Proof of the Pumping Lemma

The language L is regular, so there exists a DFA M such that $L = L(M)$. Say M has p states, $\{q_1, \ldots, q_p\}$. We are also given input string $s \in L$ with $s = s_1 s_2 \cdots s_n$ ($n = |s| \geq p$).

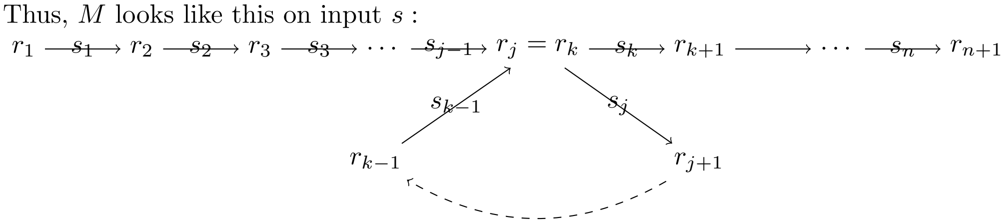
M on input s (accepts):

$$r_1 \xrightarrow{s_1} r_2 \xrightarrow{s_2} r_3 \xrightarrow{s_3} \cdots \xrightarrow{s_{p-1}} r_p \xrightarrow{s_p} r_{p+1} \xrightarrow{s_{p+1}} \cdots \xrightarrow{s_n} r_{n+1}$$

Where r_{n+1} is an accept state. (Remark: the r's are not necessarily unique – r_l and r_m may refer to the same q_p.)

M went through at least $p+1$ states, but has only p distinct states. By pigeonhole principle, some state repeats (there exists a cycle). This implies that there exists some j, k with $j \neq k$ such that $r_j = r_k$. We also know that $k \leq p+1$.

Thus, M looks like this on input s :



Let the input before the loop $s_1 s_2 \cdots s_{j-1} = x$, the input in the loop $s_j \cdots s_{k-1} = y$, and teh input after the loop $s_k \cdots s_n = z$. By assumption, $s = xyz \in L(M)$.

We have shown that

1. For all $i \geq 0$, $xy^i z \in L$ (because we may exploit the loop)
2. $|y| \geq 1$ (because j, k are distinct)
3. $|xy| \leq p$ (because $|xy| = k - 1$ and $k \leq p + 1$.)

This is really useful to show that certain languages are not regular.

Example: Given $L = \{0^n 1^n | n \geq 0\}$, show that L is not regular.
Proof (by contradiction):

1. Assume L is regular
2. There exists a p (pumping length) from pumping lemma
3. Choose $s = 0^p 1^p$ (s satisfies $|s| \geq p$ because $|s| = 2p$)
4. For any x, y, z such that $s = xyz$, $|y| \geq 1$ and $|xy| \leq p$, so y contains only $0s$
5. Choose some i such that $xy^i z \notin L$. Here, we choose $i = 2$. $xy^2 z = xyyz = 0^{p+|y|} 1^p$, which is not in L because $|y| \neq 0$. This contradicts the pumping lemma, which implies that L is not regular.

Example: Given $L = \{ww | w \in \{0,1\}^*\}$, show L is not regular.
Proof (by contradiction):

1. Assume L is regular
2. There exists a p (pumping length) from pumping lemma
3. Choose $s = 0^p 1 0^p$.
4. For any x, y, z such that $s = xyz$ and $|y| \geq 1$ and $|xy| \leq p$, so y contains only $0s$.
5. Choose some i such that $xy^i z \notin L$. Here, we choose $i = 2$. $xy^2 z = xyyz = 0^{p+|y|} 1 0^p$, but $|y| \neq 0$ so this string is not in L, contradicting the pumping lemma. Thus, L is not regular.

   (The example $0^p 0^p$ will not work because it may still remain in the language after pumping in step 5)