INF2080 Non-Context-Free Langugaes

Daniel Lupp

Universitetet i Oslo

15th February 2018





University of Oslo

Definition (Context-Free Grammar)

A context-free grammar is a 4-tuple (V, Σ, R, S) where

- I V is a finite set of variables
- **2** Σ is a finite set disjoint from V of *terminals*
- R is a finite set of *rules*, each consisting of a variable and of a string of variables and terminals
- and S is the start variable

Definition (Context-Free Grammar)

A context-free grammar is a 4-tuple (V, Σ, R, S) where

- I V is a finite set of variables
- **2** Σ is a finite set disjoint from V of *terminals*
- R is a finite set of *rules*, each consisting of a variable and of a string of variables and terminals
- and S is the start variable

A language generated by a context-free grammar is a *context-free language*

Context-Free Grammar

$$S \rightarrow aSa \mid bSb \mid cSc \mid a \mid b \mid c \mid \varepsilon$$

 $S
ightarrow aSa \mid bSb \mid cSc \mid a \mid b \mid c \mid \varepsilon$

Palindromes over $\{a, b, c\}$

Palindromes over $\{a, b, c\}$

$$S \rightarrow aSa \mid bSb \mid cSc \mid a \mid b \mid c \mid \varepsilon$$
 $S \rightarrow 01S10 \mid \varepsilon$

Palindromes over $\{a, b, c\}$ $\{(01)^n | n \ge 0\}$

·

 $S \rightarrow aSa \mid bSb \mid cSc \mid a \mid b \mid c \mid \varepsilon$

Palindromes over $\{a, b, c\}$ $\{(01)^n | n \ge 0\}$

 $S \rightarrow 01S10 \,|\, \varepsilon$

 $S \rightarrow 1 \mid 1B1 \mid 0 \mid 0B0$

 $B \rightarrow 1B \mid 0B \mid \varepsilon$

$$S \rightarrow aSa \mid bSb \mid cSc \mid a \mid b \mid c \mid \varepsilon \qquad S \rightarrow 01S10 \mid \varepsilon$$

Palindromes over $\{a, b, c\}$ $\{(01)^n (10)^n \mid n \ge 0\}$ Al

$$B \rightarrow 1B \mid 0B \mid \varepsilon$$

All $\{0, 1\}$ words beginning
and ending with the same
digit

 $S \rightarrow 1 | 1B1 | 0 | 0B0$

Every context-free grammar can be rewritten into Chomsky normal form:

Definition (Chomsky Normal Form)

A grammar in Chomsky normal form consists only of rules of the following form:

 $S
ightarrow \varepsilon$ A
ightarrow BCA
ightarrow d,

where S is the start variable, A, B, C are variables with B, C distinct from S, and d is a terminal.

A pushdown automaton (PDA) is, informally, a NFA with an additional stack to store information.

A pushdown automaton (PDA) is, informally, a NFA with an additional stack to store information.

Definition (PDA)

A PDA is a tuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$ where Q, Σ, Γ, F are finite states and

- \bigcirc Σ is the input alphabet,
- Γ is the stack alphabet,
- $\ \, \bullet \ \, \delta: Q \times \Sigma_{\varepsilon} \times \mathsf{\Gamma}_{\varepsilon} \to \mathcal{P}(Q \times \mathsf{\Gamma}_{\varepsilon}) \ \, \text{is the transition function}$
- **(3)** $q_0 \in Q$ is the start state, and
- $F \subseteq Q$ is the set of accepting states.



Examples:



Examples:





PDA

Examples:



Theorem

Theorem

A language is context-free if and only if a PDA recognizes it.

• Proof was quite involved, check last lecture/book for details (a very good exercise!)

Theorem

- Proof was quite involved, check last lecture/book for details (a very good exercise!)
- So PDAs and CFLs have a similar relationship as NFAs and regular languages

Theorem

- Proof was quite involved, check last lecture/book for details (a very good exercise!)
- So PDAs and CFLs have a similar relationship as NFAs and regular languages
- What about DPDAs (deterministic PDAs)? Are they the same as PDAs?

