# Recitation 7 - Pumping Lemma for Contex-Free Languages

John Chilton

June 8, 2007

- Homework 3 Questions
- The Pumping Lemma
- Pumping Lemma Examples

Problem 1. Provide CFGs for:

$$L_1 = \{a^m b^n c^p d^q | m + n = p + q\}$$

$$L_2 = \{x_1 \# x_2 \# \ldots \# x_k \mid x_i \in \{a, b\}^* \text{ and for some } i \text{ and } j,\ x_i = x_j^R\}$$

Problem 2. Prove

$$S \rightarrow aSb \mid bY \;\; Ya$$
$$Y \rightarrow bY \mid aY \mid \epsilon$$

generates the language. $\{w \mid w$ is not of the form $a^n b^n$ for some $n\}$

Problem 3. Convert CFGs from problem 1 into PDAs. Just follow the algorithm laid out on pages 115-118.

Problem 4. Given a CFG convert it into Chompsky Normal form.

Problem 5. Show that the class of context-free languages is closed under union, concatenation, and start.

Problem 6. Convert the given CFG to a PDA.

Problem 7. Context-free languages are not closed under intersection.

Problem 8. Let $L$ be a given context-free language and $R$ be a given regular language.

- Part 1. Show $L - R$ must be context-free.
- Part 2. Show $R - L$ isn't necessarily context-free.

Problem 10. Let $G$ be a CFG in Chomsky normal form that contains $b$ variables. Show that, if $G$ generates some string with a derivation having at least $2^b$ steps, $L(G)$ contains an infinite number of strings.
*Hint*: Look at the proof of the pumping lemma.

Some printings of the book containing errors while proving pumping lemma.

Problem 9. Use the pumping lemma for context-free languages to show three languages are not context-free.

$$A = \{a^q \mid q \text{ is prime}\}$$

- $aaa \in A$

$$A = \{a^q \mid q \text{ is prime}\}$$

- $aaa \in A$
- $aaaa \notin A$

$$A = \{a^q \mid q \text{ is prime}\}$$

- $aaa \in A$
- $aaaa \notin A$
- Pick a string of at least length $p$ where $p$ is the pumping length, and pump to a string that is not of prime length.

$$A = \{a^q \mid q \text{ is prime}\}$$

- $aaa \in A$
- $aaaa \notin A$
- Pick a string of at least length $p$ where $p$ is the pumping length, and pump to a string that is not of prime length.
- How do you show the pumped string is not of prime length?

$$A = \{a^q \mid q \text{ is prime}\}$$

- $aaa \in A$
- $aaaa \notin A$
- Pick a string of at least length $p$ where $p$ is the pumping length, and pump to a string that is not of prime length.
- How do you show the pumped string is not of prime length?
- Somewhat analogous to proof that it is not regular found at the end of these slides.

$$B = \{0^n 1^n 0^n 1^n \mid n \geq 0\}$$

- This is similar to examples from the book. Start with this one.

$$C = \{t_1\#t_2\#\ldots\#t_k \mid t_i \in \{a,b\}^* \text{ and for some } i \neq j,\ x_i = x_j\}$$

- Somewhat similar to the examples from the book, and second example in these slides

For a given context-free language, say $A$, there exists a number $p$ called the pumping length of the language. Any string $s \in A$, of length at least $p$ can be divided into $s = uvxyz$ such that:

- $uv^ixy^iz \in A$ for any $i \geq 0$

For a given context-free language, say $A$, there exists a number $p$ called the pumping length of the language. Any string $s \in A$, of length at least $p$ can be divided into $s = uvxyz$ such that:

- $uv^ixy^iz \in A$ for any $i \geq 0$
  - $u$ and $z$ prefix and suffix, not intresting.

