



# Regular Language (RL)

---

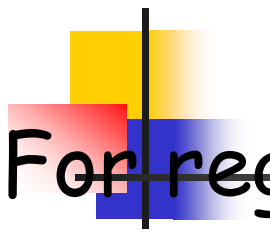
- Basic languages
  - $\phi$ : empty language
  - $\{\epsilon\}$ : null language
  - $\{0\}$ : simple language,  $0 \in \Sigma$
- A regular language over an alphabet  $\Sigma$  is one that can be obtained from these basic languages using the **operations** of union, concatenation and kleene \*

# Some Properties of Regular Languages



---

## properties



For regular languages  $L_1$  and  $L_2$   
we will prove that:

Union:  $L_1 \cup L_2$

Concatenation:  $L_1 L_2$

Star:  $L_1^*$

Are regular  
Languages



We Say:

Regular languages are **closed under**

Union:  $L_1 \cup L_2$

Concatenation:  $L_1 L_2$

Star:  $L_1^*$



# Regular Expression (RE)

---

- Let  $\Sigma$  be an alphabet, a RL over  $\Sigma$  can be described by an **explicit formula** known as **regular expression**. In other words, value of a RE is a language.
- REs are slightly simple formulas by replacing brackets  $\{ \}$  with parentheses  $( )$  and  $\cup$  by  $+$ .
- If  $r, s$  are RE over  $\Sigma$  denoting the RL  $R, S$  over  $\Sigma$ , then  $(r+s)$ ,  $(rs)$  and  $(r^*)$  are RE over  $\Sigma$  denoting  $R \cup S$ ,  $RS$  and  $R^*$  respectively.



# Examples of Regular Language

---

Let  $\Sigma = \{0, 1\}$  then RL over inputs are

- $\{0\}$
- $\{001\}$
- $\{0,1\}$  i.e.  $\{0\} \cup \{1\}$
- $\{00,10,11,01\}$  i.e.  $\{00\} \cup \{10\} \cup \{11\} \cup \{01\}$
- $\{001\}\{0,1\}$  i.e.  $\{001\}\{0\} \cup \{001\}\{1\}$
- $\{111\}^*$  i.e.  $\{\lambda, 111, 111111, 111111111, \dots\}$
- $\{1, \epsilon\}\{001\}$  i.e.  $\{1001\} \cup \{001\}$



# Examples of RE

---

Let  $\Sigma = \{0, 1\}$  then RE over inputs are

- $\varepsilon$  (empty)
- $\phi$  (null)
- 0
- 001
- $0+1$
- $00+10+11+01$
- $001(0+1)$
- $(111)^*$
- $(1+\varepsilon)001$



## Note on Regular Expression (RE)

---

- If  $r$  is a set of strings,  $r^*$  denotes the set of all strings formed by concatenating zero or more strings from  $r$ .
- If  $r$  is a set of strings,  $r^+$  denotes the set of all strings formed by concatenating one or more strings from  $r$  or  $r^+ = ((r^*)r) = (r(r^*))$ .
- We can neglect the parentheses assuming that  $*$  has a higher precedence than concatenation and concatenation has a higher precedence than  $+$





---

□  $01^* + 1$  is grouped as  $((0(1^*))) + 1$

Language : string 1 plus all strings consisting of a 0 followed by any no of 1's.

□  $(01)^* + 1$  - Language : string 1 plus all strings of 01 , zero or more times.

□  $0(1^*+1)$  – Language : set of strings that start with 0 and followed by any no of 1's



# More Examples of RE

---

- all strings ending with "011"
- $(0+1)^*011$
- all strings with no "0" after "1"
- $0^*1^*$
- all strings with at least one "0" and one "1", and no "0" after "1"
- $00^*11^*$  or  $0^+1^+$
- all strings with an even number of 0's followed by odd number of 1's.
- $(00)^*(11)^*1$
- all string containing "00" as substring.
- $(0+1)^*00(0+1)^*$



# Class Discussion

---

What languages do the following RE represent?

- $(1+0)^*(01+110)$
- $11^*(0+1)$
- $(0+1+\epsilon)$

# Context-free grammar: definition



---

**Definition.** A **context-free grammar**

$G = (N, T, P, S)$  consists of

- a finite set  $N$  of **non-terminal / variable symbols**;
- a finite set  $T$  of **terminal symbols** not in  $N$ ;
- a finite set  $P$  of **production rules** of the form  
 $U \rightarrow v$ , where  $U$  is in  $N$  and  $v$  is a string in  $(T \cup N)^*$
- a **start symbol**  $S$  in  $N$ .



# Example

---

The grammar  $G$  with non-terminal symbols

$N = \{S\}$ , terminal symbols  $= \{a, b\}$ , and productions

$$S \rightarrow aSb$$
$$S \rightarrow ba$$

Following a common practice, we use capital letters for non-terminal symbols and small letters for terminal symbols.



## Example :

$G = (\{S\}, \{a, b\}, S, P)$  with productions

$S \rightarrow aSa$  ,

$S \rightarrow bSb$

$S \rightarrow \epsilon$

- This is context free as well as linear.
- $S \rightarrow aSa \rightarrow aaSaa \rightarrow aabSbaa \rightarrow aabbbaa$
- $L(G) = \{ww^R : w \in \{a,b\}^*\}$

# Exercise



---

Give a CFG for the following languages:

- $L = \{ a^n b^n \mid n \geq 0 \}$
- $L = \{ x \mid \text{even number of } a\text{'s and } b\text{'s} \}$
- $L = \{ x \mid x \text{ is a palindrome over } \{a, b\}^* \}$