

# 字母表 (*Alphabet*)

◇ 概念      形式符号的非空有限集合

◇ 记号      常用  $\Sigma$  表示

◇ 举例

- 英文字母表  $\{ a, b, \dots, z, A, B, \dots, Z \}$
- 英文标点符号表  $\{ , ; : . ? ! ' " " ( ) [ ] - - \dots \}$
- 汉字表  $\{ \dots, \text{自}, \dots, \text{动}, \dots, \text{机}, \dots \}$
- 化学元素表  $\{ \text{H}, \text{He}, \text{Li}, \dots, \text{Une} \}$
- $\Sigma = \{ a, n, y, \text{任,意} \}$

◇ 概念 字母表  $\Sigma$  上的一个字符串 (串), 或称为字 (*word*), 为  $\Sigma$  中字符构成的一个有限序列。空串 (*empty string*), 常用  $\varepsilon$  表示, 不包含任何字符。

举例 设  $\Sigma = \{ a, b \}$ , 则  $\varepsilon, a, aaa, baba$  等都是串

◇ 字符串  $w$  的长度, 记为  $|w|$ , 是包含在  $w$  中字符的个数

举例  $|\varepsilon| = 0, |bbaba| = 5$

## ✧ 连接 (*concatenation*)

设  $x, y$  为串, 且  $x = a_1 a_2 \dots a_m$ ,  $y = b_1 b_2 \dots b_n$ ,  
则  $x$  与  $y$  的连接

$$xy = a_1 a_2 \dots a_m b_1 b_2 \dots b_n$$

## ✧ 连接运算的性质

- $(xy)z = x(yz)$
- $\varepsilon x = x \varepsilon = x$
- $|xy| = |x| + |y|$

## ◇ 幂运算

设  $\Sigma$  为字母表,  $n$  为任意自然数, 定义

$$(1) \quad \Sigma^0 = \{ \varepsilon \}$$

$$(2) \quad \text{设 } x \in \Sigma^{n-1}, a \in \Sigma, \text{ 则 } ax \in \Sigma^n$$

$$(3) \quad \Sigma^n \text{ 中的元素只能由 (1) 和 (2) 生成}$$

举例 设  $\Sigma = \{ 0, 1 \}$ , 则

$$\Sigma^0 = \{ \varepsilon \}, \quad \Sigma^1 = \{ 0, 1 \},$$

$$\Sigma^2 = \{ 00, 01, 10, 11 \}, \quad \dots$$

✧ \* 闭包  $\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \dots$

✧ + 闭包  $\Sigma^+ = \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \dots$

✧  $\Sigma^* = \Sigma^+ \cup \{ \varepsilon \}$ ,  $\Sigma^+ = \Sigma^* - \{ \varepsilon \}$  ?

✧ 举例 设  $\Sigma = \{ 0, 1 \}$ , 则

$$\Sigma^+ = \{ 0, 1, 00, 01, 10, 11, \dots \}$$

$$\Sigma^* = \{ \varepsilon, 0, 1, 00, 01, 10, 11, \dots \}$$

◇ 概念 设  $\Sigma$  为字母表，则任何集合  $L \subseteq \Sigma^*$  是字母表  $\Sigma$  上的一个语言 (language)

◇ 举例

- 英文单词集  $\{ \dots, \text{English}, \dots, \text{words}, \dots \}$
- C++ 语言程序集  $\{ \dots \}$       字母表?
- 汉语四字成语集  $\{ \dots, \text{语不惊人}, \dots, \text{言之有物}, \dots \}$
- 部分化学分子式集  $\{ \dots, \text{H}_2\text{O}, \dots, \text{NaCl}, \dots \}$
- $\{ \text{any}, \text{任意} \}$

◇ 比较 空语言  $\emptyset$  与仅含空字的语言  $\{ \varepsilon \}$

✧ 两个语言  $L$  和  $M$  的并 (*union*)

$$L \cup M = \{ w \mid w \in L \vee w \in M \}$$

✧ 举例

设  $L = \{ 001, 10, 111 \}$ ,  $M = \{ \varepsilon, 001 \}$ , 则

$$L \cup M = \{ \varepsilon, 10, 001, 111 \}$$

✧ 两个语言  $L$  和  $M$  的连接 (concatenation)

$$L \cdot M = \{w_1w_2 \mid w_1 \in L \wedge w_2 \in M\}$$

✧ 通常记  $L \cdot M$  为  $LM$

✧ 举例

设  $L = \{001, 10, 111\}$ ,  $M = \{\varepsilon, 001\}$ , 则

$$LM = \{001, 10, 111, 001001, 10001, 111001\}$$



## ✧ 语言 $L$ 的闭包 (closure)

$$L^* = L^0 \cup L^1 \cup L^2 \cup \dots = \bigcup_{i \geq 0} L^i, \text{ 其中}$$

$$L^0 = \{\varepsilon\}, \quad L^1 = L, \quad L^2 = LL, \dots$$

$$L^n = L^{n-1}L$$

## ✧ 举例

设  $L = \{0, 11\}$ , 则

$$L^* = \{\varepsilon, 0, 11, 00, 011, 110, 1111, 000, 0011, \\ 0110, 01111, 1100, 11011, 11110, 111111, \dots\}$$