Es:

$L = \{a^n b^n \mid n \geq 1\}$

$L' = \{a,b\}^+$

·) is LL' regular? $\Rightarrow$ Not regular (the first part is NOT a regular language)

Pumping lemma:

$a^N b^N a$

$\downarrow$

for $k=2 \Rightarrow xy^kz = a^{N+|y|} b^n a$

$\hookrightarrow \notin LL'$

$a^N b^N b \Rightarrow xy^2z = a^{N+|y|} b^n b$

$\downarrow$

$(|y| \geq 1)$

if $|y| = 1 \Rightarrow a^{N+1} b^{N+1}$

$\notin LL'$

if $|y| > 1 \Rightarrow a^{N+|y|} b^{N+1}$

•) Is $L'L$ regular?

$$w = aa^N b^N$$

for $k=0$: $xy^kz = a^{N+1-|y|} b^N$

if $|y|=1$: $xy^kz = a^N b^N$ (There's no L') $\notin L'L$

if $|y|>1$: $xy^kz = a^{N+1-|y|} b^N$ $(N+1-|y| < N)$ $\notin L'L$

there's not even L

•) is $L'LL'$ regular? It's a trap!

$(a+b)(a+b)^* ab (a+b)(a+b)^*$ ⟹ It's regular

Automata, Languages and Computation

Chapter 7 : Properties of Context-Free Languages
Part II

Master Degree in Computer Engineering
University of Padua
Lecturer : Giorgio Satta

Lecture based on material originally developed by :
Gösta Grahne, Concordia University

Pumping lemma for CFLs

In each sufficiently long string of a CFL we can find two substrings "next to each other" that

- can be eliminated
- can be iterated (synchronously)

still resulting in strings of the language

This property can be used to prove that some languages are not CFL

Parse trees

Theorem Let G be some CFG in CNF. Let T be a parse tree for a string $w \in L(G)$. If the longest path in T has n arcs, then $|w| \leqslant 2^{n-1}$

Proof By induction on $n \geqslant 1$

Base $n = 1$. T has one leaf and one inner node (root), and represents a derivation $S \Rightarrow a$. We have $|w| = 1 \leqslant 2^{n-1} = 2^0 = 1$

Parse trees

Induction $n > 1$. T's root uses a production $S \to AB$, and we can write $S \Rightarrow AB \stackrel{*}{\Rightarrow} w = uv$, where $A \stackrel{*}{\Rightarrow} u$ and $B \stackrel{*}{\Rightarrow} v$

We are using factorization here

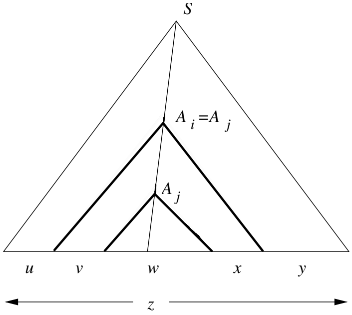
No path under the subtree rooted at A or B can have length greater than $n - 1$. By the inductive hypothesis we have $|u| \leqslant 2^{n-2}$ and $|v| \leqslant 2^{n-2}$

We can conclude that $|w| = |u| + |v| \leqslant 2^{n-2} + 2^{n-2} = 2^{n-1}$ $\square$

Pumping lemma for CFLs

Theorem Let L be some CFL. There exists a constant n such that, if $z \in L$ and $|z| \geqslant n$, we can factorize $z = uvwxy$ under the following conditions :

- $|vwx| \leqslant n$
- $vx \neq \epsilon$
- $uv^iwx^iy \in L$, for each $i \geqslant 0$



Pumping lemma for CFLs

Proof Let G be some CFG in CNF such that $L(G) = L \setminus \{\epsilon\}$. Let m be the number of variables of G. We choose $n = 2^m$

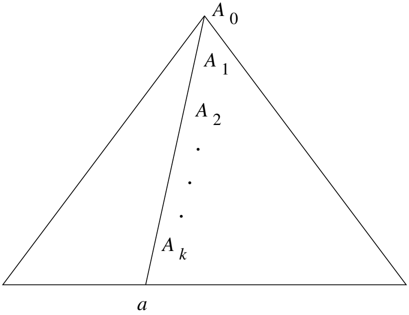
pumping lemma constant

Let $z \in L$ such that $|z| \geqslant n$

From a previous theorem, the parse tree for z must have some path of length greater than m, otherwise we would get $|z| \leqslant 2^{m-1} = n/2$

Pumping lemma for CFLs

Consider all occurrences of variables in a path of length $k+1$, where $k \geqslant m$

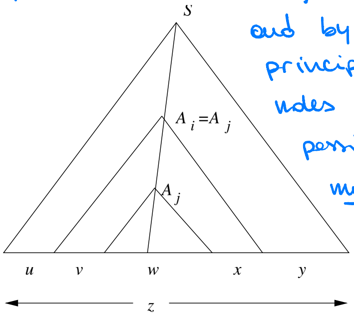


Pumping lemma for CFLs

→ pidgeon hole principle

Since G has only m variables, at least one variable occurs more than once in the path. Let us assume $A_i = A_j$, where $k - m \leqslant i < j \leqslant k$, that is, we choose A_i in the lower part of the path

↳ we can just look in the deepest m+1 nodes, since there are just m variables and by the pidgeon hole principle between m+1 nodes (with only m possible variables) there <u>must</u> be a duplicate

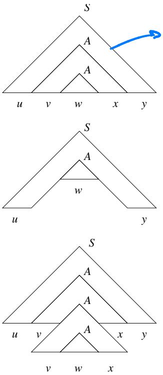


Pumping lemma for CFLs

We can then edit the parse tree in (a) in such a way that

- its yield becomes uv^0wx^0y, as shown in (b)
- its yield becomes uv^2wx^2y, as shown in (c)

If there's more than 2 equal variable, consider the two that are deeper



A generates either the big

(a) or the small tree

We can "concatenate" these trees to form different strings that belong to the language

(b)

(c)

Pumping lemma for CFLs

In the general case, we can edit the parse tree in (a) in such a way that its yield becomes uv^iwx^iy, for any $i \geqslant 0$

Since the longest path in the subtree rooted at A_i has length no longer than $m + 1$, a previous theorem allows us to assert that $|vwx| \leqslant 2^m = n$ □