Which of the following languages are CFL?

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

**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 $i = 0, 1, 2, \ldots$.
  Case 1: Neither $v$ nor $y$ can contain 2 or more distinct symbols. If $v$ contains $a$'s and $b$'s, then $uv^2xy^2z \not\in L$ since there will be $b$'s before $a$'s.
  Thus, $v$ and $y$ can be only $a$'s, $b$'s, or $c$'s (not mixed).
  Case 2: $v = a^t$, then $y = a^{t_2}$ or $b^{t_3}$ ($|vxy| \leq m$)
  If $y = a^{t_2}$, then $uv^2xy^2z = a^{m+t_1+t_2}b^mc^m \not\in L$ since $t_1 + t_2 > 0$, $n(a) > n(b)$'s (number of $a$'s is greater than number of $b$'s)
  If $y = b^{t_3}$, then $uv^2xy^2z = a^{m+t_1}b^{m+t_3}c^m \not\in L$ since $t_1 + t_3 > 0$, either $n(a) > n(c)$'s or $n(b) > n(c)$'s.
  Case 3: $v = b^{t_1}$, then $y = b^{t_2}$ or $c^{t_3}$
  If $y = b^{t_2}$, then $uv^2xy^2z = a^mb^{m+t_1+t_2}c^m \not\in L$ since $t_1 + t_2 > 0$, $n(b) > n(a)$'s.
  If $y = c^{t_3}$, then $uv^2xy^2z = a^mb^{m+t_1}c^{m+t_3} \not\in L$ since $t_1 + t_3 > 0$, either $n(b) > n(a)$'s or $n(c) > n(a)$'s.
  Case 4: $v = c^{t_1}$, then $y = c^{t_2}$
  then, $uv^2xy^2z = a^mb^mc^{m+t_1+t_2} \not\in L$ since $t_1 + t_2 > 0$, $n(c) > n(a)$'s.
  Thus, there is no breakdown of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$ and for all $i \geq 0$, $uv^ixy^iz$ 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 \( L = \{a^n b^n c^p : p > n > 0 \} \). 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 \( w = \) __________. Note \( |w| \geq m \).

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

Thus, there is no breakdown of \( w \) into \( u v x y z \) such that \( |v y| \geq 1, |v x y| \leq m \) and for all \( i \geq 0, u v^i x y^i z \) 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 $w = \ldots$
  
  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).

Thus, there is no breakdown of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$ and for all $i \geq 0$, $uv^i xy^i z$ 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 $L = \{a^{2n} b^{2p} c^n d^p : n, p \geq 0 \}$. Show $L$ is not a CFL.
Example: Consider $L = \{ w\bar{w} : w \in \Sigma^* \}$, $\Sigma = \{a, b\}$, where $\bar{w}$ is the string $w$ with each occurrence of $a$ replaced by $b$ and each occurrence of $b$ replaced by $a$. For example, $w = baaa$, $\bar{w} = abbb$, $w\bar{w} = baaaabbb$. 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 $w = \underline{w}$. Show there is no division of $w$ into $uvxyz$ such that $|vy| \geq 1$, $|vxy| \leq m$, and $uv^ixy^iz \in L$ for $i = 0, 1, 2, \ldots$.

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:** Consider $L = \{ a^n b^p b^q a^n \}$. $L$ 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:

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

  - Union:
    Construct $G_3$ s.t. $L(G_3) = L(G_1) \cup L(G_2)$.
    $G_3 = (V_3, T_3, S_3, P_3)$

  - Concatenation:
    Construct $G_3$ s.t. $L(G_3) = L(G_1) \circ L(G_2)$.
    $G_3 = (V_3, T_3, S_3, P_3)$
– Star-Closure
  Construct $G_3$ s.t. $L(G_3) = L(G_1)^*$
  $G_3 = (V_3, T_3, S_3, P_3)$

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 $w \in L(M_1)$ and $w \in L(M_2)$.

QED.
Questions about CFL:

1. Decide if CFL is empty?

2. Decide if CFL is infinite?

Example: Consider $L = \{a^{2n}b^{2m}c^{n}d^{m} : n, m \geq 0\}$. 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 $uvwz$ such that $|vy| \geq 1$, $|vxy| \leq m$, and $uv^{i}xy^{j}z \in L$ 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 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 $uv^{2}xy^{2}z = a^{2m+t_{1}+t_{2}}b^{2m}c^{m}d^{m} \notin L$ since $t_{1} + t_{2} > 0$, the number of $a$'s is not twice the number of $c$'s.

  If $y = b^{t_{3}}$, then $uv^{2}xy^{2}z = a^{2m+t_{1}+t_{2}}b^{2m+t_{3}}c^{m}d^{m} \notin L$ since $t_{1} + t_{3} > 0$, either the number of $a$’s (denoted $n(a)$) is not twice $n(c)$ or $n(b)$ is not twice $n(d)$.

  **Case 3**: $v = b^{t_{1}}$, then $y = b^{t_{2}}$ or $c^{t_{3}}$

  If $y = b^{t_{2}}$, then $uv^{2}xy^{2}z = a^{2m}b^{2m+t_{1}+t_{2}}c^{m}d^{m} \notin L$ since $t_{1} + t_{2} > 0$, $n(b) > 2n(d)$.

  If $y = c^{t_{3}}$, then $uv^{2}xy^{2}z = a^{2m}b^{2m+t_{1}}c^{m+t_{3}}d^{m} \notin L$ since $t_{1} + t_{3} > 0$, either $n(b) > 2n(d)$ or $2n(c) > n(a)$.

  **Case 4**: $v = c^{t_{1}}$, then $y = c^{t_{2}}$ or $d^{t_{3}}$

  If $y = c^{t_{2}}$, then $uv^{2}xy^{2}z = a^{2m}b^{2m+t_{1}+t_{2}}c^{m}d^{m} \notin L$ since $t_{1} + t_{2} > 0$, $2n(c) > n(a)$.

  If $y = d^{t_{3}}$, then $uv^{2}xy^{2}z = a^{2m}b^{2m+t_{1}}c^{m+t_{3}}d^{m} \notin L$ since $t_{1} + t_{3} > 0$, either $2n(c) > n(a)$ or $2n(d) > n(b)$.

  **Case 5**: $v = d^{t_{1}}$, then $y = d^{t_{2}}$

  then $uv^{2}xy^{2}z = a^{2m}b^{2m+t_{1}+t_{2}}c^{m}d^{m} \notin L$ since $t_{1} + t_{2} > 0$, $2n(d) > n(c)$.

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