pumping lemma per CFL

La forma normale di Chomsky (CNF) per le CFG prevede che tutte le produzioni siano:

A -> a, con a in T

A-> BC, con B e C variabili

non c'è ε !!

Teorema. Per ogni CFG G, esiste una CFG G' in CNF tale che L(G')=L(G)-{ε}

ripuliamo la CFG
--dai simboli inutili
--dalle produzioni unitarie, i.e. A->B
--dalla ε

un simbolo X è <u>utile</u> per una CFG se

$S \Rightarrow^* \alpha X \beta \Rightarrow^* w$

un simbolo è generatore se

$X \Rightarrow^* w$

è raggiungibile se

$S \Rightarrow^* \alpha X \beta$

prima di eliminano i non generatori e poi i non raggiungibili

restano solo i simboli utili

esempio:

S-> AB | a

A -> b

non generatori : B

resta S -> a e A -> b

ma A è diventato non raggiungibile, quindi resta

S -> a

calcolo dei simboli generatori $\prod$:
-base: i simboli terminali sono generatori, $\prod$ =T
-induzione: aggiungiamo a ∏ tutte le variabili X t.c. esiste X -> α in cui α ha solo simboli in ∏

fino a quando $\prod$ cambia

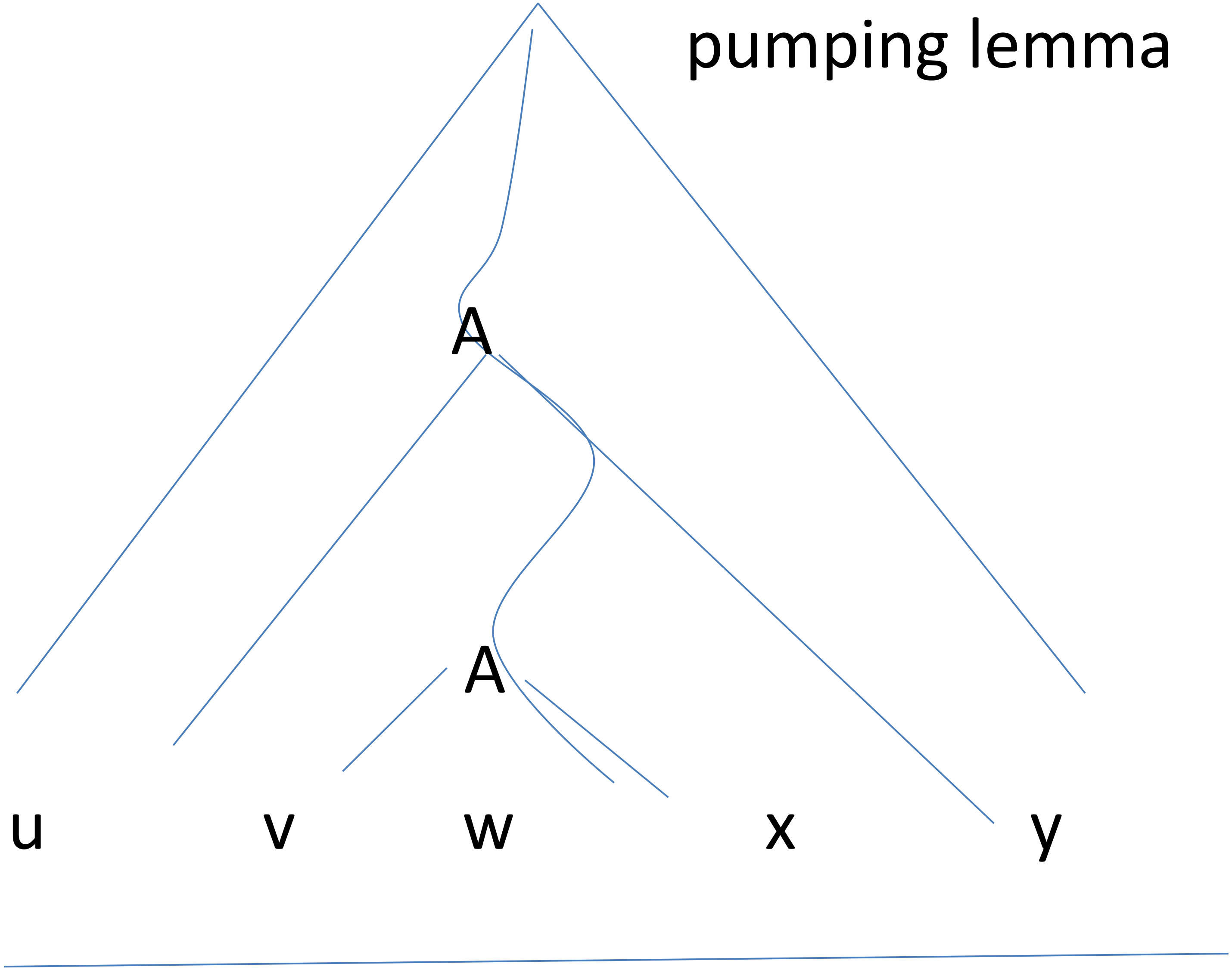
calcolo dell’insieme Z dei simboli che producono ε

base: X -> ε X in Z
induzione: aggiunguamo a Z ogni Y tale che Y-> α con i simboli di α tutti in Z

negli alberi di derivazione di una CFG in CNF è vero che se il cammino più lungo è di lunghezza n, allora il prodotto w è t.c. $|w| <= 2^{n-1}$

se m è il numero delle variabili:
un almeno con prodotto w t.c. $|w|=2^m$ deve avere un cammino di lunghezza m+1

e allora su quel cammino almeno un terminale ripete pumping lemma



$z = uvwxy$

$z = uvwxy$

ma anche uv^nwx^ny, con $n>=0$, è generabile e quindi è nel linguaggio