Theorem

- Proof was quite involved, check last lecture/book for details (a very good exercise!)
- So PDAs and CFLs have a similar relationship as NFAs and regular languages
- What about DPDAs (deterministic PDAs)? Are they the same as PDAs? → Nope! Have, in fact, different expressive power. See book (not exam-relevant).

Non-Context-Free Languages

• Today, we deal with non-context-free languages, i.e., we are interested in seeing whether a language is non-context free.

Non-Context-Free Languages

- Today, we deal with non-context-free languages, i.e., we are interested in seeing whether a language is non-context free.
- Recall our discussion on nonregular languages, in particular the pumping lemma:

Non-Context-Free Languages

- Today, we deal with non-context-free languages, i.e., we are interested in seeing whether a language is non-context free.
- Recall our discussion on nonregular languages, in particular the pumping lemma:

Lemma (Pumping Lemma for Regular Languages)

- $xy^i z \in A$ for every $i \ge 0$,
- **2** |y| > 0,
- $|xy| \le p.$

Lemma (Pumping Lemma for Regular Languages)

•
$$xy^i z \in A$$
 for every $i \ge 0$

- **2** |y| > 0,
- $|xy| \leq p.$

Lemma (Pumping Lemma for Regular Languages)

•
$$xy^i z \in A$$
 for every $i \ge 0$

- **2** |y| > 0,
- $|xy| \le p.$
- Recall that each component *x*, *y*, *z* corresponed to one part in an accepting path in the DFA.

Lemma (Pumping Lemma for Regular Languages)

•
$$xy^i z \in A$$
 for every $i \ge 0$

- **2** |y| > 0,
- $|xy| \le p.$
- Recall that each component x, y, z corresponed to one part in an accepting path in the DFA.
- y corresponded to a cycle in the DFA

Lemma (Pumping Lemma for Regular Languages)

•
$$xy^i z \in A$$
 for every $i \ge 0$

- **2** |y| > 0,
- $|xy| \leq p.$
- Recall that each component x, y, z corresponed to one part in an accepting path in the DFA.
- y corresponded to a cycle in the DFA
- x corresponded to a path from the initial node to the start of the cycle.

Lemma (Pumping Lemma for Regular Languages)

•
$$xy^i z \in A$$
 for every $i \ge 0$

- **2** |y| > 0,
- $|xy| \leq p.$
- Recall that each component x, y, z corresponed to one part in an accepting path in the DFA.
- y corresponded to a cycle in the DFA
- x corresponded to a path from the initial node to the start of the cycle.
- *z* corresponded to a path from the end of the cycle to an accepting state.

Lemma (Pumping Lemma for Regular Languages)

- $xy^i z \in A$ for every $i \ge 0$,
- **2** |y| > 0,
- $|xy| \le p.$
- Recall that each component x, y, z corresponed to one part in an accepting path in the DFA.
- y corresponded to a cycle in the DFA
- x corresponded to a path from the initial node to the start of the cycle.
- z corresponded to a path from the end of the cycle to an accepting state.
- Condition (3) in the lemma was a useful tool when proving the nonregularity of a language.

Lemma (Pumping Lemma for CFLs)

For every context-free language A there exists a number p (called the pumping length) where, if s is a word in A of length $\geq p$, then s can be divided into five parts, s = uvxyz, satisfying the following conditions:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Lemma (Pumping Lemma for CFLs)

For every context-free language A there exists a number p (called the pumping length) where, if s is a word in A of length $\geq p$, then s can be divided into five parts, s = uvxyz, satisfying the following conditions:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

 \rightarrow Things are a bit more involved this time!

