Grammatiche context-free

Proprieta' di chiusura
Proprieta' del pompaggio

Proprieta' di chiusura

Siano $G_1 = \langle V_1, \Sigma_1, P_1, S_1 \rangle$ e $G_2 = \langle V_2, \Sigma_2, P_2, S_2 \rangle$ due grammatiche libere con insiemi di non terminali disgiunti (condizione che si puo' sempre ottenere) e S un nome "fresh":

- l'unione di due linguaggi liberi e' un linguaggio libero:

  $$G_U = \langle V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, P_1 \cup P_2 \cup \{S \to S_1 \mid S_2\}, S \rangle$$

- il concatenamento di due linguaggi liberi e' un linguaggio libero:

  $$G_C = \langle V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, P_1 \cup P_2 \cup \{S \to S_1 S_2\}, S \rangle$$

- la chiusura di Kleene di un linguaggio libero e' un linguaggio libero:

  $$G_* = \langle V_1 \cup \{S\}, \Sigma_1, P_1 \cup \{S \to S\, S_1 \mid \varepsilon\}, S \rangle$$

  $$G_* = \langle V_1 \cup \{S\}, \Sigma_1, P_1 \cup \{S \to S_1\, S \mid \varepsilon\}, S \rangle$$

- la riflessione di un linguaggio libero è un linguaggio libero

  $$G_R = \langle V_1, \Sigma_1, P_1^R, S \rangle$$

  dove P_1^R è l'insieme delle produzioni ottenute da quelle di P_1 scrivendo al contrario i membri destri.

Proprietà “pumping” I

Sappiamo che un linguaggio generato da una grammatica G senza variabili inutili e priva di derivazioni circolari (pulita), è infinito se e solo se ha una derivazione ricorsiva. Consideriamo queste derivazioni

1 - $S \Rightarrow^* \alpha_2 A \beta_2 \Rightarrow^* xAy \Rightarrow^* xwy$ (assumo $A \Rightarrow^* w$ *non* ricorsiva)

2 - $A \Rightarrow^* \alpha_1 A \beta_1 \Rightarrow^* uAv \Rightarrow^* uwv$ $(uv \neq \varepsilon)$ (considero una ricorsione su A)

Mettendo insieme le derivazioni 1 e 2:

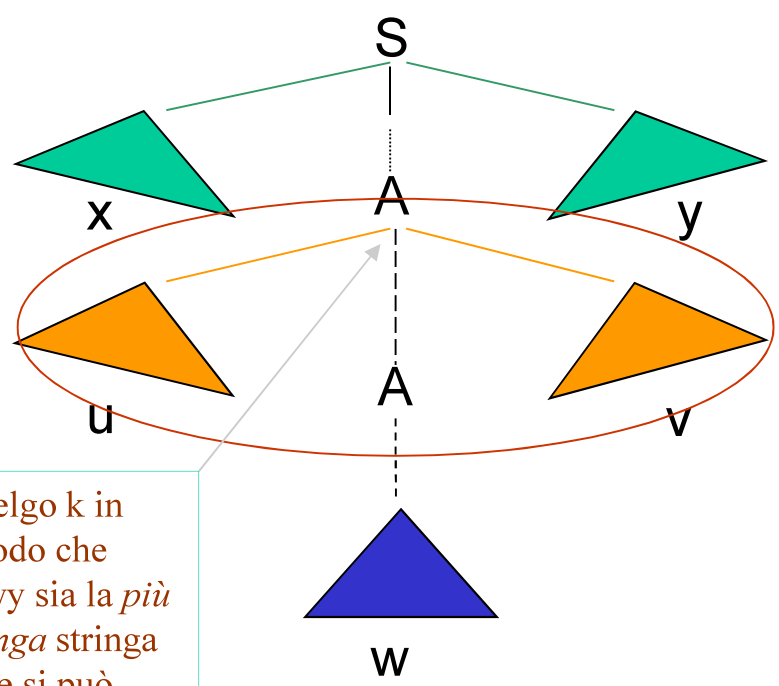
1

$$S \Rightarrow^* \alpha_2 A \beta_2 \Rightarrow^* xAy \Rightarrow^* xuAvy \Rightarrow^* xuuAvvy \Rightarrow^*$$

$$\Rightarrow^* xu^iAv^iy \Rightarrow^* xu^iwv^iy$$

Sia k la *massima* lunghezza di una stringa che si può derivare in G da una variabile con 1 sola ricorsione. Allora $|uwv| \leq k$
Si può visualizzare la derivazione nel modo seguente:

