Formal Languages

- Symbol
- Alphabet
- String
- Language and operations on languages
- Chomsky classification of languages
- Languages and automata
- Languages and grammar
- Derivation

Symbol

- A symbol is a sign that conveys some meaning.
- Example
- 0,1,2, 3, 4, ...
- a, b, c, d, e, ...
- +, -, *, /, (,), { , }

Alphabet

- An alphabet is a finite set of symbols.
- Example
- { 0, 1 }
- { a, b }
- { a, b, c , d, e, ... z }
- It is denoted by Σ.

String

- A string over an alphabet is a sequence of symbols from the alphabet.
- It is usually denoted by w, x, y, z letters.
- Example
- $\Sigma=\{0,1\}$
- Strings 010011, 1101, 1, 01, 111, ...
- $\Sigma=\{a,b\}$
- Strings ab, abbba, bbba, a, bbaab, aaa, bbbbbb ...

String Operations

- **Length of String**
- Length of a string w is the number of symbols in the string w.
- It is denoted as $|w|$
- Example
- w=01101
- $|w|=5$

String Operations

- **Concatenation of string**
- Suppose x and y are two strings then concatenation of string xy is defined as join of two strings x and y.
- Example 1
- x=sun and y=day
- xy=sunday and yx=daysun
- Example 2
- x=10011 and y=1011
- xy=100111011 and yx=101110011

Empty String

- It is a very special string.
- It is defined as the string consisting of zero symbols (no symbols).
- Length of empty string is 0.
- It is denoted as ε(epsilon).

Powers of alphabet Σ^i

- Power of an alphabet Σ^i is defined as follows:
- $\Sigma^i=\{w|w \text{ is a string over alphabet } \Sigma \text{ and } |w|=i\}$
- It is the set of strings of length i over an alphabet Σ
- Example
- $\Sigma=\{0,1\}$
- $\Sigma^0=\{\varepsilon\}$: set of strings of length 0
- $\Sigma^1=\{0,1\}$: set of strings of length 1
- $\Sigma^2=\{00,01,10,11\}$: set of strings of length 2
- $\Sigma^3=\{000,001,010,011,100,101,110,111\}$: set of strings of length 3

Kleene Closure of Σ

- It is defined as set of all possible strings over alphabet Σ
- It is denoted as Σ^*
- $\Sigma^* = \{w \mid w \text{ belongs to } \Sigma^i \text{ for } i >= 0\}$
- $\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \Sigma^4 \cup \ldots$ Upto infinity

Positive Closure of Σ

- It is defined as set of all possible strings over alphabet Σ whose length is >=1
- It is denoted as Σ^+
- $\Sigma^+ = \{w | w \text{ belongs to } \Sigma^i \text{ for } i >= 1\}$
- $\Sigma^+ = \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \Sigma^4 \cup \ldots$ Upto infinity
- Note
- $\Sigma^* = \Sigma^+ \cup \{\varepsilon\}$

Formal Languages
Symbol -> Alphabet -> String -> Language

1. Symbol
2. Alphabet Σ
3. String
4. Empty String ε
5. String Length and Concatenation
6. Kleene closure Σ^*
7. Positive closure Σ^+
8. Language L

Formal Language

- A formal language L over an alphabet Σ is a subset of Σ^*
- It is defined as **set of strings in Σ^***
- L $\underline{C}$ Σ^*
- **Example 1: Language over alphabet Σ={0,1}**
- L={ w ε Σ^*| w ends with a ‘1’}
- L={1,01,11,001,011,101,111......}

Formal Language Examples

- **Example 2** : Language over alphabet Σ={0,1}
- L={ w ε Σ^*| w begins with a ‘0’}
- L={0, 00,01,000,001,010,011......}
- **Example 3 :** Language over alphabet Σ={0,1}
- L={ w ε Σ^*| w begins and ends with a ‘0’}
- L={0, 00,010,000,0010,010,0110......}

Formal Language Examples

- **Example 4 :** Language over alphabet Σ={0,1}
- L={ w ε Σ^*| w has even number of '0's}
- L={00, 010,001,0000,00010,01000,0101......}
- **Example 5 :** Language over alphabet Σ={0,1}
- L={ w ε Σ^*| w has odd number of '1's}
- L={01, 010,01011,0010,001110,01000,01110......}

Formal Language Examples

- **Example 6 :** Language over alphabet Σ={0,1}
- L={ w ε Σ^*| w has even number of ‘0’s and odd number of ‘1’s}
- L={010, 01110,001,00100,00010,0111000,01011......}
- **Example 7 :** Language over alphabet Σ={0,1}
- L={ w ε Σ^*| w has substring ‘01’}
- L={01, 010,00101,0010,001110,01000,01110, 1110111.....}

Formal Language Examples

- **Example 8 :** Language over alphabet $\Sigma=\{0,1\}$
- L={ w ε Σ^*| length of w i.e |w|=2}
- L={00,01,10,11}
- **Example 9 :** Language over alphabet $\Sigma=\{0,1\}$
- L={ w ε Σ^*|length of w<=2 }
- L={ε,0,1,00,01,10,11}

Formal Language Examples

- **Example 10 :** Language over alphabet $\Sigma=\{0,1\}$
- L={ w ε Σ^*| length of w i.e |w|>=2}
- L={00,01,10,11,000,001,010,011,100,010,110,111,.....}
- **Example 11 :** Language over alphabet $\Sigma=\{0,1\}$
- L={ w ε Σ^*|length of w is divisible by 2}
- Or { w ε Σ^*|length of w is even}
- Or { w ε Σ^*| |w| mod 2=0}
- L={ε,00,01,10,11,0000,0011,1001,1100,1111....}

Formal Language Examples

- **Example 12 :** Language over alphabet Σ={0,1}
- L={ w ε Σ^*| w is of the form 0^m1^n m,n>=1}
- It means L contains strings of the form where series of 0s followed by series of 1s
- L={001,01,00011,011,0001,001111,0111,000011,00111111,.....}
- **Example 13** : Language over alphabet Σ={0,1}
- L={ w ε Σ^*|| w is of the form 0^n1^n,n>=1}
- It means L contains strings of the form where series of 0s followed by series of 1s and number 0s and 1s are equal
- L={01,0011,000111,00001111....}

English Language

- Alphabet Σ={a,b,c,d,e,f,g....x,y,z}
- A sentence is nothing but string over alphabet.
- There can be infinitely many strings possible. Some of them may belong to the language.
- String 1 : I am a student .
- String 1 belongs to english language
- String 2: student i a am
- String 2 does not belong to english language.
- How do we say a given string is in the language or not? By checking grammar.(Set of rules)
- So we need grammar to define language. Grammar is a language generating device.