Formal Languages

An introduction

ALPHABETS, STRINGS AND LANGUAGES

An alphabet is a set of symbols.

ex: { a, b, c, d, . . , z }

{ 0, 1 }

ALPHABETS, STRINGS AND LANGUAGES

- A string (word) is a concatenation of symbols of an alphabet.

  ex: abc

  jfhl

  rnl

- The length of a string is the number of symbols in the string.

ALPHABETS, STRINGS AND LANGUAGES

A (formal) language is a set of strings.

ALPHABETS, STRINGS AND LANGUAGES

- The empty string, λ , is the string with no symbols

- A Capital Letter is used to denote an alphabet.

  ex: $A = \{ a, b, c, d, . . . , z \}$

  $B = \{ 0,1 \}$

ALPHABETS, STRINGS AND LANGUAGES

- **Some Notations**

- W notation for the set of all possible strings that can be formed out of a given alphabet.

- W_k represents a set of strings with length k, where length is equal to the number of symbols.

  ex: P = { 0, 1, a, b, c, d, else, for }

  length (elsefor) = 2

- Given an alphabet:

  $W_k = \{ \varphi \mid \varphi \text{ is a string with length} = k \}$

  $W_0 = \{ \lambda \}$

  $W = \cup W_k = W_0 \cup W_1 \cup W_2 \cup W_3 \cup \ldots$

ALPHABETS, STRINGS AND LANGUAGES

- **CONCATENATION**
  ("multiplication" of strings)

- The concatenation of $\omega = \omega_1\omega_2\omega_3 \ldots \omega_n$ and $\varphi = \varphi_1 \varphi_2 \varphi_3 \ldots \varphi_n$ is

$$\omega\varphi = \omega_1\omega_2\omega_3 \ldots \omega_n \varphi_1 \varphi_2 \varphi_3 \ldots \varphi_n = \omega\varphi$$

ALPHABETS, STRINGS AND LANGUAGES

- **OPERATIONS WITH LANGUAGES**

  1. Union
  2. Intersection
  3. Concatenation
  4. Complement (with reference to W)

ALPHABETS, STRINGS AND LANGUAGES

Exercise:

Let $L1 = \{ a, ab \}$ be a language
Let $L2 = \{ c, bc \}$ be a language

Union: $L1 \cup L2 = \{ a, ab, c, bc \}$\

Intersection: $L1 \cap L2 = \{ \} = \phi$

ALPHABETS, STRINGS AND LANGUAGES

Exercise:

Let L1 = { a, ab } be a language
Let L2 = { c, bc } be a language

Concatenation:

L1 . L2 = { ac, abc, abbc }

L2 . L1 = { ca, cab, bca, bcab }

L1 . λ = { aλ, abλ } = { a, ab }

ALPHABETS, STRINGS AND LANGUAGES

Exercise:

Let L1 = { a, ab } be a language

Let L2 = { c, bc } be a language

Note: $\lambda \neq \phi$;

$\boldsymbol{\lambda} = \{ \lambda \}$

$\phi = \{ \ \}$

L2 . $\phi = \phi$

L2 . $\boldsymbol{\lambda} = \{ c\lambda \ . \ bc\lambda \} = \{ c, bc \}$

ALPHABETS, STRINGS AND LANGUAGES

- KLEENE-STAR (*)

Suppose A = { 0,1 }, the kleene-star of A, A*={λ, 0, 1, 00, 01, 10, 11, . . .} is the set of strings that can be formed out of the elements of the alphabet/set A including the null string.

ALPHABETS, STRINGS AND LANGUAGES

Suppose L_1 = { a, ab }

The kleene-star of L_1,

L_1* = { . a, ab, aab, . . . } = the set of all strings that can be formed out of the strings of L_1.

ALPHABETS, STRINGS AND LANGUAGES

- Exercise:

Let A = { a, ab } B = { c, bc } C = { c }

Verify that:

a) concatenation of languages is associative

b) concatenation distributes over union

Solution

a) A . (B . C) = (A . B) . C

{a, ab} . ({c, bc} . {c}) = ({a, ab} . {c, bc}) . {c}

{a, ab} . {cc, bcc} = {ac, abc, abbc} . {c}

{acc, abcc, abbcc} = {acc, abcc, abbcc}

Therefore, concatenation is associative

ALPHABETS, STRINGS AND LANGUAGES

b) A . (B U C) = A . B U A . C

{a, ab} . ({c, bc} U {c}) = ({a, ab} . {c, bc}) U ({a, ab} . {c})

{a, ab} . {c, bc} = {ac, abc, abbc} U {ac, abc}

{ac, abc, abbc} = {ac, abc, abbc}

Therefore, concatenation distributes over union

ALPHABETS, STRINGS AND LANGUAGES

Note: $A^3 = A . A . A$

Let $A = \{a,ab\}$

$A^3 = ?$

$A^3 = (\{a, ab\} . \{a, ab\} . \{a.ab\})$

$A^3 = \{aa, aab, aba, abab\} . \{a, ab\}$

$A^3 = \{aaa, aaab, aaba, aabab, abaa, abaab, ababa, ababab\}$

ALPHABETS, STRINGS AND LANGUAGES

Note: ω^R = Reverse of ω

ω = ab ω^R = ba

Let A = $\{\omega \mid \omega$ is a string$\}$ = {a, ab}

Determine the strings in B if $B=\{\omega\omega^R \mid \omega \in A\}$

ALPHABETS, STRINGS AND LANGUAGES

Note: ω^R = Reverse of ω

ω = ab ω^R = ba

Let A = {a, ab}

Determine the strings in B if $B=\{\omega\omega^R \mid \omega \in A\}$

Answer

B = {aa, abba} = mirror image language

ALPHABETS, STRINGS AND LANGUAGES

SINGLETON SET NOTATION

The singleton set notation for {0} is the boldfaced **0**.

Similarly, {a} = **a**

ALPHABETS, STRINGS AND LANGUAGES

- Exercise:

Let V= {0, 1}. Using singleton sets and operators $\cup$, ., and *, find an expression

for the language that has

a. All $\omega \in V^*$ such that the string 101 appears as a substring of ω.

b. All $\omega \in V^*$ such that each 0 appearing in ω is immediately followed by at least two 1's.

c. All $\omega \in V^*$ such that each 1 appearing in ω is immediately followed by the substring 10.

ALPHABETS, STRINGS AND LANGUAGES

- Exercise:

Let v = {0, 1}. Using singleton sets and operators U, ., and *, find an expression for all $\omega \in v^*$ such that the string 101 appears as substring of ω

Answer: (**0** U **1**)* . 101 . (**0** U **1**)*