Which of the following languages are CFL?

- \( L = \{ a^n b^n c^j \mid 0 < n \leq j \} \)
- \( L = \{ a^n b^n a^n b^j \mid n > 0, j > 0 \} \)
- \( L = \{ a^n b^i a^k b^p \mid n + j \leq k + p, n > 0, j > 0, k > 0, p > 0 \} \)

**Pumping Lemma for Regular Language’s** Let \( L \) be a regular language. Then there is a constant \( m \) such that \( w \in L \) if \( |w| \geq m \) then \( w \) may be partitioned \( w = xyz \) 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:

- \( |wxyz| \leq m \) (limit on size of substring)
- \( |xy| \geq 1 \) (\( v \) 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 \( w \) some nonterminal \( N \) must appear twice in the path.
Example: Consider $L = \{a^n b^n c^n : n \geq 1\}$. 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 $w = a^m b^m c^m$. Note $|w| \geq m$.

  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 $uv^2 y^2 z \notin L$ since there will be $b$'s before $a$'s.

  Thus $v$ and $y$ can only $a$'s or $b$'s (not mixed).

  Case 2: $v = a^t \Gamma$ then $y = a^t_2$ or $b^t_2$ ($|vxy| \leq m$)

  If $y = b^t_2$ then $uv^2 x y^2 z = a^{m+t_1+t_2} b^m c^m \notin L$ since $t_1 + t_2 > 0 \Gamma (a) > (b)$'s (number of $a$'s is greater than number of $b$'s)

  If $y = b^t_2$ then $uv^2 x y^2 z = a^{m+t_1+t_2} b^m c^m \notin L$ since $t_1 + t_2 > 0 \Gamma (a) > (n) (c)$'s or $b(n) > (c)$'s.

  Case 3: $v = b^t_2 \Gamma$ then $y = b^t_2$ or $c^t_2$

  If $y = b^t_2$ then $uv^2 x y^2 z = a^{m+b^t_1+b^t_2} c^m \notin L$ since $t_1 + t_2 > 0 \Gamma (b) > (a)$'s.

  If $y = c^t_2$ then $uv^2 x y^2 z = a^{m+b^t_1+b^t_2} c^m \notin L$ since $t_1 + t_2 > 0 \Gamma (b) > (a)$'s or $b(n) > (b)$'s.

  Case 4: $v = c^t \Gamma$ then $y = c^t_2$

  Then $uv^2 x y^2 z = a^{m+c^t+c^t_1+c^t_2} \notin L$ since $t_1 + t_2 > 0 \Gamma (c) > (a)$'s.

  Thus there is no breakdown of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$ and for all $i \geq 0$ $uv^i x y^i z$ 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 \( L = \{a^n b^n c^p : p > n > 0\} \). 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 \( w = \text{____________} \) Note \( |w| \geq m \).

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

Thus there is no breakdown of \( w \) into \( uvxyz \) such that \( |vy| \geq \Gamma |x| \leq m \) and for all \( i \geq 0 \Gamma uv^i xy^i z \) is in \( L \). Contradiction thus \( L \) is not a CFL. Q.E.D.
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 \( w = \) _______.

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

  Case 1: Neither \( v \) nor \( y \) 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).

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

Exercise: Prove the following is not a CFL by applying the pumping lemma. (answer is at the end of this handout).

Consider \( L = \{ a^{2n} b^p c^n d^p : n, p \geq 0 \} \). Show \( L \) is not a CFL.
Example: Consider $L = \{w\bar{w}w : w \in \Sigma^* \}$ where $\bar{w}$ is the string $w$ with each occurrence of $a$ replaced by $b$ and each occurrence of $b$ replaced by $a$. For example, $w = baaa \bar{w} = abbb \bar{w} = baaaabbb$. 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 $w = \ldots$

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

Thus there is no breakdown of $w$ into $uvwxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$, and for all $i \geq 0$, $uv^i xy^i z \in L$. Contradiction, thus $L$ is not a CFL. Q.E.D.
Example: Consider \( L = \{a^n b^p b^n a^m\} \). \( L \) 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:

If you apply the pumping lemma to a CFL then you should find a partition of \( w \) that works!

