Azzolini Riccardo 2020-09-23

Preliminari

1 Alfabeti e stringhe

Un **alfabeto** Σ è un *insieme finito e non vuoto di simboli*, dove un simbolo è un "oggetto qualunque": può essere un carattere, una sequenza di caratteri (come ad esempio una parola chiave di un linguaggio di programmazione), ecc.

Una **stringa** (o **parola**) su Σ è una qualunque *sequenza finita di simboli di* Σ (si considerano solo sequenze finite perché quelle infinite non potrebbero essere analizzate, trattate per intero, dunque risulterebbe impossibile assegnarvi un significato). In particolare, è una stringa anche la sequenza vuota, composta da zero simboli di Σ, che prende il nome di **stringa vuota** e viene indicata con ϵ.

Per convenzione, si indicheranno

- i simboli di Σ con le lettere latine minuscole della parte iniziale dell'alfabeto, eventualmente indiciate:

  $$a, b, c, \dots, a_0, a_1, \dots, a_n, \dots$$

  (ma, in alcuni esempi, si indicheranno invece i simboli con delle parole aventi un significato intuitivo nel contesto di tali esempi);

- le stringhe su Σ con le lettere latine minuscole della parte finale dell'alfabeto, oppure con le lettere greche minuscole, eventualmente indiciate:

  $$u, v, w, x, y, z, \alpha, \beta, \gamma, \delta, \dots, \alpha_0, \alpha_1, \dots, \alpha_n, \dots$$

Se $a_1, \dots, a_n \in \Sigma$ (con $n \geq 0$), si scrive

$$w = a_1 \dots a_n$$

per indicare la stringa costituita da n simboli di Σ che ha a_i come i-esimo simbolo della sequenza (per $i = 1, \dots, n$). Come caso particolare, se $n = 0$ si ha $w = \epsilon$.

Data una stringa w, si indica con $|w|$ la sua **lunghezza**, definita come il numero di occorrenze di simboli di Σ in w. La lunghezza della stringa vuota è $|\epsilon| = 0$.

1.1 Esempi di stringhe

Sia $\Sigma = \{0, 1\}$. I seguenti sono alcuni esempi di stringhe su Σ, con le rispettive lunghezze:

w	$\|w\|$
ϵ	0
0	1
1	1
00	2
11	2
01	2
10	2
010011	6

2 Concatenazione di stringhe

Date due stringhe α e β su Σ, la **concatenazione** di α con β, indicata con $\alpha\beta$, è la stringa costituita dai simboli di α seguiti dai simboli di β. Ad esempio, considerando ancora l'alfabeto $\Sigma = \{0, 1\}$, date $\alpha = 01$ e $\beta = 11$ si ha $\alpha\beta = 0111$.

Formalmente, la concatenazione è un *operatore binario* sulle stringhe, in quanto "prende" due stringhe e produce come risultato una nuova stringa. Si osserva che ϵ è l'*elemento neutro* di tale operatore: per ogni stringa α, $\epsilon\alpha = \alpha\epsilon = \alpha$.

3 Prefissi e postfissi

Siano α e β due stringhe su Σ.

- β è un **prefisso** di α se esiste una stringa γ tale che $\alpha = \beta\gamma$ (cioè, informalmente, se i simboli di β costituiscono la parte iniziale di α).
- β è un **postfisso** di α se esiste una stringa γ tale che $\alpha = \gamma\beta$ (ovvero se i simboli di β costituiscono la parte finale di α).

Si osserva che la stringa vuota, essendo l'elemento neutro della concatenazione,

$$\alpha\epsilon = \epsilon\alpha = \alpha$$

è sia un prefisso sia un postfisso di α.

4 Insiemi di stringhe

- La **potenza k-esima** dell'alfabeto Σ (con $k \geq 0$), indicata con Σ^k, è l'insieme delle stringhe di lunghezza k su Σ:

$$\Sigma^k = \{w \mid w \text{ è una stringa su } \Sigma \text{ e } |w| = k\}$$

- L'**insieme di tutte le stringhe** su Σ, indicato con Σ^* (che si legge "sigma star"), è definito come:

$$\Sigma^* = \bigcup_{k \geq 0} \Sigma^k$$

  Si osserva che tale insieme è sempre infinito, perché contiene le stringhe di tutte le infinite lunghezze possibili. Ad esempio, anche nel caso estremo di un alfabeto composto da un unico simbolo, $\Sigma = \{a\}$, si ha $\Sigma^* = \{\epsilon, a, aa, aaa, \dots\}$.

- L'**insieme di tutte le stringhe non vuote** su Σ, indicato con Σ^+ (che si legge "sigma più"), è:

$$\Sigma^+ = \bigcup_{k \geq 1} \Sigma^k = \{w \in \Sigma^* \mid |w| > 0\}$$

  Si osserva dunque che $\Sigma^* = \Sigma^+ \cup \{\epsilon\}$.

- Un **linguaggio** sull'alfabeto Σ è un qualunque insieme di stringhe su Σ, cioè un qualunque insieme $L \subseteq \Sigma^*$. Si osserva che un linguaggio può essere finito o infinito.

4.1 Esempi

Dato l'alfabeto $\Sigma = \{0, 1\}$, si hanno:

$$\Sigma^0 = \{\epsilon\} \quad \Sigma^1 = \{0, 1\} \quad \Sigma^2 = \{00, 01, 10, 11\} \quad \dots$$
$$\Sigma^* = \{\epsilon, 0, 1, 00, 01, 10, 11, \dots\}$$
$$\Sigma^+ = \{0, 1, 00, 01, 10, 11, \dots\}$$

Alcuni esempi di linguaggi su Σ sono allora:

- $L_1 = \{\epsilon\}$, il linguaggio formato solo dalla stringa vuota;
- $L_2 = \varnothing$, il linguaggio vuoto, che non comprende alcuna stringa;
- $L_3 = \Sigma^*$, il linguaggio infinito formato da tutte le stringhe su Σ;
- $L_4 = \{w \in \Sigma^* \mid w \text{ termina con } 0\}$, il linguaggio infinito formato dalle stringhe che terminano con 0 (ovvero il linguaggio dei numeri naturali pari, se si interpretano le stringhe su Σ come rappresentazioni binarie dei numeri naturali).