For a given context-free language, say $A$, there exists a number $p$ called the pumping length of the language. Any string $s \in A$, of length at least $p$ can be divided into $s = uvxyz$ such that:

- $uv^ixy^iz \in A$ for any $i \geq 0$
  - $u$ and $z$ prefix and suffix, not intresting.
  - Intresting part is $vxy$, $v$ and $y$ can be pumped around $x$
- $|vxy| \leq p$

For a given context-free language, say $A$, there exists a number $p$ called the pumping length of the language. Any string $s \in A$, of length at least $p$ can be divided into $s = uvxyz$ such that:

- $uv^ixy^iz \in A$ for any $i \geq 0$
  - $u$ and $z$ prefix and suffix, not intresting.
  - Intresting part is $vxy$, $v$ and $y$ can be pumped around $x$
- $|vxy| \leq p$
  - Intresting part is a window of length at most $p$

For a given context-free language, say $A$, there exists a number $p$ called the pumping length of the language. Any string $s \in A$, of length at least $p$ can be divided into $s = uvxyz$ such that:

- $uv^ixy^iz \in A$ for any $i \geq 0$
  - $u$ and $z$ prefix and suffix, not intresting.
  - Intresting part is $vxy$, $v$ and $y$ can be pumped around $x$
- $|vxy| \leq p$
  - Intresting part is a window of length at most $p$
  - Need to make an argument for every such possible window, not just all windows in the first $p$ symbols.

For a given context-free language, say $A$, there exists a number $p$ called the pumping length of the language. Any string $s \in A$, of length at least $p$ can be divided into $s = uvxyz$ such that:

- $uv^ixy^iz \in A$ for any $i \geq 0$
  - $u$ and $z$ prefix and suffix, not intresting.
  - Intresting part is $vxy$, $v$ and $y$ can be pumped around $x$
- $|vxy| \leq p$
  - Intresting part is a window of length at most $p$
  - Need to make an argument for every such possible window, not just all windows in the first $p$ symbols.
- $|vy| > 0$

For a given context-free language, say $A$, there exists a number $p$ called the pumping length of the language. Any string $s \in A$, of length at least $p$ can be divided into $s = uvxyz$ such that:

- $uv^ixy^iz \in A$ for any $i \geq 0$
  - $u$ and $z$ prefix and suffix, not intresting.
  - Intresting part is $vxy$, $v$ and $y$ can be pumped around $x$
- $|vxy| \leq p$
  - Intresting part is a window of length at most $p$
  - Need to make an argument for every such possible window, not just all windows in the first $p$ symbols.
- $|vy| > 0$
  - Either $v$ or $y$ needs to be non-empty, but either one of them could be.

- Cannot choose decomposition
  - NEED to argue about every possible decomposition

- Cannot choose decomposition
  - NEED to argue about every possible decomposition
- Cannot use the pumping lemma to prove a language is context-free

- Cannot choose decomposition
  - NEED to argue about every possible decomposition
- Cannot use the pumping lemma to prove a language is context-free
- When using PL, pick a specific, explicitly stated string.

- Cannot choose decomposition
  - NEED to argue about every possible decomposition
- Cannot use the pumping lemma to prove a language is context-free
- When using PL, pick a specific, explicitly stated string.
- Make sure your string is at least of length $p$.

- Cannot choose decomposition
  - NEED to argue about every possible decomposition
- Cannot use the pumping lemma to prove a language is context-free
- When using PL, pick a specific, explicitly stated string.
- Make sure your string is at least of length $p$.
- $w \in A$ cannot be pumped and $w \in B$, does not imply $B$ is not context-free.

$$D = 0^n \# 0^{2n} \# 0^{3n}$$

- $0\#00\#000 \in D$
- $0\#00\#00 \notin D$

Assume $D$ is context-free with pumping length $p$

$$0^p\#0^{2p}\#0^{3p}$$

- $vxy$ cannot lie completely in first group 0s.
  $uv^2xy^2z = 0^{p+i}\#0^{2p}\#0^{3p}$ for $i > 0$. NOT IN D.

Assume $D$ is context-free with pumping length $p$

$$0^p \# 0^{2p} \# 0^{3p}$$

- $vxy$ cannot lie completely in first group 0s.
  $uv^2xy^2z = 0^{p+i} \# 0^{2p} \# 0^{3p}$ for $i > 0$. NOT IN D.
- $vxy$ cannot lie completely in second group of 0s.
  $uv^2xy^2z = 0^{p} \# 0^{2p+i} \# 0^{3p}$ for $i > 0$. NOT IN D.

