Section: Properties of Context-free Languages

Which of the following languages are CFL?

- $L = \{a^n b^n c^j | 0 < n \leq j\}$ NOT CFL
- $L = \{a^n b^j a^n b^j | n > 0, j > 0\}$ Not CFL
- $L = \{a^n b^j a^k b^p | n + j \leq k + p, n > 0, j > 0, k > 0, p > 0\}$ CFL
- $L = \{a^n b^j a^j b^n | n > 0, j > 0\}$ CFL
Pumping Lemma for Regular Language’s: Let $L$ be a regular language, Then there is a constant $m$ such that $w \in L$, $|w| \geq m$, $w = xyz$ such that

- $|xy| \leq m$
- $|y| \geq 1$
- for all $i \geq 0$, $xy^i z \in L$
Pumping Lemma for CFL’s Let $L$ be any infinite CFL. Then there is a constant $m$ depending only on $L$, such that for every string $w$ in $L$, with $|w| \geq m$, we may partition $w = uvxyz$ such that:

- $|vxy| \leq m$, (limit on size of substring)
- $|vy| \geq 1$, ($v$ and $y$ not both empty)

For all $i \geq 0$, $uv^i xy^i z \in L$

**Proof:** (sketch) There is a CFG $G$ s.t. $L= L(G)$. 
Consider the parse tree of a long string in $L$. 
For any long string, some nonterminal $N$ must appear twice in the path.
Example: Consider $L = \{a^nb^nc^n : n \geq 1\}$. Show $L$ is not a CFL.

• Proof: (by contradiction)  
  Assume $L$ is a CFL and apply the pumping lemma.  
Let $m$ be the constant in the pumping lemma and consider  
$w = a^mb^mc^m$. Note $|w| \geq m$.  
Show there is no division of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$,  
and $uv^ixy^iz \in L$ for $i = 0, 1, 2, \ldots$. 
\[ w = a^m b^m c^m \]

Case 1: \( V \) nor \( y \) or \( \) distinct symbols

Case 2: \( V = a \) \( t_1 \) \( \rightarrow \) \( y = a^t_2 \)

Case 3: \( V = b \) \( t_1 \) \( \rightarrow \) \( y = b^t_3 \)

Case 4: \( V = C \)
Thus, there is no breakdown of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$ and for all $i \geq 0$, $uv^ixy^iz$ is in $L$. Contradiction, thus, $L$ is not a CFL. Q.E.D.
Example Why would we want to recognize a language of the type \( \{a^n b^n c^n : n \geq 1 \} \)?
Example: Consider
$L = \{a^nb^nc^p : p > n > 0\}$. Show $L$ is not a CFL.

- Proof: Assume $L$ is a CFL and apply the pumping lemma. Let $m$ be the constant in the pumping lemma and consider

$w = \phantom{=} \text{Note } |w| \geq m.$

Show there is no division of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$, and $uv^ixy^iz \in L$ for $i = 0, 1, 2, \ldots$. 

Strategy → Choose a tight string to get contradiction
easily!!

$w = a^mb^m c^{m+1}$

$c^p, p > n$

$|w| \geq m$
Strategy $\rightarrow$ Choose a tight string to get contradiction easily!!

$\omega = a^{m} b^{m} c^{m+1}$

Case 1: $\nu$ nor y distinct symbols $\Rightarrow \nu$ contains $a's + b'$

$\Rightarrow u \nu^2 x y^2 z \notin L$

$|\omega| \geq m$

Case 2: $\nu$ is all $a$'s
Example: Consider $L = \{a^j b^k : k = j^2\}$. Show $L$ is not a CFL.

• Proof: Assume $L$ is a CFL and apply the pumping lemma. Let $m$ be the constant in the pumping lemma and consider

$$w = a^m b^{m^2}$$

Show there is no division of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$, and $uv^i xy^i z \in L$ for $i = 0, 1, 2, \ldots$.

Case 1: Neither $v$ nor $y$ can contain 2 or more distinct symbols. If $v$ contains $a$’s and $b$’s, then $uv^2 xy^2 z \notin L$ since there will be $b$’s before $a$’s.

Thus, $v$ and $y$ can be only $a$’s, and $b$’s (not mixed).
\[ a^m b^n \]

\[ v = a^r, \quad y = b^t \]

\[ i = 2, \quad \ln^2 x y^z = m + 1, \quad m^2 + t = 2 \]

\[ m + 1 > a \]

\[ f - 1 \equiv (m + 1)^2 = m^2 + 2m + 1 \]

There will be too few b's.

You have to prove all the cases.

Give a contradiction.

