Fundamentele Informatica 1 (I&E)

najaar 2015

`http://www.liacs.leidenuniv.nl/~vlietrvan1/fi1ie/`

Rudy van Vliet
kamer 124 Snellius, tel. 071-527 5777
rvvliet(at)liacs(dot)nl

college 11, dinsdag 1 december 2015

6. Context-Free and Non-Context-Free Languages

6.1. The Pumping Lemma for Context-Free Languages

6. Context-Free and Non-Context-Free Languages

6.1. The Pumping Lemma for Context-Free Languages

reg. languages	FA	reg. grammar	reg. expression
determ. cf. languages	DPDA		
cf. languages	PDA	cf. grammar	
re. languages	TM	unrestr. grammar	

A slide from lecture 3:

2.4 The Pumping Lemma

Theorem 2.29.
The Pumping Lemma for Regular Languages.
Suppose L is a language over the alphabet Σ.
If L is accepted by a finite automaton $M = (Q, \Sigma, q_0, A, \delta)$,
and if n is the number of states of M,

then for every $x \in L$ satisfying $|x| \geq n$, there are three strings u, v, and w such that $x = uvw$ and the following three conditions are true:

1. $|uv| \leq n$.
2. $|v| > 0$ (i.e., $v \neq \Lambda$).
3. For every $i \geq 0$, the string uv^iw also belongs to L.

A slide from lecture 3:

Example 2.30. The language *AnBn*.

Let $L = \{a^i b^i \mid \; i \geq 0\}$.

Now, context-free languages.

Intuitively clear that PDA cannot accept $AnBnCn$ or $XX\ldots$

Now, context-free languages.

Intuitively clear that PDA cannot accept *AnBnCn* or *XX*...

Pumping lemma based on derivation in CFG (not on PDA):

$$S \Rightarrow^* v\underline{A}z \Rightarrow^* v\ w\underline{A}y\ z \Rightarrow^* vw\ x\ yz$$

$$S \Rightarrow^* v\underline{A}z \Rightarrow^* v\ w\underline{A}y\ z \Rightarrow^* vw\ wAy\ yz \Rightarrow^* vw^m xy^m z$$

Theorem 6.1.
The Pumping Lemma for Context-Free Languages.

Suppose L is a context-free language. Then there is an integer n so that for every $u \in L$ with $|u| \geq n$, u can be written as $u = vwxyz$, for some strings v, w, x, y and z satisfying

1. $|wy| > 0$
2. $|wxy| \leq n$
3. for every $m \geq 0$, $vw^m xy^m z \in L$

Proof. . .

A slide from lecture 7:

Definition 4.29. Chomsky Normal Form

A context-free grammar is said to be in *Chomsky normal form* if every production is of one of these two types:

$$A \rightarrow BC \text{ (where } B \text{ and } C \text{ are variables)}$$
$$A \rightarrow \sigma \text{ (were } \sigma \text{ is a terminal symbol)}$$

A slide from lecture 8:

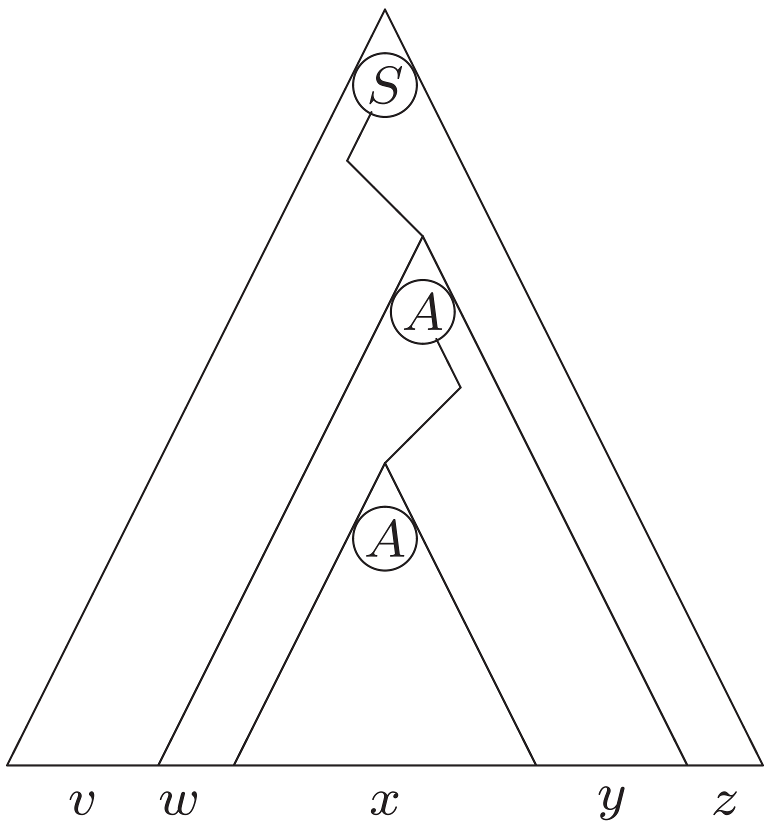
Theorem 4.30. (not Theorem 4.31!)