Lemma (Pumping Lemma for CFLs)

For CFL A there exists a pumping length p where, if $s \in A$ of length $\geq p$, then s can be divided into five parts, s = uvxyz, such that:

•
$$uv^i xy^i z \in A$$
 for all $i \ge 0$,

- **2** |vy| > 0,
- $|vxy| \leq p.$
- Proof idea: Let G be the CFG generating A and τ a parse tree of s. If s is sufficiently large, we can argue that, along a path, a variable must occur twice.

Parse tree for s:



Parse tree for s:












Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- 2 |vy| > 0,
- $|vxy| \leq p.$

Proof:

• Let G be the CFG generating A and b be the maximal number of symbols on the right side of any rule in G.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Let G be the CFG generating A and b be the maximal number of symbols on the right side of any rule in G.
- We can assume $b \ge 2$, since any CFG is equivalent to a CFG in Chomsky normal form.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- 2 |vy| > 0,
- $|vxy| \leq p.$

- Let G be the CFG generating A and b be the maximal number of symbols on the right side of any rule in G.
- We can assume $b \ge 2$, since any CFG is equivalent to a CFG in Chomsky normal form.
- Thus, in a parse tree, any node can have maximally b children.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- 2 |vy| > 0,
- $|vxy| \leq p.$

- Let G be the CFG generating A and b be the maximal number of symbols on the right side of any rule in G.
- We can assume $b \ge 2$, since any CFG is equivalent to a CFG in Chomsky normal form.
- Thus, in a parse tree, any node can have maximally b children.
- \Rightarrow If a parse tree has height *h*, then it has at most b^h leaves.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof:

• If a parse tree has height h, then it has at most b^h leaves.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- If a parse tree has height h, then it has at most b^h leaves.
- Hence if a string has length $b^{h} + 1$, then the parse tree must have height at least h + 1.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- If a parse tree has height h, then it has at most b^h leaves.
- Hence if a string has length $b^{h} + 1$, then the parse tree must have height at least h + 1.
- Let |V| be the number of variables in the grammar G and set the pumping length $p = b^{|V|+1}$.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- If a parse tree has height h, then it has at most b^h leaves.
- Hence if a string has length $b^{h} + 1$, then the parse tree must have height at least h + 1.
- Let |V| be the number of variables in the grammar G and set the pumping length $p = b^{|V|+1}$.
- If $|s| \ge p$, then its parse tree must have height at least |V| + 1.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof: • If $|s| \ge p$, then its parse tree must have height at least |V| + 1.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof: • If $|s| \ge p$, then its parse tree must have height at least |V| + 1.

 $\bullet\,$ Let τ be the parse tree with the smallest number of nodes

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof: • If $|s| \ge p$, then its parse tree must have height at least |V| + 1.

 $\bullet\,$ Let $\tau\,$ be the parse tree with the smallest number of nodes



Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all i > 0.
- **2** |vy| > 0,
- **3** |vxy| < p.

- Proof: If $|s| \ge p$, then its parse tree must have height at least |V| + 1.
 - Let τ be the parse tree with the smallest number of nodes
 - It has height at least |V| + 1, so the longest path from T to a leaf has length > |V| + 1, i.e., it contains > |V| + 2 nodes.



Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all i > 0,
- **2** |vy| > 0,
- **3** |vxy| < p.

- Proof: If $|s| \ge p$, then its parse tree must have height at least |V| + 1.
 - Let τ be the parse tree with the smallest number of nodes
 - It has height at least |V| + 1, so the longest path from T to a leaf has length > |V| + 1, i.e., it contains > |V| + 2 nodes.
 - The path must have at least |V| + 1 variables. But G only
 - has |V| variables, hence one must occur twice!

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all i > 0,
- **2** |vy| > 0,
- **3** |vxy| < p.

