Example Proof using the Pumping Lemma for Regular Languages

Andrew P. Black

22 April 2008

Prove that the language $E = \{w \in (01)^* \mid w$ has an equal number of 0s and 1s$\}$ is not regular.

Proof

We prove the required result by contradiction. So, we assume that E *is* regular. Then, by the pumping lemma, there is a pumping length p such that all strings s in E of length p or more can be written as $s = xyz$ where

1. $y \neq \Lambda$
2. $\mid xy \mid \leq p$, and
3. $xy^iz \in E$, for all $i \geq 0$

Consider the string $s = 0^p1^p$. Clearly, $p \in E$ and $\mid s \mid \geq p$, so we should be able to find a decomposition of s into xyz that meets conditions 1–3 above.

How about $x = z = \Lambda,\ y = 0^p1^p$? This meets conditions 1 and 3. But no, it fails to meet condition 2. If $\mid xy \mid \leq p$, then xy *must* contain just 0s and no 1s. Hence, y *must* contain just 0s and no 1s. So, if $s = xyz \in E$, it follows that $xz \notin E$, since xz has fewer 0s than xyz but the same number of 1s.

Thus, we have found a string in E that cannot be pumped, which contradicts the assumption that E is regular.

□

Note that we get to choose a string s to suit our purposes. If, instead, we had chosen $(01)^p$, then we would not have been able to complete the proof. Why not? Because that particular string *can* be pumped. But this is not a problem: the lemma says that *all* strings can be pumped, so all that we need do is find *one* string that cannot be pumped, and we have the contradiction that we are looking for.

What is the minimum pumping length for the language $L = 0001^*$?

The minimum pumping length for a language L is the smallest p such that all strings of length p or more can be pumped.

In $L = 0001^*$:

1. 000 can't be pumped (because 0000 is not in L)
2. 0001 *can* be pumped: put $x = 000,\ y = 1,\ z = \Lambda$.

So the minimum pumping length is 4.

How many states would you expect to find in a DFA recognizing L?