For every context-free grammar G,
there is another CFG G_1 in Chomsky normal form
such that $L(G_1) = L(G) - \{\Lambda\}$.

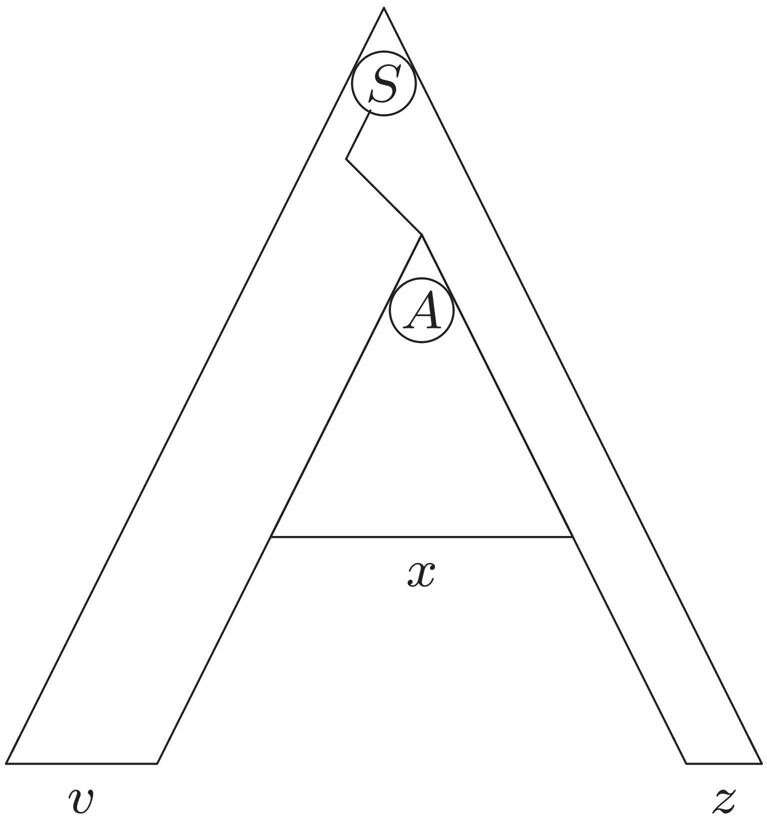
What if $\Lambda \notin L(G)$?