- Proof: If $|s| \ge p$, then its parse tree must have height at least |V| + 1.
 - Let τ be the parse tree with the smallest number of nodes
 - It has height at least |V| + 1, so the longest path from T to a leaf has length > |V| + 1, i.e., it contains > |V| + 2 nodes.
 - The path must have at least |V| + 1 variables. But G only
 - has |V| variables, hence one must occur twice!

Lemma (Pumping Lemma for CFLs)

```
CFL A, s \in A with |s| \ge p, then s = uvxyz, such that:
```

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof:

• But the path could be much longer than |V| + 1, so we choose R to be a repeating variable from the lowest |V| + 1 variables.



Lemma (Pumping Lemma for CFLs)

```
CFL A, s \in A with |s| \ge p, then s = uvxyz, such that:
```

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof:

INF2080

- But the path could be much longer than |V| + 1, so we choose R to be a repeating variable from the lowest |V| + 1 variables.
- We get the situation we looked at before and subdivide

s = uvxyz.

Lemma (Pumping Lemma for CFLs)

```
CFL A, s \in A with |s| \ge p, then s = uvxyz, such that:
```

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof:

- But the path could be much longer than |V| + 1, so we choose R to be a repeating variable from the lowest |V| + 1 variables.
- We get the situation we looked at before and subdivide

s = uvxyz.



Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof:

• The first occurence of *R* generates *vxy*, while the second generates *x* (remember, only using rules from G),



Lemma (Pumping Lemma for CFLs)

```
CFL A, s \in A with |s| \ge p, then s = uvxyz, such that:
```

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \le p.$

- The first occurrence of R generates vxy, while the second generates x (remember, only using rules from G),
- Thus we can replace one subtree with the other and get valid parse trees for uxz and uv^ixy^iz , i > 1.



Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \le p.$

- The first occurence of R generates vxy, while the second generates x (remember, only using rules from G),
- Thus we can replace one subtree with the other and get valid parse trees for *uxz* and *uvⁱxyⁱz*, *i* > 1.



Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- The first occurence of *R* generates *vxy*, while the second generates *x* (remember, only using rules from G),
- Thus we can replace one subtree with the other and get valid parse trees for uxz and uv^ixy^iz , i > 1.



Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- The first occurence of *R* generates *vxy*, while the second generates *x* (remember, only using rules from G),
- Thus we can replace one subtree with the other and get valid parse trees for uxz and uv^ixy^iz , i > 1.
- This proves condition 1.



Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \le p.$

Proof:

• For condition 2, assume, to the contrary, that $v = y = \varepsilon$.



Lemma (Pumping Lemma for CFLs)

```
CFL A, s \in A with |s| \ge p, then s = uvxyz, such that:
```

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \le p.$

- For condition 2, assume, to the contrary, that $v = y = \varepsilon$.
- Then replacing the first *R* (generating *vxy*) with the second *R* (generating *x*) gives us a *smaller* parse tree for s.



Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- For condition 2, assume, to the contrary, that $v = y = \varepsilon$.
- Then replacing the first *R* (generating *vxy*) with the second *R* (generating *x*) gives us a *smaller* parse tree for s.



Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof:

- For condition 2, assume, to the contrary, that $v = y = \varepsilon$.
- Then replacing the first *R* (generating *vxy*) with the second *R* (generating *x*) gives us a *smaller* parse tree for s. Contradiction! We assumed the parse tree was the smallest



ш

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Proof:

• Condition 3: We chose the Rs such that they lie in the lowest |V| + 1 variables on the path.



Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Condition 3: We chose the Rs such that they lie in the lowest |V| + 1 variables on the path.
- Hence the subtree generated by the first R (which generates vxy) has height at most |V| + 1.



Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Condition 3: We chose the Rs such that they lie in the lowest |V| + 1 variables on the path.
- Hence the subtree generated by the first R (which generates vxy) has height at most |V| + 1.
- Thus that subtree has at most $b^{|V|+1}$ leaves.