Assume $D$ is context-free with pumping length $p$

$$0^p \# 0^{2p} \# 0^{3p}$$

- $vxy$ cannot lie completely in first group 0s. $uv^2xy^2z = 0^{p+i}\#0^{2p}\#0^{3p}$ for $i > 0$. NOT IN D.
- $vxy$ cannot lie completely in second group of 0s. $uv^2xy^2z = 0^{p}\#0^{2p+i}\#0^{3p}$ for $i > 0$. NOT IN D.
- $vxy$ cannot lie completely in third group of 0s. $uv^2xy^2z = 0^{p}\#0^{2p}\#0^{3p+i}$ for $i > 0$. NOT IN D.

$$0^p \# 0^{2p} \# 0^{3p}$$

- So we know vxy must contain a $\#$.

$$0^p \# 0^{2p} \# 0^{3p}$$

- So we know vxy must contain a $\#$.
- v and y cannot contain the $\#$, because then $uv^2xy^2z$ would have more than two $\#$s, and thus cannot be contained in $D$.

$$0^p\#0^{2p}\#0^{3p}$$

- So we know vxy must contain a $\#$.
- v and y cannot contain the $\#$, because then $uv^2xy^2z$ would have more than two $\#$s, and thus cannot be contained in $D$.
- So $x$ must contain the $\#$.

$$0^p \# 0^{2p} \# 0^{3p}$$

- So we know vxy must contain a $\#$.
- v and y cannot contain the $\#$, because then $uv^2xy^2z$ would have more than two $\#$s, and thus cannot be contained in $D$.
- So $x$ must contain the $\#$.
- $vxy$ cannot contain 0s from more than two of the groups because $|vxy| \leq p$.

$$0^p \# 0^{2p} \# 0^{3p}$$

- So we know vxy must contain a $\#$.
- v and y cannot contain the $\#$, because then $uv^2xy^2z$ would have more than two $\#$s, and thus cannot be contained in $D$.
- So $x$ must contain the $\#$.
- $vxy$ cannot contain 0s from more than two of the groups because $|vxy| \leq p$.
- So $uv^2xy^2z$ is of the form $0^{p+i}\#0^{2p+j}\#0^{3p}$ or $0^p\#0^{2p+i}\#0^{3p+j}$
  - $i$ and $j$ cannot both be 0, since $v$ and $y$ cannot both be empty.
  - Any such string is not a member of $D$.

$$0^p \# 0^{2p} \# 0^{3p}$$

- So we know vxy must contain a $\#$.
- v and y cannot contain the $\#$, because then $uv^2xy^2z$ would have more than two $\#$s, and thus cannot be contained in $D$.
- So $x$ must contain the $\#$.
- $vxy$ cannot contain 0s from more than two of the groups because $|vxy| \leq p$.
- So $uv^2xy^2z$ is of the form $0^{p+i}\#0^{2p+j}\#0^{3p}$ or $0^p\#0^{2p+i}\#0^{3p+j}$
  - $i$ and $j$ cannot both be 0, since $v$ and $y$ cannot both be empty.
  - Any such string is not a member of $D$.
- Hence $D$ cannot be pumped and is not context-free

2.30 Part C. Show the following string is not context-free:

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

2.30 Part C. Show the following string is not context-free:

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

- $00\#1001 \in E$
- $1001\#00 \notin E$

2.30 Part C. Show the following string is not context-free:

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

- $00\#1001 \in E$
- $1001\#00 \notin E$
- Some candidate strings for pumping:
  - a) $0^p\#1^p0^p$
  - b) $0^p\#0^p$
  - c) $1^p0^p\#1^p0^p$
  - d) $0^p\#0^{p+1}$

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

- Consider $s = 1^p0^p\#1^p0^p$.

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

- Consider $s = 1^p0^p\#1^p0^p$.
- If $vxy$ lies completely on the left side of the $\#$, then $uv^2xy^2z$ will have more symbols on the left side then right, so the left couldn't be a substring of the right. Result not in $E$!