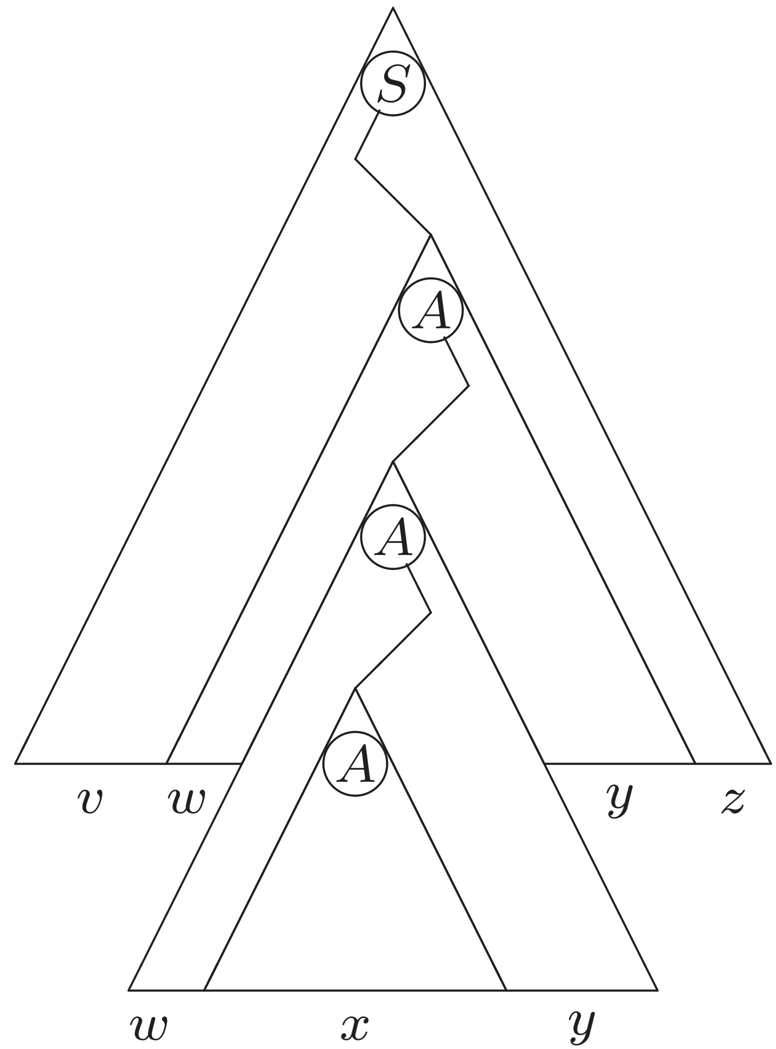
Number of leaf nodes in a binary tree of a given height

Pumping Lemma for CFLs



Pumping Lemma for CFLs





Theorem 6.1.
The Pumping Lemma for Context-Free Languages.

Proof

Let G be CFG in Chomsky normal form with $L(G) = L - \{\Lambda\}$.

Derivation tree in G is binary tree
(where each parent of a leaf node has only one child).

Height of a tree is number of edges in longest path from root to leaf node.

At most 2^h leaf nodes in binary tree of height h: $|u| \leq 2^h$.

At most 2^h leaf nodes in binary tree of height h: $|u| \leq 2^h$.

Let p be number of variables in G,
let $n = 2^p$
and let $u \in L(G)$ with $|u| \geq n$.

(Internal part of) derivation tree of u in G has height at least p. Hence, longest path in (internal part of) tree contains at least $p+1$ (internal) nodes.

Consider final portion of longest path in derivation tree.
(leaf node + $p+1$ internal nodes),
with ≥ 2 occurrences of a variable A.

Pump up derivation tree, and hence u.

Application of pumping lemma:
mainly to prove that a language L cannot be generated by a context-free grammar.

How?
Find a string $u \in L$ with $|u| \geq n$ that cannot be pumped up!

What is n?

What should u be?

What can v, w, x, y and z be?

Suppose that there exists context-free grammar G with $L(G) = L$. Let n be the integer from the pumping lemma.

Pumping lemma:

For every $u \in L$ with $|u| \geq n$,
there are five strings v, w, x, y and z such that $u = vwxyz$
and the following three conditions are true:

1. $|wy| > 0$
2. $|wxy| \leq n$
3. for every $m \geq 0$, $vw^m xy^m z \in L$

Suppose that there exists context-free grammar G with $L(G) = L$. Let n be the integer from the pumping lemma.

We prove:

NOT
(

For every $u \in L$ with $|u| \geq n$,
there are five strings v, w, x, y and z such that $u = vwxyz$
and the following three conditions are true:

1. $|wy| > 0$
2. $|wxy| \leq n$
3. for every $m \geq 0$, $vw^m xy^m z \in L$

)

Suppose that there exists context-free grammar G with $L(G) = L$. Let n be the integer from the pumping lemma.

We prove:

There exists $u \in L$ with $|u| \geq n$, such that

NOT
(

there are five strings v, w, x, y and z such that $u = vwxyz$ and the following three conditions are true:

1. $|wy| > 0$
2. $|wxy| \leq n$
3. for every $m \geq 0$, $vw^mxy^mz \in L$

)

Suppose that there exists context-free grammar G with $L(G) = L$. Let n be the integer from the pumping lemma.

