Non-Context-Free Languages
Sections: 2.3
October 24, 2016

Pumping Lemma: Regular Languages
Example

Initial Grammar G:

\[
S \rightarrow ASA \mid aB \\
A \rightarrow B \mid S \\
B \rightarrow b \mid \varepsilon
\]

Chomsky Normal Form (CNF):

\[
S_0 \rightarrow AA_1 \mid UB \mid a \mid SA \mid AS \\
S \rightarrow AA_1 \mid UB \mid a \mid SA \mid AS \\
A \rightarrow b \mid AA_1 \mid UB \mid a \mid SA \mid AS \\
A_1 \rightarrow SA \\
U \rightarrow a, B \rightarrow b
\]

CNF allows for binary parse trees!

Note bab is in G since production is:

\[
S_0 \rightarrow AA_1 \rightarrow ASA \rightarrow bab
\]
Example

Note bab is in G since production is
S0 → AA1 → ASA → bab
Parse tree? What is height (root is at height 0)?

Height of parse tree versus max word length
(root at height 0)

<table>
<thead>
<tr>
<th>Height of parse tree</th>
<th>Max word length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>h</td>
<td>$2^h$</td>
</tr>
<tr>
<td>$</td>
<td>\mathcal{V}</td>
</tr>
</tbody>
</table>
Pumping Lemma: CFG

Theorem: For any CFG there is an equivalent grammar in CNF (thus binary parse tree).

Pumping lemma (CFG): Suppose A is a CFG. There exists a number $p$ such that if $s \in A$ and $|s| \geq p$ then $s = uvxyz$ where
- $uv^i xy^i z \in A$, for all $i \geq 0$
- $|vy| > 0$ (so not both $v$, $y$ are empty)
- $|vxy| \leq p$

Proof

Suppose A is a CFG in CNF and $s \in A$,
- $|s| \geq p = 2|V|+1$
1) Height of parse tree for $s$ is at least $|V|+1$
2) Let $R = \text{variable that is repeated latest}$
3) Let $x = \text{substring of } s \text{ derived from last appearance of } R$, $vxy$ derived from second to last occurrence, $u$, $z$ the rest
4) Then $s = uvxyz$. 
Pumping a Parse Tree

Should be able to pump up or pump down and still be within the language!

Example

$L = \{a^ib^i \mid i \geq 0\}$

A PDA **can** represent this. Why?

Pumping Lemma:

\[ s = \]
\[ u = \]
\[ v = \]
\[ x = \]
\[ y = \]
\[ z = \]
Example

L = \{a^ib^jc^i \mid i \geq 0\}

A PDA **cannot** represent this. Why?

Pumping Lemma:

\[
s = \\
u = \\
v = \\
x = \\
y = \\
z =
\]

Example

L = \{a^ib^jc^k \mid k \geq j \geq i \geq 0\}

A PDA cannot represent this. Why?

Pumping Lemma:

\[
s = \\
u = \\
v = \\
x = \\
y = \\
z =
\]
Example

\[ L = \{ \text{ww} | w \in \{0, 1\}^* \} \]

Pump-able?

\[ s = \]

Example

\[ L = \{ \text{w # x} | \text{w}^R \text{ is substring of } x; \text{w,x} \in \{0, 1\}^* \} \]

Pump-able?

\[ s = \]
Operations

• What operations are closed over context-free languages?
  – Union
  – Intersection
  – Complement
  – Kleene Star
  – Concatenation

Exercise Examples

• p 131
  – 2.30 Show the following are not CFL
    • \( \{0^n\#0^{2n}\#0^{3n} \mid n \geq 0\} \)
    • \( \{w\#t \mid w \text{ is a substring of } t; \ w, t \text{ in } \{a,b\}^*\} \)
  – 2.31
    – Show that \( \{0^n1^{n2^n} \mid n \geq 0\} \) is non context-free
    – Show that the complement of \( \{0^n1^{n2^n} \mid n \geq 0\} \) is context-free.