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

- Consider $s = 1^p0^p\#1^p0^p$.
- If $vxy$ lies completely on the left side of the $\#$, then $uv^2xy^2z$ will have more symbols on the left side then right, so the left couldn't be a substring of the right. Result not in $E$!
- If $vxy$ lies completely on the right side, then what can be said about $uv^2xy^2z$?

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

- Consider $s = 1^p0^p\#1^p0^p$.
- If $vxy$ lies completely on the left side of the $\#$, then $uv^2xy^2z$ will have more symbols on the left side then right, so the left couldn't be a substring of the right. Result not in $E$!
- If $vxy$ lies completely on the right side, then what can be said about $uv^2xy^2z$?
- If $vxy$ lies completely on the right side, then $uv^0xy^0z = uxz$ has more symbols to the left of $\#$ than the right. Result not in $E$!

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

$$s = 1^p 0^p \# 1^p 0^p$$

- $vxy$ must contain the $\#$.

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

$$s = 1^p 0^p \# 1^p 0^p$$

- $vxy$ must contain the $\#$.
- Because $|vxy| \leq p$, $|vxy|$ must lie completely between the first 0 and last 1 in $s$.

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

$$s = 1^p 0^p \# 1^p 0^p$$

- $vxy$ must contain the $\#$.
- Because $|vxy| \leq p$, $|vxy|$ must lie completely between the first 0 and last 1 in $s$.
- As before $v$ and $y$ cannot contain the $\#$, $v$ may contain some 0s and $y$ may contain some 1s.

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

$$s = 1^p 0^p \# 1^p 0^p$$

- $vxy$ must contain the $\#$.
- Because $|vxy| \leq p$, $|vxy|$ must lie completely between the first 0 and last 1 in $s$.
- As before $v$ and $y$ cannot contain the $\#$, $v$ may contain some 0s and $y$ may contain some 1s.
- What is the form of $uv^2xy^2z$?

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

$$s = 1^p 0^p \# 1^p 0^p$$

- $vxy$ must contain the $\#$.
- Because $|vxy| \leq p$, $|vxy|$ must lie completely between the first 0 and last 1 in $s$.
- As before $v$ and $y$ cannot contain the $\#$, $v$ may contain some 0s and $y$ may contain some 1s.
- What is the form of $uv^2xy^2z$?
- $1^p 0^{p+i} \# 1^{p+j} 0^p$ where $i$ and $j$ cannot both be 0

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

$$s = 1^p0^p\#1^p0^p$$

- $vxy$ must contain the $\#$.
- Because $|vxy| \leq p$, $|vxy|$ must lie completely between the first 0 and last 1 in $s$.
- As before $v$ and $y$ cannot contain the $\#$, $v$ may contain some 0s and $y$ may contain some 1s.
- What is the form of $uv^2xy^2z$?
- $1^p0^{p+i}\#1^{p+j}0^p$ where $i$ and $j$ cannot both be 0
- Done?

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

$$s = 1^p 0^p \# 1^p 0^p$$

- If $v$ is $\epsilon$, then this has the form $1^p 0^p \# 1^{p+j} 0^p$, which is in $E$

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

$$s = 1^p0^p\#1^p0^p$$

- If $v$ is $\epsilon$, then this has the form $1^p0^p\#1^{p+j}0^p$, which is in $E$
- Well, if $v$ is $\epsilon$, consider $uv^0xy^0z = uxz = 1^p0^p\#1^{p-k}0^p$.

$$E = \{w\#t \mid w \text{ is a substring of t}\}$$

$$s = 1^p0^p\#1^p0^p$$

- If $v$ is $\epsilon$, then this has the form $1^p0^p\#1^{p+j}0^p$, which is in $E$
- Well, if $v$ is $\epsilon$, consider $uv^0xy^0z = uxz = 1^p0^p\#1^{p-k}0^p$.
- This is not in $E$, so this string cannot be pumped

Consider $A = \{a^p \mid \text{p is prime}\}$.