Another case:

\[ v = b^{t+1}, \quad y = b^t \]

\[ i = 2, \quad \ln^2 x y^z = m + m^2 + t+1, \quad m^2 + t+1 \]

\[ \text{num of b's is not an even squared.} \]
Example: Consider
$L = \{ww\bar{w} : w \in \Sigma^*\}$, $\Sigma = \{a, b\}$, where $\bar{w}$ is the string $w$ with each occurrence of $a$ replaced by $b$ and each occurrence of $b$ replaced by $a$. Show $L$ is not a CFL.

**Proof:** Assume $L$ is a CFL and apply the pumping lemma. Let $m$ be the constant in the pumping lemma and consider

\[ w = \overline{ab}^m \overline{ab}^m \overline{ab}^m \]

Show there is no division of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$, and $uv^ixy^iz \in L$ for $i = 0, 1, 2, \ldots$.
\begin{align*}
\text{Case 1} & \quad v = a^1, \quad y = b^{1+3} \\
& \quad i = 0 \quad m = 1, \quad m = 1, \quad m = 1, \quad m = 1,
\end{align*}

\text{num of a's won't match num of A's}

\begin{align*}
\text{Case 2} & \quad v = a^{1+4}, \quad y = b^{1+3} \\
& \quad i = 0 \quad \text{num of y's} = a \quad b \quad A
\end{align*}

\text{num of a's < num of A's}
Example: Consider $L = \{a^n b^p b^p a^n\}$. $L$ is a CFL. The pumping lemma should apply!

Let $m \geq 4$ be the constant in the pumping lemma. Consider $w = a^m b^m b^m a^m$.

We can break $w$ into $uvxyz$, with:

$u = a^m b^{m^2}$, $v = b$, $x = b b$, $y = b$, $z = b^{m-2} a^m$

$u v^i x y^i z = a^m b^{m+i} b^{m+i} a^m \in L$ for any $i$.
Chap 8.2 Closure Properties of CFL’s

Theorem CFL’s are closed under union, concatenation, and star-closure.

• Proof:

Given 2 CFG $G_1 = (V_1, T_1, S_1, P_1)$ and $G_2 = (V_2, T_2, S_2, P_2)$

– Union:

Construct $G_3$ s.t. $L(G_3) = L(G_1) \cup L(G_2)$.

$G_3 = (V_3, T_3, S_3, P_3)$

$V_3 = V_1 \cup V_2 \cup \{S_3\}$, $T_3 = T_1 \cup T_2$

$P_3 = P_1 \cup P_2 \cup \{S_3 \rightarrow S_1 \mid S_2\}$
- Concatenation:
  Construct $G_3$ s.t. $L(G_3) = L(G_1) \circ L(G_2)$.
  $G_3 = (V_3, T_3, S_3, P_3)$

  $T_3 = T_1 \cup T_2$

  Similar $V_3 = V_1 \cup V_2 \cup \{S_3\}$

  $P_3 = P_1 \cup P_2 \cup \{S_3 \rightarrow S_1 S_2\}$

- Star-Closure
  Construct $G_3$ s.t. $L(G_3) = L(G_1)^*$
  $G_3 = (V_3, T_3, S_3, P_3)$

  $V_3 = V_1 \cup \{S_3\}$

  $T_3 = T_1$

  $P_3 = P_1 \cup \{S_3 \rightarrow S_1 S_2\}$
Theorem CFL’s are NOT closed under intersection and complementation.

- **Proof:**

  - **Intersection:**
    \[ L_1 = \{ a^nb^n c^m \mid n,m \geq 0 \} \]
    \[ L_2 = \{ a^n b^m c^n \mid n,m \geq 0 \} \]
    \[ L_1 \cap L_2 \]
    is not CFL!
– Complementation:
Theorem: CFL’s are closed under regular intersection. If $L_1$ is CFL and $L_2$ is regular, then $L_1 \cap L_2$ is CFL.

• Proof: (sketch) We take a NPDA for $L_1$ and a DFA for $L_2$ and construct a NPDA for $L_1 \cap L_2$.

$M_1 = (Q_1, \Sigma, \Gamma, \delta_1, q_0, z, F_1)$ is an NPDA such that $L(M_1) = L_1$.

$M_2 = (Q_2, \Sigma, \delta_2, q_0, F_2)$ is a DFA such that $L(M_2) = L_2$.