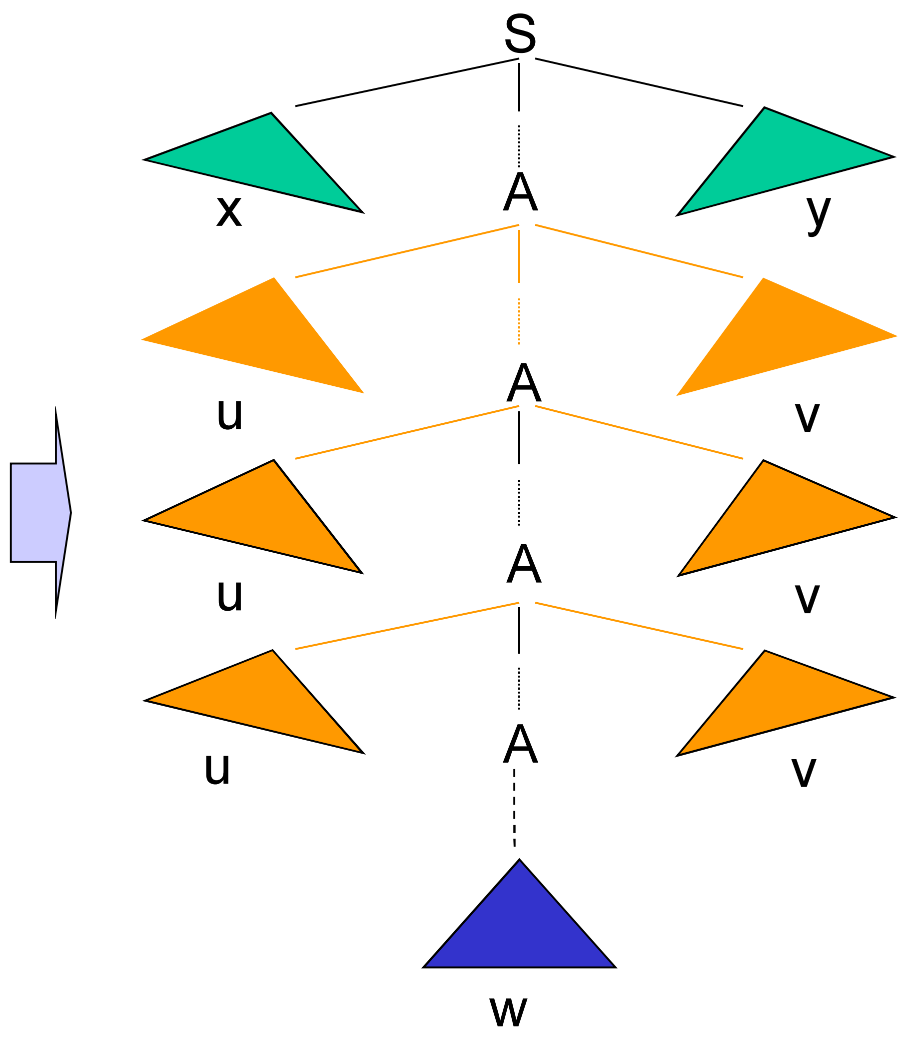
Proprietà “pumping” II



scelgo k in modo che uwy sia la *più lunga* stringa che si può generare con una sola ricorsione

x u w v y

dove $|uwv| \leq k$ e $uv \neq \varepsilon$



$x\, u^i\, w\, v^i\, y$

Proprietà “pumping” III

La costante k dipende ovviamente dalla grammatica G.

Se considero un *linguaggio* libero posso prendere, tra *tutte* le grammatiche equivalenti che lo generano, il k *più piccolo*.

Proprietà P1 (pumping lemma per linguaggi CF)

Per ogni linguaggio context-free L *esiste* una costante caratteristica k>0, che dipende solo da L, tale che ogni frase $z \in L$, di lunghezza maggiore o uguale a k (cioè $|z| \geq k$), si può scrivere come la concatenazione di cinque sottostringhe, $z = xuwvy$ con:

- $|uwv| \leq k$,
- $uv \neq \varepsilon$ e

tali che $xu^iwv^iy \in L$ *per ogni* $i \geq 0$.

Per esempio (esercizio): $\{0^n1^n2^n \ / \ n > 0\}$ non è libero dal contesto

Nota: il pumping lemma vale banalmente per i linguaggi finiti (perchè?).

Oltre le grammatiche context-free

Il seguente linguaggio non è libero dal contesto, cioè non esiste nessuna grammatica context-free che lo genera.

$L = \{0^n 1^n 2^n \ / \ n > 0\}$

Le grammatiche libere, mentre riescono a verificare che due caratteri siano ripetuti lo stesso numero di volte, non sono in grado di generare stringhe con tre gruppi di caratteri ripetuti lo stesso numero di volte.

Una grammatica che genera L e' la seguente, in cui le riscritture non sono definite per le variabili, ma per le variabili "in un contesto".

$< \{S, B, C\}, \{0, 1, 2\}, P, S >$ dove $P = \{ \ S \rightarrow 0SBC \mid 01C,$

$CB \rightarrow BC,$

$1B \rightarrow 11,$

$1C \rightarrow 12,$

$2C \rightarrow 22 \ \}$

Classificazione di Chomsky

Chomsky ha introdotto una classificazione “storica” dei linguaggi, disringuendoli in tipo 3 (regolari), tipo 2 (context free), tipo 1 (contestuali) e tipo 0 (ricorsiamente enumerabili).