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Condition 3: We chose the Rs such that they lie in the lowest |V| + 1 variables on the path.
- Hence the subtree generated by the first R (which generates vxy) has height at most |V| + 1.
- Thus that subtree has at most $b^{|V|+1}$ leaves. $\Rightarrow |vxy| \le b^{|V|+1} = p$

Lemma (Pumping Lemma for CFLs)

CFL A, $s \in A$ with $|s| \ge p$, then s = uvxyz, such that:

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Condition 3: We chose the Rs such that they lie in the lowest |V| + 1 variables on the path.
- Hence the subtree generated by the first R (which generates vxy) has height at most |V| + 1.
- Thus that subtree has at most $b^{|V|+1}$ leaves. $\Rightarrow |vxy| \le b^{|V|+1} = p \quad \Box$

Pumping Lemma - Examples

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_1 = \{a^n b^n c^n \mid n \ge 0\}.$

• Assume A_1 is context-free.
Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Assume A_1 is context-free. Then let p be its pumping length and $s = a^p b^p c^p$.
- $|s| = 3p \ge p$ and $s \in A_1$

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Assume A_1 is context-free. Then let p be its pumping length and $s = a^p b^p c^p$.
- $|s| = 3p \ge p$ and $s \in A_1$ \checkmark

- $uv^i xy^i z \in A \ \text{for all} \ i \geq 0,$
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Assume A_1 is context-free. Then let p be its pumping length and $s = a^p b^p c^p$.
- $|s| = 3p \ge p$ and $s \in A_1$ \checkmark
- Now we consider how we can subdivide s into uvxyz.
- Condition 2: v or y must be nonempty.

- $uv^i xy^i z \in A \ \text{for all} \ i \geq 0,$
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_1 = \{a^n b^n c^n \mid n \ge 0\}.$

• Case 1: Either v or y contains more than one type of symbol.

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 1: Either v or y contains more than one type of symbol.
- Then uv^2xy^2z does not have the symbols in the correct order.

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \le p.$

- Case 1: Either v or y contains more than one type of symbol.
- Then uv^2xy^2z does not have the symbols in the correct order. Contradiction!

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_1 = \{a^n b^n c^n \mid n \ge 0\}.$

• Case 2: Both v and y contains only one type of symbol.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2: Both v and y contains only one type of symbol.
- Then v and y cannot contain both a's and b's or both b's and c's.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2: Both v and y contains only one type of symbol.
- Then v and y cannot contain both a's and b's or both b's and c's.
- But then uv^2xy^2z cannot have an equal number of a's, b's, and c's.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2: Both v and y contains only one type of symbol.
- Then v and y cannot contain both a's and b's or both b's and c's.
- But then uv^2xy^2z cannot have an equal number of a's, b's, and c's. Contradiction!
- s cannot be pumped.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2: Both v and y contains only one type of symbol.
- Then v and y cannot contain both a's and b's or both b's and c's.
- But then uv^2xy^2z cannot have an equal number of a's, b's, and c's. Contradiction!
- s cannot be pumped.
- A_3 is, in fact, *context-sensitive* (more on that in a later lecture)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_2 = \{a^i b^j c^k \mid 0 \le i \le j \le k\}.$

• Same as before, assume A_2 is context-free and let $s = a^p b^p c^p$. Again, s = uvxyz.

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_2 = \{a^i b^j c^k \mid 0 \le i \le j \le k\}.$

- Same as before, assume A_2 is context-free and let $s = a^p b^p c^p$. Again, s = uvxyz.
- Case 1: Same argument as before. If either v or y contains more than one type of symbol, then uv^2xy^2z will not have the right order of symbols.

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_2 = \{a^i b^j c^k \mid 0 \le i \le j \le k\}.$

- Same as before, assume A_2 is context-free and let $s = a^p b^p c^p$. Again, s = uvxyz.
- Case 1: Same argument as before. If either v or y contains more than one type of symbol, then uv^2xy^2z will not have the right order of symbols. Contradiction!