Chap 8.2 Closure Properties of CFL’s

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

- **Proof:**
  Given 2 CFG \( G_1 = (V_1, T_1, R_1, S_1) \) and \( G_2 = (V_2, T_2, P_2, S_2) \)

  - Union:
    Construct \( G_3 \) s.t. \( L(G_3) = L(G_1) \cup L(G_2) \).
    \( G_3 = (V_3, T_3, R_3, S_3) \)

  - Concatenation:
    Construct \( G_3 \) s.t. \( L(G_3) = L(G_1) \circ L(G_2) \).
    \( G_3 = (V_3, T_3, R_3, S_3) \)
- Star-Closure

  Construct $G_3$ s.t. $L(G_3) = L(G_1)^*$
  $G_3 = (V_3, T_3, R_3, S_3)$

QED.

**Theorem** CFL’s are NOT closed under intersection and complementation.

- **Proof:**
  - Intersection:
  - 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) This proof is similar to the construction proof in which we showed regular languages are closed under intersection. 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 $L(M_2) = L_2$.

Example of replacing arcs (NOT a Proof!):
Note this is not a proof but sketches how we will combine the DFA and NPDA. We must formally define $\Delta_3$. If

then

Must show

if and only if

Must show:

\[ w \in L(M_3) \iff w \in L(M_1) \text{ and } w \in L(M_2). \]

QED.
Questions about CFL:

1. Decide if CFL is empty?

2. Decide if CFL is infinite?

Example: Consider $L = \{a^{2n}b^{2m}c^nd^n : n, m \geq 0\}$. 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 $w = a^{2m}b^{2m}c^md^m$.

Show there is no division of $w$ into $uxyz$ such that $|vy| \geq 1\Gamma|vxy| \leq m\Gamma$ and $uv^i xy^i z \in L$ for

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 or $b$’s or $c$’s or $d$’s (not mixed).

Case 2: $v = a^i \Gamma$ then $y = a^iz$ or $b^i \Gamma$ ($|vxy| \leq m$)

If $y = a^i \Gamma$ then $uv^2 xy^2 z = a^{2m+i+t_1}b^{2m+i}c^md^n \notin L$ since $t_1 + t_2 > 0\Gamma$ the number of $a$’s is not twice the number of $c$’s.

If $y = b^i \Gamma$ then $uv^2 xy^2 z = a^{2m+i+t_2}b^{2m+i}c^md^n \notin L$ since $t_1 + t_3 > 0\Gamma$ the number of $a$’s (denoted $n(a)$) is not twice $n(c)$ or $n(b)$ is not twice $n(d)$.

Case 3: $v = b^i \Gamma$ then $y = b^iz$ or $c^i \Gamma$

If $y = b^i \Gamma$ then $uv^2 xy^2 z = a^{2m+2i+t_1}b^{2m+i}c^md^n \notin L$ since $t_1 + t_2 > 0\Gamma$ the number of $b$’s is not twice $n(b)$.

If $y = c^i \Gamma$ then $uv^2 xy^2 z = a^{2m+i}b^{2m+i+t_2}c^md^n \notin L$ since $t_1 + t_3 > 0\Gamma$ the number of $a$’s is not twice $n(c)$ or $2n(c)>n(a)$.

Case 4: $v = c^i \Gamma$ then $y = c^iz$ or $d^i \Gamma$

If $y = c^i \Gamma$ then $uv^2 xy^2 z = a^{2m+i}b^{2m+i+t_3}c^md^n \notin L$ since $t_1 + t_3 > 0\Gamma$ the number of $c$’s is not twice $n(c)$.

If $y = d^i \Gamma$ then $uv^2 xy^2 z = a^{2m+i}b^{2m+i}c^{m+i}d^n \notin L$ since $t_1 + t_2 > 0\Gamma$ the number of $d$’s is not twice $n(d)$.

Case 5: $v = d^i \Gamma$ then $y = d^iz$

then $uv^2 xy^2 z = a^{2m+i}b^{2m+i}c^{m+i}d^{m+i} \notin L$ since $t_1 + t_2 > 0\Gamma$ the number of $d$’s is not twice $n(d)$.

Thus there is no breakdown of $w$ into $uxyz$ such that $|vy| \geq 1\Gamma|vxy| \leq m$ and for all $i \geq 0\Gamma uv^i xy^i z$ is in $L$. Contradiction! thus $L$ is not a CFL. Q.E.D.