MATH 2301 22/09/2023 ①

* Recap/formal def

Def : Fix an alphabet Σ. A regular expression (or regex) is a word r in the letters of Σ, together with symbols $*$, $|$, ε, ϕ, satisfying one of the following:

(1) $r = a$ for $a \in \Sigma$

(2) $r = \varepsilon$

(3) $r = \phi$

(4) $r = r_1 r_2$ where r_1, r_2 are valid regular expressions

(5) $r = r_1 | r_2$ ← "OR"

(6) $r = r_1^*$

* Note : Parentheses ( ) may also be used for dis-ambiguation or grouping

* As always, assume that none of the symbols $*, |, \varepsilon, \phi, (,)$ are in Σ.

②

* Today : Matching and the language of a regexp.

Def : Given a regex r, the language of r, denoted $L(r)$, is a language on Σ defined as follows:

(1) If $r = a$ for $a \in \Sigma$, $L(r) = \{a\}$

(2) If $r = \varepsilon$, $L(r) = \{\varepsilon\}$

(3) If $r = \phi$, $L(\phi) = \phi$ i.e $L(r) = \phi$

(4) If $r = r_1 r_2$, $L(r) = L(r_1) \circ L(r_2)$

(5) If $r = r_1 | r_2$, $L(r) = L(r_1) \cup L(r_2)$

(6) If $r = r_1^*$, $L(r) = L(r_1)^*$

Examples :

(1) $r = (01 | 11)^*$

$L(r) = L((01|11))^*$

Sub-question: What is:

$L(01|11)$? $= L(01) \cup L(11)$

$= \{01\} \cup \{11\} = \{01, 11\}$

$L((01|11))^* = \{\varepsilon, 01, 11, 0111, 0101, 1111, 1101, 111101, \ldots\}$

* Examples

(2) $r = (10)^* 1 \mid (01)^* 0$

$L(r) = L((10)^* 1) \cup L((01)^* 0)$

$L((10)^* 1) = L((10)^*) \circ L(1)$

$= L(10)^* \circ \{1\}$

$= \{\varepsilon, 10, 1010, 101010, \ldots\} \circ \{1\}$

$= \{1, 101, 10101, 1010101, \ldots\}$

$L((01)^* 0) = \{0, 010, 01010, \ldots\}$

$L(r) =$ strings with at least one letter that begin & end with the same letter, and alternate letters

* Notation: The regex $(0 \mid 1)^*$ has language Σ^*

Shorthand for $(0 \mid 1)$ is Σ

(more generally, Σ means $(a_1 \mid a_2 \mid \cdots \mid a_n)$ if $\Sigma = \{a_1, \ldots, a_n\}$)

Shorthand for $(0 \mid 1)^*$ is Σ^*

(3) $r = (\Sigma \Sigma 0)^*$

$L(r) =$ words of length $3k$, for $k \geq 0$, such that every third letter is a 0.

* Def: Given a regex r and a word w on Σ, we say that w matches r if $w \in L(r)$.

E.g. 000100 matches $r = (\Sigma \Sigma 0)^*$.

Example

(4) $L = \{ w \in \Sigma^* \mid w$ contains an even number of 1's$\}$

r $(11)^*$? $(1 0^* 1)^*$? $(0^* 1 0^* 1 0^*)^*$?

→ Return on Monday.