Documenti XML

Sono composti da testo + tag

tag = marcatore di testo che da informazioni semantiche al testo

es <rubrica>

<nome> Bianchi </nome>

<tel> 023456710 </tel>

...

</rubrica>

Notazione: tag aperto <t> ← t

tag chiuso </t> ← $\bar{t}$

condizioni:

Un documento XML deve soddisfare:

① deve $\exists$ un tag che contiene tutto il doc.

$$S \ldots\ldots \bar{S}$$

② Ogni tag aperto → tag chiuso

③ i tag devono essere innestati correttamente

$$x \ldots y \ldots \bar{y} \ldots \bar{x}$$

Definizione: Un documento XML é CORRETTO se valgono ①, ②, ③

Oltre ad ①, ② e ③, il documento deve soddisfare anche un DTD (= grammatica tipo 2)

DTD = definisce come i tag possono essere innestati tra loro

es nel doc rubrica

<tel> non può stare dentro a <nome>

Definizione: Un documento XML é VALIDO secondo un certo DTD se é generato da quel DTD
" grammatica di tipo 2

Forma delle regole per i DTD

$$A \to a \; R_a \; \bar{a}$$

R_a = espressione di variabili con $\overset{\cup}{+}, \cdot, *$

R_a é trasformabile in regole di tipo 2

esempio di DTD:

$$S \to s\, C^* \,\bar{s}$$

$$C \to c\, M N (E V)^* \,\bar{c}$$

$M \to m\bar{m}$

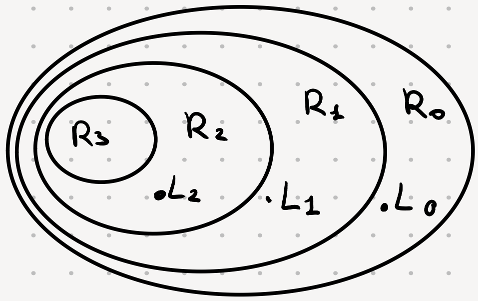
$N \to n\bar{n}$

$E \to e\bar{e}$

$V \to v\bar{v}$

Teorema sugli R_k:

$R_3 \subset R_2 \subset R_1 \subset R_0$



"$\subset$" = sottoinsieme di $\neq$ "$\subseteq$"

dim:

- L'inclusione tra gli R_k segue dal fatto che abbiamo detto sui tipi di G:

$$\text{Tipo } k \Rightarrow \text{tipo } k-1$$

- dimostriamo che l'inclusione è propria

- esiste $L_2 \in R_2$ ma $L_2 \notin R_3$

$L_2 = \{a^n b^n \mid n > 0\}$

Infatti $a^n b^n$ ammette G di tipo 2

$$S \to aSb, \quad S \to ab$$

e inoltre vedremo che $a^n b^n \notin R_3$

perché non ammette un automa a stati finiti che lo riconosce

- esiste $L_1 \in R_1$ ma $L_1 \notin R_2$

$$L_1 = \{ a^n b^n c^n \mid n > 0 \}$$

Infatti $a^n b^n c^n$ ammette G di tipo 1 e inoltre vedremo che $a^n b^n c^n \notin R_2$ perché non soddisfa il pumping lemma per i linguaggi liberi da contesto.

- esiste $L_0 \in R_0$ ma $L_0 \notin R_1$

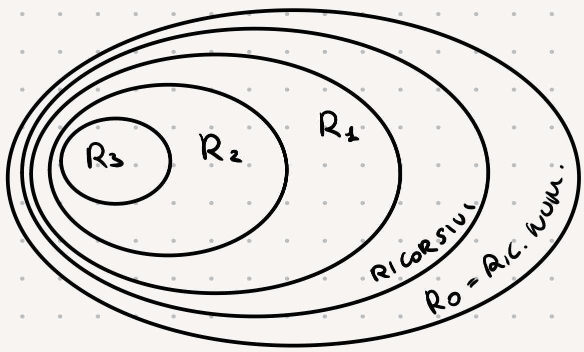
$$L_0 = D = \{ x \in \{0,1\}^* \mid F_u(x \$ x) \downarrow \}$$

dim:

