Pumping Lemma for CFL

Let $G = (V, \Sigma, S, P)$ be a CFG. Suppose we have a string $u \in L(G)$ with the following derivation:

$$S \Rightarrow^* vAz \Rightarrow^* vwAy \Rightarrow^* vwxyz = u$$

where $v, w, x, y, z \in \Sigma^*$. $A \in V$

Which other strings (besides $u$) can be generated by $G$?

$$S \Rightarrow^* vA \Rightarrow^* vxz$$

$$S \Rightarrow^* vA \Rightarrow^* vwAy \Rightarrow^* vwwAyyz \Rightarrow^* vwwxxyyz$$

$$vw^mx^ny^nz \in L(G) \text{ for } m \geq 0$$
Pumping Lemma for CFL

Let $G = (V, \Sigma, S, P)$ be a CFG. Suppose we have a string $u \in L(G)$ with the following derivation:

$$
S \Rightarrow^* vAz \Rightarrow^* vwAy \Rightarrow^* vwxyz = u
$$

where $v, w, x, y, z \in \Sigma^*$.

Which other strings (besides $u$) can be generated by $G$?

$$
vw^mxy^mz \in L(G) \quad \text{for every } m \geq 0
$$

Is it true that for every sufficiently long $u$ there exist $v, w, x, y, z$ satisfying the above derivation? \textbf{YES.}
Pumping Lemma for CFL

Let $G = (V, \Sigma, S, P)$ be a CFG. Suppose $G$ is in Chomsky normal form, i.e., every rule is of the form $A \rightarrow \sigma$ or $A \rightarrow A_1A_2$ where $A, A_1, A_2 \in V$ and $\sigma \in \Sigma$ (plus we allow the rule $S \rightarrow \Lambda$).

Let's look at a derivation tree for some $u \in L(G)$.

\[
S \Rightarrow^* vA_2 \Rightarrow^* vwAy_z \Rightarrow^* \\\nvwxyz = u
\]

if $\exists$ a path from $S$ w. repeating nonterminal, then have
if $\exists$ path of length $> |\Sigma|$, then
if the depth of deriv. tree $= d$, then $|u| \leq 2^d$
if $|u| > 2^{|\Sigma|}$ then the depth of the deriv. tree $> |\Sigma|$
Pumping Lemma for CFL

Thm (Pumping lemma for CFL):

Let $L$ be a CFL. Then there exists an integer $n > 0$ such that for every $u \in L$ with $|u| \geq n$ there exist strings $v, w, x, y, z$ satisfying:

1. $u = vwxyz$
2. $|wxy| \leq n$
3. $|wy| > 0$
4. $vw^mxy^mz \in L$ for every $m \geq 0$
Pumping Lemma for CFL

Example: \( L = \{ a^k b^k c^k \mid k \geq 0 \} \)

Pf:

suppose \( L \) is a CFL. then the PL for CFL holds. let \( n \) be the PL constant.

take \( u = a^n b^n c^n \in L \vee |u| = 3n \geq n \vee \)

should exist \( v, w, x, y, z \) sat. (1-4).

if they exist by 1)

\[
\begin{align*}
 u &= a^n b^n c^n \\
 v & \quad w & \quad x & \quad y & \quad z
\end{align*}
\]

by 2) \( wxy \) contains at most two types of symbols (and not \( a \)'s with \( c \)'s)

by 4) consider \( m = 0 \) \( vw^0 xy^0 z = vxz \)

CASE 1) \( wxy \) does not contain \( c \)'s then \( vxz \) contains exactly \( n \) \( c \)'s

by 3) \( \text{contains } < 2n \text{ a's \& b's} \)

CASE 2) \( wxy \) does not contain \( a \)'s

analogous

\[ \square \]
Lemma: CFL’s are not closed under intersection.

Pf: have to find two CFL’s $L_1, L_2$ s.t. $L_1 \cap L_2$ is not a CFL.

try for

$\begin{align*}
L_1 \cap L_2 &= \{a^k b^k c^k \mid k \geq 0\} \\
L_1 &= \left\{a^k b^k c^k \mid k \geq 0\right\} \\
L_2 &= \left\{a^k b^k c^l \mid k, l \geq 0\right\}
\end{align*}$

$\begin{align*}
S &\rightarrow AB \\
A &\rightarrow aA \mid \Lambda \\
B &\rightarrow bBc \mid \Lambda
\end{align*}$

for $L_1$
Lemma: CFL's are not closed under complement.

Pf:

**OPTION 1:** Suppose CFL's are closed under complement. We know that CFL's are closed under $\cup, \cdot, \ast$.

Let $L_1, L_2$ be CFL's.

$$L_1 \cap L_2 = \left(L_1' \cup L_2'ight)'$$

Thus this should be CFL.