- Homework 3 asks you to show $A$ is context-free.
- I will show how to prove it is not regular with the pumping lemma for regular languages.
- You should repeat this exercise with the context-free pumping lemma, to show it is not context-free.
  - Remember: Not regular does not imply not context-free.

Consider $w = 1^q$ where $q$ is the smallest prime strictly larger the pumping length,$p$, of $A$. Now consider some decomposition $s = xyz$ as promised by the pumping lemma. So $s = 1^a 1^b 1^c$.

Consider $w = 1^q$ where $q$ is the smallest prime strictly larger the pumping length,$p$, of $A$. Now consider some decomposition $s = xyz$ as promised by the pumping lemma. So $s = 1^a 1^b 1^c$.

- $|xy| \leq p \Rightarrow a + b \leq p$

Consider $w = 1^q$ where $q$ is the smallest prime strictly larger the pumping length,$p$, of $A$. Now consider some decomposition $s = xyz$ as promised by the pumping lemma. So $s = 1^a 1^b 1^c$.

- $|xy| \leq p \Rightarrow a + b \leq p$
- $|y| \geq 0 \Rightarrow b \geq 0$

Consider $w = 1^q$ where $q$ is the smallest prime strictly larger the pumping length,$p$, of $A$. Now consider some decomposition $s = xyz$ as promised by the pumping lemma. So $s = 1^a 1^b 1^c$.

- $|xy| \leq p \Rightarrow a + b \leq p$
- $|y| \geq 0 \Rightarrow b \geq 0$
- $|s| > p \Rightarrow c > 0$

Consider $w = 1^q$ where $q$ is the smallest prime strictly larger the pumping length,$p$, of $A$. Now consider some decomposition $s = xyz$ as promised by the pumping lemma. So $s = 1^a 1^b 1^c$.

- $|xy| \leq p \Rightarrow a + b \leq p$
- $|y| \geq 0 \Rightarrow b \geq 0$
- $|s| > p \Rightarrow c > 0$
- $xy^m z \in A \Rightarrow a + c + m * b$ is prime for any $m \geq 0$

Consider $w = 1^q$ where $q$ is the smallest prime strictly larger the pumping length,$p$, of $A$. Now consider some decomposition $s = xyz$ as promised by the pumping lemma. So $s = 1^a 1^b 1^c$.

- $|xy| \leq p \Rightarrow a + b \leq p$
- $|y| \geq 0 \Rightarrow b \geq 0$
- $|s| > p \Rightarrow c > 0$
- $xy^m z \in A \Rightarrow a + c + m * b$ is prime for any $m \geq 0$
- Call $a' = a + c$, and note $a' > 0$.

Consider $w = 1^q$ where $q$ is the smallest prime strictly larger the pumping length,$p$, of $A$. Now consider some decomposition $s = xyz$ as promised by the pumping lemma. So $s = 1^a 1^b 1^c$.

- $|xy| \leq p \Rightarrow a + b \leq p$
- $|y| \geq 0 \Rightarrow b \geq 0$
- $|s| > p \Rightarrow c > 0$
- $xy^m z \in A \Rightarrow a + c + m * b$ is prime for any $m \geq 0$
- Call $a' = a + c$, and note $a' > 0$.
- $xy^{a'} z = 1^{a' * b + a'} = 1^{a'(1+b)}$

Consider $w = 1^q$ where $q$ is the smallest prime strictly larger the pumping length,$p$, of $A$. Now consider some decomposition $s = xyz$ as promised by the pumping lemma. So $s = 1^a1^b1^c$.

- $|xy| \leq p \Rightarrow a + b \leq p$
- $|y| \geq 0 \Rightarrow b \geq 0$
- $|s| > p \Rightarrow c > 0$
- $xy^mz \in A \Rightarrow a + c + m * b$ is prime for any $m \geq 0$
- Call $a' = a + c$, and note $a' > 0$.
- $xy^{a'}z = 1^{a'*b+a'} = 1^{a'(1+b)}$
- $xy^{a'+b+1}z = 1^{a'+b(a'+b+1)} = 1^{a'(1+b)+b(1+b)} = 1^{(a'+1)(b+1)}$