Automata Theory :: Pumping Lemma for Context-Free Languages

Jörg Endrullis

Vrije Universiteit Amsterdam

Theorem

Let *L* be a context-free language.

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

$$w = uvxyz$$

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

Theorem

Let *L* be a context-free language.

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

$$w = uvxyz$$

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

Proof

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

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

Theorem

Let *L* be a context-free language.

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

$$w = uvxyz$$

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

Proof

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

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

Let $m = 1 + \text{maximum number of leaves of derivation trees of depth} \le (k + 2)$.

Theorem

Let *L* be a context-free language.

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

$$w = uvxyz$$

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

Proof

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

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

Let $m = 1 + \text{maximum number of leaves of derivation trees of depth} \le (k + 2)$.

continued on next slide...

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

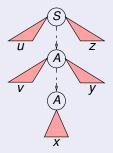
Let $w \in L$ with $|w| \ge m$. Consider a derivation tree for w.

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

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

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

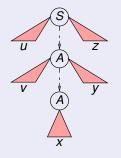
As there are only k variables, there must be a variable A that occurs twice among the last k+1 variable nodes of the path.



Let $w \in L$ with $|w| \ge m$. 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.



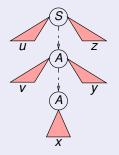
We have w = uvxyz with

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

Let $w \in L$ with $|w| \ge m$. 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

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

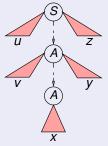
Hence

■ $S \Rightarrow^+ uv^i xy^i z$ for every $i \ge 0$.

Let $w \in L$ with $|w| \ge m$. 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

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

Hence

• $S \Rightarrow^+ uv^i xy^i z$ for every $i \ge 0$.

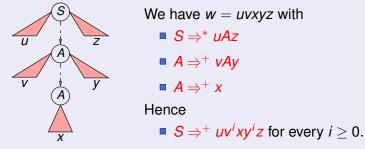
Then

• $|vxy| \le m$ as the subtree generating vxy has depth $\le k+2$,

Let $w \in L$ with $|w| \ge m$. 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.



Then

- $|vxy| \le m$ as the subtree generating vxy has depth $\le k+2$,
- |vy| > 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!

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):

```
\exists m > 0.
\forall w \in L \text{ with } |w| \ge m.
\exists u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\forall i \ge 0. uv^i xy^i z \in L
```

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):

```
\exists m > 0.
\forall w \in L \text{ with } |w| \ge m.
\exists u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\forall i \ge 0. uv^i xy^i z \in L
```

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

```
\forall m > 0.
\exists w \in L \text{ with } |w| \ge m.
\forall u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\exists i \ge 0. uv^i xy^i z \notin L
```

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

```
\forall m > 0.
\exists w \in L \text{ with } |w| \ge m.
\forall u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\exists i \ge 0. uv^i xy^i z \notin L
```

Pumping Lemma as a Game

Given is L.

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

```
\forall m > 0.
\exists w \in L \text{ with } |w| \ge m.
\forall u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\exists i \ge 0. uv^i xy^i z \notin L
```

Pumping Lemma as a Game

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

```
\forall m > 0.
\exists w \in L \text{ with } |w| \ge m.
\forall u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\exists i \ge 0. uv^i xy^i z \notin L
```

Pumping Lemma as a Game

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

1. Opponent picks m.

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

```
\forall m > 0.
\exists w \in L \text{ with } |w| \ge m.
\forall u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\exists i \ge 0. uv^i xy^i z \not\in L
```

Pumping Lemma as a Game

- 1. Opponent picks m.
- 2. We choose a word $w \in L$ with $|w| \ge m$.

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

```
\forall m > 0.
\exists w \in L \text{ with } |w| \ge m.
\forall u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\exists i \ge 0. uv^i xy^i z \notin L
```

Pumping Lemma as a Game

- 1. Opponent picks m.
- 2. We choose a word $w \in L$ with $|w| \ge m$.
- 3. Opponent picks u, v, x, y, z with $w = uvxyz, |vxy| \le m$ and $|vy| \ge 1$.

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

```
\forall m > 0.
\exists w \in L \text{ with } |w| \ge m.
\forall u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\exists i \ge 0. uv^i xy^i z \not\in L
```

Pumping Lemma as a Game

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

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

```
\forall m > 0.
\exists w \in L \text{ with } |w| \ge m.
\forall u, v, x, y, z \text{ with } w = uvxyz, |vxy| \le m, |vy| \ge 1.
\exists i \ge 0. uv^i xy^i z \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 $|w| \ge m$.
- 3. Opponent picks u, v, x, y, z with $w = uvxyz, |vxy| \le m$ and $|vy| \ge 1$.
- 4. If we can find $i \ge 0$ such that $uv^i x y^i z \notin L$, then we win.

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

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

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

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

$$= uvxyz$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Contradiction, thus *L* is not context-free.

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

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

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

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

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

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

Then uv^2xy^2z 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 $w = a^m b^m c^m$,
- \blacksquare opponent u, v, x, y, z

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

Show that $L = \{ a^n b^n \mid n \ge 0 \}$ has the pumping property. Let m = 2.

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

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

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

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

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

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

We have $|vxy| \le m$, $|vy| \ge 1$

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

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

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

We have $|vxy| \le m$, $|vy| \ge 1$ and

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

for every i > 0.

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

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

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

We have $|vxy| \le m$, $|vy| \ge 1$ and

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

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

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

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

Let m = 3.

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

Let m = 3. Every word $w \in L$ with $|w| \ge m$ has the form $w = sctcs^R$

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

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

Let
$$m = 3$$
. Every word $w \in L$ with $|w| \ge m$ has the form $w = sctcs^R$

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

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

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

Let
$$m = 3$$
. Every word $w \in L$ with $|w| \ge m$ has the form $w = sctcs^R$

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

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

We have
$$|vxy| \le m$$
, $|vy| \ge 1$

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

Let m = 3. Every word $w \in L$ with $|w| \ge m$ has the form $w = sctcs^R$

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

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

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

 $uv^ixy^iz = sc^itc^is^R \in L$

for every $i \ge 0$.

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

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

$$w = sctcs^R$$

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

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

We have $|vxy| \le m$, $|vy| \ge 1$ and

$$uv^ixy^iz = sc^itc^is^R \in L$$

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

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

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

$$w = sctcs^R$$

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

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

We have $|vxy| \le m$, $|vy| \ge 1$ and

$$uv^ixy^iz = sc^itc^is^R \in L$$

for every $i \ge 0$. 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.

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

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

Assume that *L* was context-free.

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

$$= uvxyz$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 $\it m > 0$ such that

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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