The Pumping Lemma for Context-free Languages: An Example

Claim 1 *The language*

$$\left\{ww^R w \mid w \in \{0,1\}^*\right\}$$

is not context-free.

Proof: For the sake of contradiction, assume that the language $L = \{ww^R w \mid w \in \{0,1\}^*\}$ is context-free. The Pumping Lemma must then apply; let k be the pumping length. Consider the string

$$s = \overbrace{0^k 1^k}^{w} \overbrace{1^k 0^k}^{w^R} \overbrace{0^k 1^k}^{w} = 0^k 1^{2k} 0^{2k} 1^k \in L.$$

Since $|s| \geq k$, it must be possible to split s into five pieces $uvxyz$ satisfying the conditions of the Pumping Lemma. The substrings v and y must collectively contain some symbols since $|vy| > 0$. We consider the following exhaustive cases.

1. *The substrings v and/or y contain some symbols from the first block of k 0s.* Since $|vxy| \leq k$, v and y cannot contain any 0s from the second block of $2k$ 0s. Consider the string $uv^0xy^0z = uxz$. The string uxz must be of the form $0^{i_1}1^{i_2}0^{2k}1^k$ where $i_1 < k$ and $i_2 \leq 2k$. If $uxz \in L$, it must be of the form $\alpha\alpha^R\alpha$. Since uxz is of the form $0^{i_1}1^{i_2}0^{2k}1^k$ and of length at least $5k$, the first α must begin with the block of $i_1 < k$ 0s followed by some number of 1s. Thus, $\alpha^R\alpha$ must contain a block of at most $2i_1 < 2k$ 0s. But uxz contains a block of $2k$ 0s, a contradiction.

2. *The substrings v and/or y contain some symbols from the first block of $2k$ 1s.* Since $|vxy| \leq k$, v and y cannot contain any 1s from the second block of k 1s. Consider the string $uv^0xy^0z = uxz$. The string uxz must be of the form $0^{i_1}1^{i_2}0^{i_3}1^k$ where $i_1 \leq k$, $i_2 < 2k$, and $i_3 \leq 2k$. If $uxz \in L$, it must be of the form $\alpha\alpha^R\alpha$. Since uxz is of the form $0^{i_1}1^{i_2}0^{i_3}1^k$ and of length at least $5k$, the last α must end with the block of k 1s preceded by some number of 0s. Thus, $\alpha\alpha^R$ must contain a block of $2k$ 1s. But uxz contains a block of $i_2 < 2k$ 1s, a contradiction.

3. *The substrings v and/or y contain some symbols from the second block of $2k$ 0s.* Since $|vxy| \leq k$, v and y cannot contain any 0s from the first block of k 0s. Consider the string $uv^0xy^0z = uxz$. The string uxz must be of the form $0^k1^{i_1}0^{i_2}1^{i_3}$ where $i_1 \leq 2k$, $i_2 < 2k$, and $i_3 \leq k$. If $uxz \in L$, it must be of the form $\alpha\alpha^R\alpha$. Since uxz is of the form $0^k1^{i_1}0^{i_2}1^{i_3}$ and of length at least $5k$, the first α must begin with the block of k 0s followed by some number of 1s. Thus, $\alpha^R\alpha$ must contain a block of $2k$ 0s. But uxz contains a block of $i_2 < 2k$ 0s, a contradiction.

4. *The substrings v and/or y contain some symbols from the second block of k 1s.* Since $|vxy| \leq k$, v and y cannot contain any 1s from the first block of $2k$ 1s. Consider the string $uv^0xy^0z = uxz$. The string uxz must be of the form $0^k1^{2k}0^{i_1}1^{i_2}$ where $i_1 \leq 2k$ and $i_2 < k$. If $uxz \in L$, it must be of the form $\alpha\alpha^R\alpha$. Since uxz is of the form $0^k1^{2k}0^{i_1}1^{i_2}$ and of length at least $5k$, the second α must end with the block of $i_2 < k$ 1s preceded by some number of 0s. Thus, $\alpha\alpha^R$ must contain a block of at most $2i_2 < 2k$ 1s. But uxz contains a block of $2k$ 1s, a contradiction.

Thus, the Pumping Lemma is violated under all circumstances, and the language in question cannot be context-free. $\square$

Note that the choice of a particular string s is critical to the proof. One might think that *any* string of the form ww^Rw would suffice. This is not correct, however. Consider the trivial string $0^k0^k0^k = 0^{3k}$ which is of the form ww^Rw. Letting $v = 0$, $x = \varepsilon$, and $y = 00$, we have $uv^ixy^iz = 0^{3(k+i-1)}$ which is an element of L since it is a string consisting of a multiple of three 0s. The above argument can be generalized for any string of the form p^{3k} where p is a palindrome. Furthermore, seemingly "good" strings such as

$$s = \overbrace{0^k1}^{w}\,\overbrace{10^k}^{w^R}\,\overbrace{0^k1}^{w} = 0^k110^{2k}1$$

can also be pumped: let $v = 0$, $x = 11$, and $y = 00$ (i.e., v consists of the 0 immediately preceding the 11, x is the 11, and y consists of the two 0s immediately following the 11). We then have $uv^ixy^iz = 0^{k+i-1}110^{2(k+i-1)}1$ which is an element of L for all i since $uv^ixy^iz = \alpha\alpha^R\alpha$ where $\alpha = 0^{k+i-1}1$. Again, the choice of the string to be pumped is critical.