**OPTION 2:** Give a CFL $L$ s.t. $L'$ is not CFL.

$$L = \{a^k b^k c^k \mid k \geq 0\}' = \{w \in \{a,b,c\}^* \mid w \text{ is not of the form } a^*b^*c^* \} \cup$$

$$\{a^i b^j c^k \mid i \neq j, i, j, k \geq 0\} \cup$$

$$\{a^i b^j c^k \mid j \neq k, i, j, k \geq 0\} \cup$$

$$\{a^i b^j c^k \mid i \neq k, i, j, k \geq 0\}.$$
Lemma: CFL’s are not closed under difference.

**Option 1:** \( L' = \Sigma^* - L \uparrow \text{CFL} \)

Assume CFL’s are closed under subtraction, then if \( L \) is a CFL, \( \Sigma^* - L \) should be CFL.

**Option 2:** Find two CFL \( L_1, L_2 \) s.t. \( L_1 - L_2 \) is not CFL.

\( L_1 = a^*b^*c^* \)

\( L_2 = \{ a^ib^jc^k \mid i \neq j \text{ or } i \neq k \text{ or } j \neq k \} \)

\( L_1 - L_2 = \{ a^kb^kc^k \mid k > 0 \} \)
Lemma: CFL’s are closed under intersection with a regular language.

Pf: we want to prove that if $L_1$ is CFL and $L_2$ is regular, then $L_1 \cap L_2$ is CFL.

let $M_1$ be a PDA accepting $L_1$ and $M_2$ be an FA accepting $L_2$.

then we can use a cross-product construction to give a PDA $M_3$ for $L_1 \cap L_2$.

Note: we cannot do a cross-product construction for two PDA’s because they would fight over the single stack available for the resulting PDA!
Example: \( L = \{ w \in \{a,b,c\}^* \mid n_a(w) = n_b(w) = n_c(w) \} \)

Assume \( L \) is a CFL. Then take \( L \cap a^* b^* c^* \)

\[ \underbrace{\text{regular}}_{\text{CFL}} \]

\[ \leftarrow \text{not CFL} \]

If \( L_1 \) is not CFL and \( L_1 \subseteq L_2 \), then \( L_2 \) is not CFL  CANNOT DO.
Decision Problems for CFL's

A decision problem: answer is YES or NO.

Give algorithms for the following decision problems:

- Given a CFL L (by a PDA or a CFG) and a string $x$, is $x$ in L?

**Algorithm:**

1. Check if $S \Rightarrow^* x$
2. Convert $G$ into Chomsky normal form
   - $A \rightarrow A, A_2$ or $A \rightarrow \epsilon$ - the form for rules
   - $S \rightarrow AB \mid BS$
   - $A \rightarrow AA \mid BB$
   - $S \Rightarrow BS \Rightarrow BBS \Rightarrow BBAB \Rightarrow B \rightarrow AS \mid c \mid a$
   - $\Rightarrow BBBBB \Rightarrow^* aaaca$

   Need exactly $2n-1$ rules to generate a string of length $n$

**Algorithm (Note):** Try all possible derivations of $2n-1$ rules and see if can get $x$. (Note: exists an efficient algo, CTK)
A decision problem: answer is YES or NO.

Give algorithms for the following decision problems:

- Given a CFL \( L \) (by a PDA or a CFG), is \( L = \emptyset \) ?

  a variable \( A \) is \underline{terminable} if \( \exists x \in \Sigma^* \) s.t. \( A \Rightarrow^* x \)

  1) if \( A \Rightarrow x \) where \( x \in \Sigma^* \) is a rule, then \( A \) is \underline{terminable}

  2) if \( A \Rightarrow x \) where \( x \in (\Sigma u \text{terminable})^* \) then \( A \) is terminable

  3) n.e. is term.

ALGO: check if \( S \) is terminable.
A decision problem: answer is YES or NO.

Give algorithms for the following decision problems:

- Given a CFL $L$ (by a PDA or a CFG), is $L$ finite?

Hint: pumping lemma.

If $\exists u \in L$ s.t. $|u| > n$ then I can pump $u$ and thus $L$ is infinite. (where $n = 2^{|v|}$)

Suppose $u \in L$ and $|u| > 2n$ then $\exists v, w, x, y, z$ s.t. 1) $u = vwxyz$

- take $m = 0$, then
  
  $vw^0xy^0z = vxz \in L$

- notice:
  
  $n \leq |u| - n \leq |s| < |u|$ thus, if $L$ is infinite, there

  exists a string $t \in L$ s.t. $n \leq |t| \leq 2n$

ALGO: check if there exists a string in $L$ of length $(n, 2n)$. 

[Section 8.3]