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_2 = \{a^i b^j c^k \mid 0 \le i \le j \le k\}, s = a^p b^p c^p$.

• Case 2: If both v and y contain only one type of symbol, then we cannot use the same reasoning as in the last example (the number of a's, b's, and c's are different now!)

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2: If both v and y contain only one type of symbol, then we cannot use the same reasoning as in the last example (the number of a's, b's, and c's are different now!)
- However, since each contains only one type of symbol, one symbol *a*, *b*, or *c* does not occur in *v* or *y*.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_2 = \{a^i b^j c^k \mid 0 \le i \le j \le k\}, s = a^p b^p c^p$.

• Case 2a: *a* does not appear.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_2 = \{a^i b^j c^k \mid 0 \le i \le j \le k\}, s = a^p b^p c^p$.

• Case 2a: a does not appear. Then consider the string uxz. This has the same amount (p) of a's as s, but fewer b's and c's.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2a: a does not appear. Then consider the string uxz. This has the same amount (p) of a's as s, but fewer b's and c's. Contradiction!
- Case 2b: b does not appear.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2a: a does not appear. Then consider the string uxz. This has the same amount (p) of a's as s, but fewer b's and c's. Contradiction!
- Case 2b: *b* does not appear. Then either *a*'s or *c*'s must occur in *v* or *y* (both cannot be empty).

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2a: a does not appear. Then consider the string uxz. This has the same amount (p) of a's as s, but fewer b's and c's. Contradiction!
- Case 2b: b does not appear. Then either a's or c's must occur in v or y (both cannot be empty). If a's appear, then uv²xy²z has more a's than b's.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2a: a does not appear. Then consider the string uxz. This has the same amount (p) of a's as s, but fewer b's and c's. Contradiction!
- Case 2b: *b* does not appear. Then either *a*'s or *c*'s must occur in *v* or *y* (both cannot be empty). If *a*'s appear, then uv^2xy^2z has more *a*'s than *b*'s. If *c*'s appear, then uxz has more *b*'s than *c*'s, contradiction!
- Case 2c: *c* does not appear.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Case 2a: a does not appear. Then consider the string uxz. This has the same amount (p) of a's as s, but fewer b's and c's. Contradiction!
- Case 2b: *b* does not appear. Then either *a*'s or *c*'s must occur in *v* or *y* (both cannot be empty). If *a*'s appear, then uv^2xy^2z has more *a*'s than *b*'s. If *c*'s appear, then uxz has more *b*'s than *c*'s, contradiction!
- Case 2c: c does not appear. Then the string uv^2xy^2z contains more a's or b's than c's, contradiction!

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_3 = \{ \omega \omega \, | \, \omega \in \{0, 1\}^* \}.$

• Assume A_3 is CF, and choose $s = 0^p 10^p 1$.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Assume A_3 is CF, and choose $s = 0^p 10^p 1$.
- set $u = 0^{p-1}$, v = 0, x = 1, y = 0, and $z = 0^{p-1}1$. Then s = uvxyz.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Assume A_3 is CF, and choose $s = 0^p 10^p 1$.
- set $u = 0^{p-1}$, v = 0, x = 1, y = 0, and $z = 0^{p-1}1$. Then s = uvxyz.
- But this can be pumped! $uv^i xy^i z$ is still a word in A_3 .

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Assume A_3 is CF, and choose $s = 0^p 10^p 1$.
- set $u = 0^{p-1}$, v = 0, x = 1, y = 0, and $z = 0^{p-1}1$. Then s = uvxyz.
- But this can be pumped! $uv^i xy^i z$ is still a word in A_3 .
- We need to be more careful with our choice of s.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_3 = \{ \omega \omega \, | \, \omega \in \{0, 1\}^* \}.$

