MACM-300: Intro to Formal Languages and Automata

Anoop Sarkar – anoop@cs.sfu.ca

Pumping Lemma. If L is a regular language then there is some $p \geq 1$ such that for every $s \in \Sigma^*$, if $s \in L$ and $|s| \geq p$, then there is some decomposition of s as $s = xyz$, where

1. $xy^iz \in L$ for all $i \geq 0$, and
2. $|y| > 0$ or $y \neq \varepsilon$
3. $|xy| \leq p$

The proof of the pumping lemma is covered in the textbook. Here we will examine the various uses of the pumping lemma.

Typically the pumping lemma is used to prove that a language is not regular. The method is to proceed by contradiction, i.e., to assume (contrary to what we want to prove) that a language L is regular, and derive a contradiction of the pumping lemma. Thus, it is helpful to see exactly what the negation of the pumping lemma is, and for this, we state the pumping lemma as a logical formula (which we can then negate).

Let:

- $\mathcal{N}$ be the set of natural numbers $\{0, 1, 2, \ldots\}$,
- $\mathcal{P}$ be the set of positive numbers $\{1, 2, \ldots\}$,
- $P \equiv (\forall i \in \mathcal{N}\ xy^iz \in L)$
- $A \equiv (s = xyz)$,
- $B \equiv (y \neq \varepsilon)$,
- $C \equiv (|xy| \leq p)$,

The pumping lemma can be stated with respect to the DFA D for each regular language L:

$$\forall D\ \exists p \in \mathcal{P}\ \forall s \in \Sigma^*\ ((s \in L \wedge |s| \geq p) \rightarrow (\exists x, y, z \in \Sigma^*\, A \wedge B \wedge C \wedge P)) \tag{1}$$

where, $\forall$ means for all, $\exists$ means there exists, $\wedge$ means logical AND, and $\rightarrow$ means logical implication ($\wedge$ and $\rightarrow$ are defined in the Boolean Logic section of Sipser, Chp 0).

Note that for a language L that we know to be a regular language, we have a DFA D such that for all strings $s \in L$ there exists the strings $x, y, z \in \Sigma^*$ such that $s = xyz$ for which all the conditions of the pumping lemma are true.

Now let us try to take the negation ($\neg$) of equation (1). From Sipser, Chp 0, we know that:

$$\neg(A \wedge B \wedge C \wedge P) \equiv \neg(A \wedge B \wedge C) \vee \neg P \equiv (A \wedge B \wedge C) \rightarrow \neg P \tag{2}$$

and that:

$$\neg(R \rightarrow S) \equiv R \wedge \neg S \tag{3}$$

The negation of the pumping lemma in equation (1) can be written as:

$$\exists D\ \forall p \in \mathcal{P}\ \exists s \in \Sigma^*\ ((s \in L \wedge |s| \geq p) \wedge (\forall x, y, z \in \Sigma^*\, A \wedge B \wedge C \rightarrow \neg P)) \tag{4}$$

Note that $\neg P \equiv (\exists i \in \mathcal{N}\ xy^iz \notin L)$ and so in order to show that the pumping lemma is contradicted we need to show that for some DFA D for every $p \geq 1$ there is some string $s \in L$ of length at least p such that for every possible decomposition $s = xyz$ satisfying $y \neq \varepsilon$ and $|xy| \leq p$, there is some $i \geq 0$ such that $xy^iz \notin L$. When proving by contradiction, we have a language L that we are assuming to be regular and we can use any DFA D recognizing L. The trick is to find the right candidate $s \in L$.