We prove:

There exists $u \in L$ with $|u| \geq n$, such that
for every five strings v, w, x, y and z such that $u = vwxyz$

NOT
(

the following three conditions are true:

1. $|wy| > 0$
2. $|wxy| \leq n$
3. for every $m \geq 0$, $vw^mxy^mz \in L$

)

Suppose that there exists context-free grammar G with $L(G) = L$. Let n be the integer from the pumping lemma.

We prove:

There exists $u \in L$ with $|u| \geq n$, such that
for every five strings v, w, x, y and z such that $u = vwxyz$

NOT all of the following three conditions are true:

1. $|wy| > 0$
2. $|wxy| \leq n$
3. for every $m \geq 0$, $vw^m xy^m z \in L$

)

Suppose that there exists context-free grammar G with $L(G) = L$. Let n be the integer from the pumping lemma.

We prove:

There exists $u \in L$ with $|u| \geq n$, such that
for every five strings v, w, x, y and z such that $u = vwxyz$

if
1. $|wy| > 0$
2. $|wxy| \leq n$

then NOT
(

3. for every $m \geq 0$, $vw^m xy^m z \in L$

)

Suppose that there exists context-free grammar G with $L(G) = L$. Let n be the integer from the pumping lemma.

We prove:

There exists $u \in L$ with $|u| \geq n$, such that
for every five strings v, w, x, y and z such that $u = vwxyz$

if
1. $|wy| > 0$
2. $|wxy| \leq n$

then

3. there exists $m \geq 0$, such that $vw^m xy^m z$ does not belong to L

Example 6.3. Applying the Pumping Lemma to *AnBnCn*

$\mathit{AnBnCn} = \{a^i b^i c^i \mid i \geq 0\}$

Choose $u = \ldots$

Example 6.3. Applying the Pumping Lemma to *AnBnCn*

$\mathit{AnBnCn} = \{a^i b^i c^i \mid i \geq 0\}$

Choose $u = a^n b^n c^n$

A slide from lecture 3:

2.4 The Pumping Lemma

Theorem 2.29.
The Pumping Lemma for Regular Languages.
Suppose L is a language over the alphabet Σ.
If L is accepted by a finite automaton $M = (Q, \Sigma, q_0, A, \delta)$,
and if n is the number of states of M,

then for every $x \in L$ satisfying $|x| \geq n$, there are three strings u, v, and w such that $x = uvw$ and the following three conditions are true:

1. $|uv| \leq n$.
2. $|v| > 0$ (i.e., $v \neq \Lambda$).
3. For every $i \geq 0$, the string uv^iw also belongs to L.

Theorem 6.1.
The Pumping Lemma for Context-Free Languages.

Suppose L is a context-free language. Then there is an integer n so that for every $u \in L$ with $|u| \geq n$, u can be written as $u = vwxyz$, for some strings v, w, x, y and z satisfying

1. $|wy| > 0$
2. $|wxy| \leq n$
3. for every $m \geq 0$, $vw^m xy^m z \in L$

Proof. . .

Example 6.5. Applying the Pumping Lemma to...

$\{x \in \{a, b, c\}^* \mid \; n_a(x) < n_b(x)$ and $n_a(x) < n_c(x)\}$

Choose $u = \ldots$

Example 6.5. Applying the Pumping Lemma to...

$\{x \in \{a, b, c\}^* \mid \; n_a(x) < n_b(x)$ and $n_a(x) < n_c(x)\}$

Choose $u = a^n b^{n+1} c^{n+1}$

Example 6.6. The Set of Legal C Programs is Not a CFL

Choose $u = \ldots$

Example 6.6. The Set of Legal C Programs is Not a CFL

Choose $u =$

```
main(){int aaa...a;aaa...a;aaa...a;}
```

where `aaa...a` contains $n + 1$ `a`'s

Example 6.4. Appplying the Pumping Lemma to XX

$XX = \{xx \mid x \in \{a, b\}^*\}$

Choose $u = \ldots$

Example 6.4. Appplying the Pumping Lemma to XX

$XX = \{xx \mid x \in \{a, b\}^*\}$

Choose $u = a^n b^n a^n b^n$