CSPC 105

Automata Theory and Formal Languages

L.P.Facun

Introduction

Alphabets, Strings and Languages

Alphabets: non-empty finite set of symbols

{ , }

$\Sigma_1 = \{0, 1\}$ → binary numbers

Σ - sigma

$\Sigma_2 = \{a, b, c, \ldots, x, y, z\}$ → English alphabet (lowercase letters)

to

$\Sigma_3 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ → decimal numbers

$\Sigma_4 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F\}$ → hexadecimal numbers

$\Sigma_5 = \{a, b\}$ ←

String: finite sequence of symbols from an alphabet

$\Sigma = \{0, 1\}$

$w = 0$

w = 10

w = 111111

$w = 1$

$w = \underline{10011101}$

$w = 00$

w = 01

Length of string / cardinality

$w = abcd$ 4 $w = 01$ $w = automata$

$|w| = 4$ $|w| = 2$ $|w| = 8$

$|w| = 4$

Null string or empty string

$w =$

$w = \lambda$

$|\lambda| = 0$

λ - lambda

Louis Philippe Facun

Substring

$w = \text{banana}$ $w = 01011$ $w = \text{automata}$

$z = \text{ana}$ ✓ - consecutively

$z = 010$ ✓

$w = \text{aut}$ ✓

$z = \text{anb}$ ✗

$z = 11$ ✓

$w = \text{mata}$ ✓

$z = \text{nab}$ ✗

$z = 00$ ✗

$z = 01011$ ✓

Concatenation: combine two string

prefix

suffix

x is $x_1, x_2, x_3, \ldots, x_m$

$x = 011$

y is $y_1, y_2, y_3, \ldots, y_n$

$y = 100$

x + y

$x_1x_2x_3 \ldots y_1y_2y_3 \ldots y_n$

$xy = 011100$

$yx = 100011$

w^k copies

– k copies of w

$w = ab$

$w^1 = ab$

$w^1 = ab$

$w^2 = abab$

$w^3 = ababab$

$w^4 = abababab$

Language (L): set of all possible strings from an alphabet, given a condition

$\Sigma = \{0, 1\}$ - no two consecutive 1's

$L = \{ \lambda , 0, 1, 00, 01, 10, 000, 001, 010,$

$100, 101, 0000, 0001, 0010,$

$0100, 0101, 1000, 1001, 1010, \ldots \}$

Language (L): set of all possible strings from an alphabet, given a condition

$\Sigma = \{0, 1\}$ - no two consecutive 1's

$L = \{ \lambda, 0, 1, 00, 01, 10, 000, 001, 010,$

$100, 101, 0000, 0001, 0010,$

$0100, 0101, 1000, 1001, 1010, \ldots \}$

Language notation

$L = \{x^1, x^2, x^3, x^4, \ldots\}$

$$\Sigma = \{x, y\}$$

x^1 x^2 x^3 y^4

$$L_1 = \{x, xx, xxx, xxxx, \ldots\}$$

$x = ab$

$$L_1 = \{x^k \mid \underbrace{k \geq 1}\}$$

language notation

k copies of 'x' such that k is greater than or equal to 1

Language notation

$xy, xyy, xyyy, xyyyy$

$xy^1, xy^2, xy^3, \ldots\}$

$\Sigma = \{x, y\}$

If $L_2 = \{\lambda, xy, xyxy, xyxyxy, \ldots\}$

$\{xy^0, xy^1, xy^2, xy^3, \ldots\}$

$L_2 = \{(xy)^k \mid k \geq 0\}$

$L = \{(xy)^k \mid k \geq 0\}$

$\{\ldots\}$ - Sets

Σ - Alphabet

w - string

λ - empty string

$|w|$ - length of string w

xy - concatenation of x and y

w^k - k copies of w

Questions?