Automata Theory :: Pumping Lemma for Context-Free Languages

Jörg Endrullis

Vrije Universiteit Amsterdam

Pumping Lemma for Context-Free Languages (1961)

Theorem

Let L be a context-free language.

There exists $m > 0$ such that for every word $w \in L$ with $|w| \geq m$:

$$w = uvxyz$$

with $|vxy| \leq m$ and $|vy| \geq 1$, and $uv^ixy^iz \in L$ for every $i \geq 0$.

Proof

Let G be a context-free grammar with $L(G) = L \setminus \{\lambda\}$

- with k variables, and
- without λ and unit productions.

Let m = 1 + maximum number of leaves of derivation trees of depth $\leq (k+2)$.

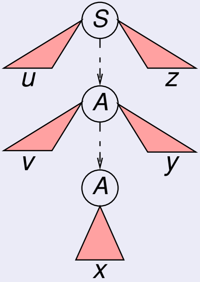
continued on next slide...

Pumping Lemma for Context-Free Languages (1961)

Let $w \in L$ with $|w| \geq m$. Consider a derivation tree for w.

There is a path of length $\geq k + 2$. Consider the longest path.

As there are only k variables, there must be a variable A that occurs twice among the last $k + 1$ variable nodes of the path.



We have $w = uvxyz$ with

- $S \Rightarrow^* uAz$
- $A \Rightarrow^+ vAy$
- $A \Rightarrow^+ x$

Hence

- $S \Rightarrow^+ uv^ixy^iz$ for every $i \geq 0$.

Then

- $|vxy| \leq m$ as the subtree generating vxy has depth $\leq k + 2$,
- $|vy| \geq 1$, since there are no λ and unit productions. □

Using the Pumping Lemma

Attention

A contradiction of the pumping property for specific values of m, or of u, v, x, y, z, is not sufficient!

Pumping lemma as formula (note the quantifiers):

$$\begin{array}{l} \exists m > 0. \\ \quad \forall w \in L \text{ with } |w| \geq m. \\ \quad\quad \exists u, v, x, y, z \text{ with } w = uvxyz,\ |vxy| \leq m,\ |vy| \geq 1. \\ \quad\quad\quad \forall i \geq 0.\, uv^i xy^i z \in L \end{array}$$

To **contradict the pumping property**, we prove the negation:

$$\begin{array}{l} \forall m > 0. \\ \quad \exists w \in L \text{ with } |w| \geq m. \\ \quad\quad \forall u, v, x, y, z \text{ with } w = uvxyz,\ |vxy| \leq m,\ |vy| \geq 1. \\ \quad\quad\quad \exists i \geq 0.\, uv^i xy^i z \notin L \end{array}$$

Pumping Lemma as a Game

To **contradict the pumping property**, we prove the negation:

$\forall m > 0.$
$\exists w \in L$ with $|w| \geq m.$
$\forall u, v, x, y, z$ with $w = uvxyz,\ |vxy| \leq m,\ |vy| \geq 1.$
$\exists i \geq 0.\, uv^ixy^iz \notin L$

Pumping Lemma as a Game

Given is L. We want to prove that L is not context-free.

1. Opponent picks m.
2. We choose a word $w \in L$ with $|w| \geq m$.
3. Opponent picks u, v, x, y, z
   with $w = uvxyz$, $|vxy| \leq m$ and $|vy| \geq 1$.
4. If we can find $i \geq 0$ such that $uv^ixy^iz \notin L$, then **we win.**

If we can always win, then L has no pumping property!

Example

Assume that $L = \{ a^n b^n c^n \mid n \geq 0 \}$ was context-free.

According to the pumping lemma there is $m > 0$ such that

$$a^m b^m c^m = uvxyz$$

with $|vxy| \leq m$, $|vy| \geq 1$, and $uv^i xy^i z \in L$ for every $i \geq 0$.

Since $|vxy| \leq m$, $vy = a^j b^k$ or $vy = b^j c^k$ for some $j, k \geq 0$.

Since $|vy| \geq 1$ we have $j + k \geq 1$.

Then $uv^2 xy^2 z$ does not contain equally many a's, b's and c's.

Contradiction, thus L is not context-free.

Intuitively:

- opponent picks m,
- we pick $w = a^m b^m c^m$,
- opponent u, v, x, y, z

Exercises (1)

Show that $L = \{ a^n b^n \mid n \geq 0 \}$ has the pumping property.

Let $m = 2$. Every word $w = a^n b^n$ with $|w| \geq m$ can be split

$$a^n b^n = uvxyz \quad u = a^{n-1} \quad v = a \quad x = \lambda \quad y = b \quad z = b^{n-1}$$

We have $|vxy| \leq m$, $|vy| \geq 1$and

$$uv^i xy^i z = a^{n-1+i} b^{n-1+i} \in L$$

for every $i \geq 0$. Thus the language has the pumping property.

Exercises (2)

Show $L = \{ w \in \{a, b\}^* \mid w = w^R \}$ for has the pumping property.

Let $m = 3$. Every word $w \in L$ with $|w| \geq m$ has the form

$$w = sctcs^R$$

where $s \in \{a, b\}^*$, $c \in \{a, b\}$ and $t \in \{ a, b, \lambda \}$. Thus

$$w = uvxyz \qquad u = s \qquad v = c \qquad x = t \qquad y = c \qquad z = s^R$$

We have $|vxy| \leq m$, $|vy| \geq 1$and

$$uv^i xy^i z = sc^i tc^i s^R \in L$$

for every $i \geq 0$. Thus the language has the pumping property.

Show that L also has the pumping property for $m = 2$.

Hint: distinguish w of even and odd length when splitting.

Exercises (3)

Show that $L = \{\ ww \mid w \in \{a,b\}^*\ \}$ is not context-free.

Assume that L was context-free.

According to the pumping lemma there is $m > 0$ such that

$$a^m b^m a^m b^m = uvxyz$$

with $|vxy| \leq m$, $|vy| \geq 1$, and $uv^i xy^i z \in L$ for every $i \geq 0$.

Since $|vxy| \leq m$, $vy = a^j b^k$ or $vy = b^k a^j$ for some $j, k \geq 0$.

Since $|vy| \geq 1$ we have $j + k \geq 1$.

Since $|vxy| \leq m$, we have:

- If $|u| < m$, then $uv^0 xy^0 z = a^{m-j} b^{m-k} a^m b^m \notin L$.
- If $m \leq |u| < 2m$, then $uv^0 xy^0 z = a^m b^{m-k} a^{m-j} b^m \notin L$.
- If $2m \leq |u|$, then $uv^0 xy^0 z = a^m b^m a^{m-j} b^{m-k} \notin L$.

Contradiction in each case! Thus L is not context-free.