CSCI/MATH 387: Computability and Complexity — Fall 2022
Reed College — Department of Computer Science
Handout 3: Notes on Pumping Lemma for CFLs — **Instructor:** Chanathip Namprempre

---

Contrapositive Form of Pumping Lemma for Context-Free Languages

> **If** **for any** number $p \geq 0$
>   **there exists** a string $s \in L$ with $|s| \geq p$
>   so that **for any** parse of $s = uvxyz$ with $|vy| \geq 1$ and $|vxy| \leq p$
>   **there exists** some $i \geq 0$ such that $uv^ixy^iz \notin L$,
> **then** L is not context-free.

Example 1

Let the input alphabet be $\{0, 1\}$. Show that $L_1 = \{0^n1^n0^n \mid n \geq 0\}$ is not context-free.

Proof. Let $p \geq 0$. Consider $s = 0^p1^p0^p \in L_1$.
Let $s = uvxyz$ with $|vy| \geq 1$ and $|vxy| \leq p$.

Case 1: vy contains only 0's. Then, $t = uv^2xy^2z$ contains more 0s and than 1s. Thus, $t \notin L_1$.

Case 2: vy contains only 1's. Then, $t = uv^2xy^2z$ contains more 1s and than 0s. Thus, $t \notin L_1$.

Case 3: vy contains a mix of 0's and 1's. Since $|vxy| \leq p$, we know that vy cannot simultaneously contain 0s in both the front segment and the back segment. Let A be the segment of 0's with which vxy overlaps, and let B be the other segment of 0's. (So $A, B \in$ { front, back }.) Thus, $t = uv^2xy^2z$ contains more 0's in the A segment, and $t \notin L_1$.

Therefore, L_1 is not context-free. □

Example 2

Let the input alphabet be $\{0, 1\}$. Show that $L_2 = \{ww \mid w \in \{0, 1\}^*\}$ is not context-free.

Proof. Let $p \geq 0$. Consider $s = 0^p1^p0^p1^p \in L_2$.
Let $s = uvxyz$ with $|vy| \geq 1$ and $|vxy| \leq p$.

Case 1: vxy is exclusively in the first half of s. Then, the midpoint of $t = uv^2xy^2z$ is to the right of the old midpoint. Thus, the first half of t begins with 0 while the second half of t begins with a 1, and $t \notin L_2$.

Case 2: vxy is exclusively in the second half of s. Then, the midpoint of $t = uv^2xy^2z$ is to the left of the old midpoint. Thus, the first half of t ends with 0 while the second half of t ends with a 1, and $t \notin L_2$.

Case 3: vxy straddles the middle part of s. Then, $t = uv^0xy^0z = uxz$ is of the form $0^p1^i0^j1^p$. We know that i and j cannot both be p. Thus, $t \notin L_2$.

Therefore, L_2 is not context-free. □

Example 3

Let the input alphabet be $\{\#, 0, 1\}$. Show that $L_3 = \{r\#s \mid r, s \in \{0,1\}^* \text{ and } r \text{ is a substring of } s\}$ is not context-free.

Proof. Let $p \geq 0$. Consider $s = 0^p 1^p \# 0^p 1^p \in L_3$.
Let $s = uvxyz$ with $|vy| \geq 1$ and $|vxy| \leq p$.

Case 1: vxy is located exclusively before #. Then, the segment of $t = uv^2xy^2z$ preceding # contains more symbols than the segment of t following #. Thus, $t \notin L_3$.

Case 2: vxy is located exclusively after #. Then, the segment of $t = uv^0xy^0z = uxz$ preceding # contains more symbols than the segment of t following #. Thus, $t \notin L_3$.

Case 3: vxy straddles the symbol #. Since $|vxy| \leq p$, we know that vxy overlaps with neither the first stretch of 0s nor the last stretch of 1s.

Case 3.1: v or y contains 1. We know that all the 1s in either v or y must precede #. Consider $t = uv^2xy^2z$. Since there are more 1s preceding # than those following #, the string preceding # is not a substring of the string following #. Thus, $t \notin L_3$.

Case 3.2: v or y contains 0. We know that all the 0s in either v or y must follow #. Consider $t = uv^0xy^0z = uxz$. Since there are more 0s preceding # than those following #, the string preceding # is not a substring of the string following #. Thus, $t \notin L_3$.

Case 3.3: $v = \#$ or $y = \#$. Then, $t = uv^2xy^2z$ contains two #. Thus, $t \notin L_3$.

Therefore, L_3 is not context-free. □