$M_3 = (Q_3, \Sigma, \Gamma, \delta_3, (q_0, p_0), z, F_3)$

$Q_3 = Q_1 \times Q_2, F_3 = \{(q, p) \mid q \in F_1, p \in F_2\}$

Example of replacing arcs (NOT a Proof!):
combine

\[ q_0 \]

\[ q_1, q_k \]

\[ q_i, q_k \]

\[ a, x, y \]

\[ a, y, y \]

\[ a, x, y, y \]

\[ a, y, y \]

\[ q_0, q_0 \]

\[ q_i, q_k \]

\[ \ldots \]

\[ m \]

\[ m^2 \]
We must formally define $\delta_3$. If

$$(q_k, x) \in S((q_i, a, b), \delta_3(q_j, a)) = q_l$$

then

$$(q_k, q_l, x) \in \delta_3((q_i, q_j), a, b)$$

Must show

$$((q_0, q_0), w, z) \neq ((q_i, q_i), z, x)$$

if and only if

$$(q_0, w, z) \neq (q_i, z, z)$$

$$(q_0', w) \neq (q_j, z)$$

$$q \in F_1, q_j \in F_2$$
Questions about CFL:

1. Decide if CFL is empty?

   get rid of useless
   prod.
   If nothing is left
   it is empty

2. Decide if CFL is infinite?

   get rid of useless again
   variable that repeats?
   \[ A \Rightarrow \gamma A \gamma \]

   Look for cycle graph
Example: Consider $L = \{a^{2n}b^{2m}c^n d^m : n, m \geq 0\}$. Show $L$ is not a CFL.

- **Proof:** Assume $L$ is a CFL and apply the pumping lemma. Let $m$ be the constant in the pumping lemma and consider $w = a^{2m}b^{2m}c^m d^m$.

Show there is no division of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$, and $uv^i xy^i z \in L$ for $i = 0, 1, 2, \ldots$.

**Case 1:** Neither $v$ nor $y$ can contain 2 or more distinct symbols. If $v$ contains $a$’s and $b$’s, then $uv^2xy^2z \notin L$ since there will be $b$’s before $a$’s.

Thus, $v$ and $y$ can be only $a$’s, $b$’s, $c$’s, or $d$’s (not mixed).

**Case 2:** $v = a^{t_1}$, then $y = a^{t_2}$ or $b^{t_3}$ ($|vxy| \leq m$)

If $y = a^{t_2}$, then
\[ uv^2xy^2z = a^{2m+t_1+t_2}b^{2m}c^m d^m \notin L \text{ since } t_1 + t_2 > 0, \text{ the number of } a'\text{s is not twice the number of } c'\text{s.} \]

If \( y = b^{t_3} \), then

\[ uv^2xy^2z = a^{2m+t_1}b^{2m+t_3}c^m d^m \notin L \text{ since } t_1 + t_3 > 0, \text{ either the number of } a'\text{s (denoted } n(a)) \text{ is not twice } n(c) \text{ or } n(b) \text{ is not twice } n(d). \]

**Case 3:** \( v = b^{t_1} \), then \( y = b^{t_2} \) or \( c^{t_3} \)

If \( y = b^{t_2} \), then

\[ uv^2xy^2z = a^{2m}b^{2m+t_1+t_2}c^m d^m \notin L \text{ since } t_1 + t_2 > 0, n(b) > 2*n(d). \]

If \( y = c^{t_3} \), then

\[ uv^2xy^2z = a^{2m}b^{2m+t_1}c^m+t_3 d^m \notin L \text{ since } t_1 + t_3 > 0, \text{ either } n(b) > 2*n(d) \text{ or } 2*n(c) > n(a). \]

**Case 4:** \( v = c^{t_1} \), then \( y = c^{t_2} \) or \( d^{t_3} \)

If \( y = c^{t_2} \), then

\[ uv^2xy^2z = a^{2m}b^{2m}c^{m+t_1+t_2}d^m \notin L \text{ since } t_1 + t_2 > 0, 2*n(c) > n(a). \]

If \( y = d^{t_3} \), then
\[ uv^2xy^2z = a^{2m}b^{2m}c^{m+t_1}d^{m+t_3} \notin L \text{ since } t_1 + t_3 > 0, \text{ either } 2*n(c) > n(a) \text{ or } 2*n(d) > n(b). \]

Case 5: \( v = d^{t_1} \), then \( y = d^{t_2} \)
then \( uv^2xy^2z = a^{2m}b^{2m}c^{m}d^{m+t_1+t_2} \notin L \text{ since } t_1 + t_2 > 0, \text{ } 2*n(d) > n(c). \)

Thus, there is no breakdown of \( w \)
into \( uvxyz \) such that \( |vy| \geq 1 \),
\( |vxy| \leq m \) and for all \( i \geq 0 \), \( uv^ixy^iz \) is
in \( L \). Contradiction, thus, \( L \) is not a
CFL. Q.E.D.