$\Sigma = \{a_1, a_2, \dots, a_k\}$ alfabeto

Σ^* = insieme di parole

$\cdot$ = prodotto di concatenazione

$(\Sigma^*, \cdot, \varepsilon)$ è monoide
non è commutativo

Un linguaggio formale su Σ : $L \subseteq \Sigma^*$

L
- finiti $\emptyset$, $\{\varepsilon\}$, $L_n = \{a^i \mid i \leq n\}$
- infiniti $L = \{\varepsilon, a^1, a^2, \dots, a^n, \dots\} = \{a\}^*$

Operazione sui linguaggi

Insiemistiche	tipiche
$A \cup B$	$\cdot$
$A \cap B$	$*$, $+$
A^c	L^k

Dati A e B linguaggi su Σ

- UNIONE:

$$A \cup B = \{w \in \Sigma^* \mid w \in A \vee w \in B\}$$

- INTERSEZIONE

$$A \cap B = \{w \in \Sigma^* \mid w \in A \wedge w \in B\}$$

- COMPLEMENTO

$$A^c = \{ w \in \Sigma^* \mid w \notin A \}$$

Operazioni tipiche sui linguaggi

- Prodotto ·

A, B linguaggi su Σ il loro prodotto è

$$A \cdot B = \{ w \in \Sigma^* \mid w = xy \text{ con } x \in A \text{ e } y \in B \}$$

$$= \{ xy \in \Sigma^* \mid x \in A \text{ e } y \in B \}$$

non si può invertire xy con yx

- Potenza (alla k) $L \subseteq \Sigma^*$

$$L^k = \underbrace{L \cdot L \cdot L \cdot \ldots \cdot L}_{k \text{ volte}}$$

<u>es</u>: $L^3 \Rightarrow$ se $w \in L^3$ $w = xyz$ dove $x, y, z \in L$

<u>Def.</u> ricorsiva di L^k

$$\begin{cases} L^0 = \{\varepsilon\} \\ L^k = L \cdot L^{k-1} \end{cases}$$

- CHIUSURA DI KLEENE : $*, +$

$$L^* = L^0 \cup L^1 \cup L^2 \cup \dots \cup L^k \cup L \dots = \bigcup_{k=0}^{\infty} L^k$$

$$L^+ = L^1 \cup L^2 \cup \dots \cup L^k \cup \dots = \bigcup_{k=1}^{\infty} L^k$$

ESERCIZI

- $L^+ \stackrel{?}{=} L^* \setminus \{\varepsilon\}$

- $A = \{mo, se\}$ $\quad B = \{ra, re\}$

$A \cdot B = \{mora, more, sera, sere\}$

$B \cdot A = \{ramo, rase, remo, rese\}$

$$A \cdot B \neq B \cdot A$$

- Numeri binari : $1\{0,1\}^*$

$A = \{1\}$ $\quad B = \{0,1\}^*$

$A \cdot B = \{1\} \cdot \{0,1\}^* = 1\{0,1\}^*$

abuso di notazione

- $L = \{a,b\}$ su $\Sigma = \{a,b\}$

$L = \Sigma$

$L^k = \Sigma^k = \underbrace{\Sigma \cdot \Sigma \cdot \dots \Sigma}_{k \text{ volte}}$

$$\Sigma^k = \{ w \in \{a,b\}^* \mid |w| = k \}$$

- Σ^* = insieme delle parole su Σ inclusa ε ①

$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \dots \Sigma^k \cup \dots$

② $= \bigcup_{k=0}^{\infty} \Sigma^k$

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

$\Sigma^k = \{ w \in \{a, b\}^* \mid |w| = k \}$

= ok si verifica che le due scritture sono uguali

1 = 2

- $A = \{ bb \}$ su $\Sigma = \{b\}$

$A^* = \{ w \in \{b\}^* \mid |w| = 2n \quad n \geq 0 \}$
$= \{ b^{2n} \mid n \in \mathbb{N} \}$

- $\bar{A} = \{ b, bb \} \qquad \Sigma = \{b\}$

$\bar{A}^* = \{ b^n \mid n \in \mathbb{N} \} = \{ b \}^*$

$\bar{A}^+ = \{ b \}^* \setminus \{ \varepsilon \} = \{ b \}^+$

$bbb \in \bar{A}^+$ si può scomporre in parole di $\bar{A}$ in questo modo:

$$\left.\begin{array}{l} bb \cdot b \\ b \cdot bb \\ b \cdot b \cdot b \end{array}\right\}$$ $\bar{A}$ non è un codice

Codice

Definizione: Un linguaggio L è un codice quando ogni parola in L^+ è decomponibile in un unico modo in parole di L

es $L = \{ aa, ab, b \}$

parola in L^+:

$$\underbrace{ab}\underbrace{ba}\underbrace{a}\underbrace{ab}\underbrace{b} \in L$$

si ha che L è un codice

Definizione: L è un codice PREFISSO o ISTANTANEO quando è un codice per cui ogni parola di L NON è prefisso di altre parole di L

Es $\{aa, ab, b\}$ è un codice ed è prefisso

$\{0, 01\}$ è un codice ma NON è prefisso perché 0 è prefisso di 01

Codice ASCII esteso

codifica ogni carattere della tastiera:

$$\{A, \dots, Z, a, \dots z, 0, 1, \dots, 9, \dots, @, \{, |, \dots\}$$

in sequenza di 8 bit

$$C_A = \{ x \in \{0, 1\}^* \mid |x| = 8 \}$$

ogni carattere è codificato con 8 bit

C_A^+ = "File binario"

Perché è importante che C_A sia un codice?