Alphabets & Languages

Kun-Mao Chao

Alphabet: a finite set of symbols.

Σ

string: a member in Σ^* empty string: e

concatenation $x \circ y$ or xy $\begin{cases} w^0 = e \\ w^{i+1} = w^i \circ w \text{ for } i \geq 0 \end{cases}$

substring

suffix

prefix

$\boxed{\begin{aligned} &(\text{stressed})^R \\ &= d(\text{stresse})^R \\ &= de(\text{stress})^R \\ &\quad\vdots \\ &= \text{desserts} \end{aligned}}$

reversal: $w = e \Rightarrow w^R = w$

w^R $|w| = n+1 > 0$ $w = ua \Rightarrow w^R = au^R$.

for some $a \in \Sigma$

$w, x \in \Sigma^*$, $(wx)^R = x^R w^R$.

pf. Basis: $|x| = 0 \Rightarrow x = e \Rightarrow (wx)^R = w^R = e^R w^R = x^R w^R$

Induction Hypothesis: If $|x| \leq n$, then $(wx)^R = x^R w^R$.

Induction Step. $|x| = n+1$. $x = ua$, for some $u \in \Sigma^*$ and $a \in \Sigma$ $(|u| = n)$

$$\begin{aligned} (wx)^R &= (w(ua))^R \\ &= (wua)^R \\ &= a(wu)^R \\ &= au^R w^R \\ &= x^R w^R \end{aligned}$$

Language: any subset of Σ^* Kun-Mao Chao

i.e. any set of strings over Σ

e.g. $\{orz, crz\}$ is a language over $\{a, \cdots, z\}$

$\{0, 01, 011, 0111, \cdots\}$

$\{w \in \{0,1\}^* : w$ has an equal number of 0's and 1's$\}$

$$L = \{w \in \Sigma^* : w \text{ has property } P\}.$$

If Σ is a finite alphabet, then Σ^* is countably infinite.

However, $\underbrace{2^{\Sigma^*}}_{\text{the set of all possible languages}}$ is uncountably infinite. Not all languages can be represented in Σ^*.

The complement of L: $\overline{L} = \Sigma^* - A$

concatenation: $L = L_1 \circ L_2 = L_1 L_2$ where

$L = \{w \in \Sigma^* : w = x \circ y$ for some $x \in L_1$ & $y \in L_2\}$.

Kleene star: $L^* = \{w \in \Sigma^* : w = w_1 \circ \cdots \circ w_k$ for some $k \geq 0$ and some $w_1, w_2, \cdots, w_k \in L\}$

Kun-Mao Chao

- $L = \{01, 1, 100\}$

  $110001110011 \in L^*$

- $L = \{w \in \{0,1\}^* : w$ has an unequal number of 0's and 1's $\}$.

  What is L^*?

  Is $10111 \in L^*$? Yes! $1 \circ 0 \circ 1 \circ 1 \circ 1$

  $\{0,1\} \subseteq L \Rightarrow \{0,1\}^* \subseteq L^*$ $\{0,1\}^*$

  On the other hand, $L^* \subseteq \Sigma^*$ by definition.

  We have $L^* = \{0,1\}^*$.

$L^+ = LL^* = \{w \in \Sigma^* : w = w_1 \circ \cdots \circ w_k$ for some $k \geq 1$ and some $w_1, \ldots, w_k \in L\}$.