $(K, \Sigma, s, A, \delta)$, where:

- K is a finite set of states,

- Σ is the input alphabet,
- $s \in K$ is the start state,
- $A \subseteq K$ is the set of accepting states, and
- δ is the transition function. It maps from:

$$K \times \Sigma \rightarrow K$$

A configuration of a DFSM M is an element of $K \times \Sigma^*$. Think of it as a snapshot of M . It captures the two things that can make a difference to M 's future behavior:

- its current state
- the input that is still left to read.

M halts whenever it enters either an accepting or a rejecting configuration. It will do so immediately after reading the last character of its input.

The language accepted by M , denoted $L(M)$, is the set of all strings accepted by M .