4 Nonregular languages

Regular versus nonregular languages: Consider the following languages:

- $A = \{0^*1^*\}$
- $B = \{0^*1^n | n \geq 0 \}$
- $C = \{w \in \{0,1\}^* | w \text{ has an equal number of 0s and 1s} \}$
- $D = \{w \in \{0,1\}^* | w \text{ has an equal number of substrings 01 and 10} \}$

4.1 Proving nonregularity by pumping lemma

*Reading: Sipser 1.4 (pp. 77-82)*

**Theorem** (The pumping lemma for regular languages): Let $A$ be a regular language. Then there exists a constant $p$ (the pumping length which depends on $A$) such that $\forall s \in A$ with $|s| \geq p$, we can break $s$ into three substrings $s = xyz$ such that

1. $|y| > 0$;
2. $|xy| \leq p$; and
3. $\forall i \geq 0$, string $xy^iz \in A$.

**Proof:** Let $A = L(M)$ for some DFA $M$ with $p$ states. Consider any $s \in A$ with $s = s_1s_2 \cdots s_n$, where $n \geq p$ and $s_i \in \Sigma$ for $i = 1, 2, \ldots, n$. Assume that $\delta(q_0, s_1 \cdots s_i) = q_i$. On the path $q_0 \rightarrow q_1 \rightarrow \cdots \rightarrow q_n$ that accepts $s$, there are $n + 1 \geq p + 1$ states. By the pigeonhole principle, there are at least two identical states on the path. Let $q_i = q_j$ for some $0 \leq i < j \leq n$ be the first such pair on the path. Now we can break $s = xyz$ as follows:

- $x = s_1 \cdots s_i$;
- $y = s_{i+1} \cdots s_j$; and
- $z = s_{j+1} \cdots s_n$.

We can then easily verify that this partition of $s$ satisfies all three requirements stated in the theorem, i.e., $|y| > 0$, $|xy| \leq p$, and $xy^iz \in A$ for any $i \geq 0$. This completes the proof.

How to use the pumping lemma to prove that a language is not regular:

- Assume that $A$ was regular by contradiction. Then the pumping lemma applies to $A$. Let $p$ be the constant in the pumping lemma.
- Select $s \in A$ with $|s| = f(p) \geq p$.
- By the pumping lemma, $\exists x, y, z$ such that $s = xyz$ with $|y| > 0$, $|xy| \leq p$ and $xy^iz \in A$ for any $i \geq 0$.
- For any $x, y, z$ such that $s = xyz$, $|y| > 0$, and $|xy| \leq p$, find $i \geq 0$ such that $xy^iz \notin A$. A contradiction!

**Example** (Sipser p. 80): Prove that $B = \{0^*1^n | n \geq 0 \}$ is not regular.

**Example:** Prove that $A = \{1^r | r \text{ is a prime} \}$ is not regular.

**Example** (Sipser p. 81): Prove that $A = \{ww | w \in \{0,1\}^* \}$ is not regular.

**Example:** Prove that $A = \{(01)^a0^b | a > b \geq 0 \}$ is not regular.

**Example:** Prove that $A = \{0^n0^m | m \neq n \}$ is not regular.

4.2 Proving nonregularity by closure properties

To prove that $A$ is not regular, assume it was. Find a regular language $B$ and a language operator that preserves regularity, and then apply the operator on $A$ and $B$ to get a regular language $C$. If $C$ is known to be nonregular, a contradiction is found.

**Example:** Prove that $C = \{w \in \{0,1\}^* | w \text{ has an equal number of 0s and 1s} \}$ is not regular.