• Let $s = 0^{p}1^{p}0^{p}1^{p}$. We will make use of condition 3, so $|vxy| \le p$.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Let $s = 0^{p}1^{p}0^{p}1^{p}$. We will make use of condition 3, so $|vxy| \le p$.
- Two options: vxy is contained in the first or second half of s, or covers the middle.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Let $s = 0^{p}1^{p}0^{p}1^{p}$. We will make use of condition 3, so $|vxy| \le p$.
- Two options: vxy is contained in the first or second half of s, or covers the middle.
- If vxy is contained in the first half, then pumping up to uv^2xy^2z "pushes" a 1 to the first position of the second half.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Let $s = 0^{p}1^{p}0^{p}1^{p}$. We will make use of condition 3, so $|vxy| \le p$.
- Two options: vxy is contained in the first or second half of s, or covers the middle.
- If vxy is contained in the first half, then pumping up to uv^2xy^2z "pushes" a 1 to the first position of the second half. But this is not a word in A_3 .

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Let $s = 0^{p}1^{p}0^{p}1^{p}$. We will make use of condition 3, so $|vxy| \le p$.
- Two options: vxy is contained in the first or second half of s, or covers the middle.
- If vxy is contained in the first half, then pumping up to uv²xy²z "pushes" a 1 to the first position of the second half. But this is not a word in A₃.
 Example: 000111000111, where v and y are green. pumping up to uv²xy²z yields 00001111000111

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Let $s = 0^{p}1^{p}0^{p}1^{p}$. We will make use of condition 3, so $|vxy| \le p$.
- Two options: vxy is contained in the first or second half of s, or covers the middle.
- If vxy is contained in the second half, then pumping up to uv^2xy^2z "pushes" a 0 to the last position of the first half. Contradiction.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

- Let $s = 0^{p}1^{p}0^{p}1^{p}$. We will make use of condition 3, so $|vxy| \le p$.
- Two options: vxy is contained in the first or second half of s, or covers the middle.
- If vxy is contained in the second half, then pumping up to uv²xy²z "pushes" a 0 to the last position of the first half. Contradiction.
 Example: 000111 000111, where v and y are green. pumping up to uv²xy²z yields 0001110 0000111

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_3 = \{\omega \omega \, | \, \omega \in \{0, 1\}^*\}.$

• Now assume *vxy* covers the middle.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_3 = \{\omega \omega \mid \omega \in \{0,1\}^*\}.$

Now assume vxy covers the middle. Then pumping s down to uxz yields something of the form 0^p1ⁱ0^j1^p.
Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_3 = \{\omega \omega \mid \omega \in \{0,1\}^*\}.$

• Now assume vxy covers the middle. Then pumping s down to uxz yields something of the form $0^{p}1^{i}0^{j}1^{p}$. Since v and y are not both empty, i and j cannot both equal p.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_3 = \{ \omega \omega \, | \, \omega \in \{0, 1\}^* \}.$

Now assume vxy covers the middle. Then pumping s down to uxz yields something of the form 0^p1ⁱ0^j1^p. Since v and y are not both empty, i and j cannot both equal p. ⇒ not a word in A₃

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_3 = \{\omega \omega \, | \, \omega \in \{0, 1\}^*\}.$

- Now assume vxy covers the middle. Then pumping s down to uxz yields something of the form 0^p1ⁱ0^j1^p. Since v and y are not both empty, i and j cannot both equal p. ⇒ not a word in A₃
- Hence *s* cannot be pumped.

Lemma (Pumping Lemma for CFLs)

- $uv^i xy^i z \in A$ for all $i \ge 0$,
- **2** |vy| > 0,
- $|vxy| \leq p.$

Language: $A_3 = \{\omega \omega \, | \, \omega \in \{0, 1\}^*\}.$

- Now assume vxy covers the middle. Then pumping s down to uxz yields something of the form 0^p1ⁱ0^j1^p. Since v and y are not both empty, i and j cannot both equal p. ⇒ not a word in A₃
- Hence s cannot be pumped. $\Rightarrow A_3$ is not context-free