I passo: $R_1 \subseteq$ Ricorsivi

II passo: D non è ricorsivo $\Rightarrow D \notin R_1$

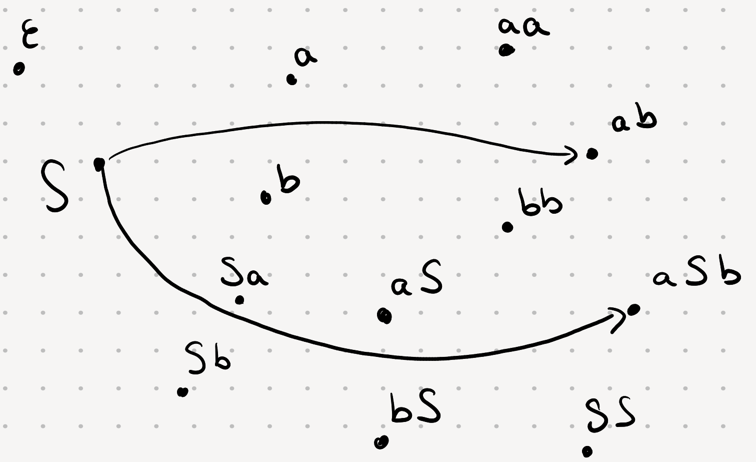
D è ric. enum. $\Rightarrow D \in R_0$



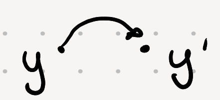
Definizione di $GR(x)$ con $x \in \Sigma^*$

$$GR(x) = \langle V_x, E_x \rangle \text{ dove:}$$

$$V_x = \{ y \in (\Sigma \cup M)^* \mid |y| \leq |x| \}$$



$$E_x = \{ (y, y') \mid y \Rightarrow y' \}$$



Algoritmo w $(x \in \Sigma^*)$ {

costruisci $GR(x)$ ← serve G

IF (esiste un cammino $S \rightsquigarrow x$)

then return 1
else return 0

}

Oss: Il problema della generazione di x si trasforma nella ricerca di un cammino

in un grafo.
Il tempo dell'algoritmo w é esponenziale,
richiesto da costruisci $GR(x)$

Correttezza di w:

$x \in L(G) \Rightarrow$ esiste una derivazione in G

$S \Rightarrow z_1 \Rightarrow z_2 \Rightarrow \dots \Rightarrow z_k = x$ t.c.

per ogni i $\quad |z_i| \leq |z_{i+1}| \Rightarrow$
per ogni i $\quad |z_i| \leq |x|$ e quindi
per ogni i $\quad z_i \in V_x$ inoltre $z_i \rightarrow z_{i+1}$

$\Rightarrow$ nel grafo $GR(x)$ si ha

$$S \rightarrow z_1 \rightarrow z_2 \rightarrow z_3 \quad \dots \rightarrow z_k = x$$

$\Rightarrow \underline{F_w(x) = 1}$

$x \in L(G) \Rightarrow F_w(x) = 0$

$\Updownarrow$

$F_w(x) = 1 \Rightarrow x \in L(G)$

$F_w(x) = 1 \Rightarrow$ esiste un cammino

$S \rightarrow z_1 \rightarrow z_2 \dots \rightarrow z_k = x \Rightarrow$ in G esistono i seg. passi

$S \Rightarrow z_1, \; z_1 \Rightarrow z_2, \dots, z_{k-1} \Rightarrow z_k = x$

$\Rightarrow S \Rightarrow z_1 \Rightarrow z_2 \Rightarrow \dots \Rightarrow z_k = x \quad \Rightarrow x \in L(G)$