FORMAL LANGUAGES AND AUTOMATA, 2024 FALL SEMESTER

Lec 05. Pumping Lemma

Eunjung Kim

Limit of finite automata and tools for investigation

Which of the following languages are regular?

1. $B = \{0^n 1^n : n \geq 0\}$.
2. $C = \{w : w$ has equal number of 0's and 1's$\}$.
3. $D = \{w : w$ has equal number of 01's and 10's$\}$.
4. For a DFA D, the set of strings in $L(D)$ accepted via a computation history visiting all states.

Pumping lemma

Pumping lemma: tool to prove nonregularity

Let A be a regular language. Then there exists a number p (called the pumping length) such that any string $w \in A$ of length at least p, w can be written as $w = xyz$ such that the following holds:

1. $|y| \geq 1$,
2. $|xy| \leq p$,
3. $xy^iz \in A$ for every $i \geq 0$.

Proof idea: DFA for A has a finite (constant) number of states.

Pumping lemma

Pumping lemma: tool to prove nonregularity

Let A be a regular language. Then there exists a number p (called the pumping length) such that any string $w \in A$ of length at least p, w can be written as $w = xyz$ such that the following holds:

1. $|y| \geq 1$,
2. $|xy| \leq p$,
3. $xy^iz \in A$ for every $i \geq 0$.

Proof idea: DFA for A has a finite (constant) number of states.

Pumping lemma, proof

There exists DFA M with $L(M) = A$.

1. Let p be the number of states of this DFA.
2. Consider the accepting computation history $r_0 = q_0, r_1, \ldots, r_s$ for w (with $r_s \in F$) such that $r_{i+1} = \delta(r_i, w_{i+1})$ for all $i = 0, \ldots, s-1$, where w_i is the i-th symbol of w.
3. In the first $q + 1$ states $r_0, \ldots, r_p$, there exist two identical states, say r_a and r_b, with $a \neq b$.
4. Take $x = w_1 \cdots w_a$, $y = w_{a+1} \cdots w_b$ and $z = w_{b+1} \cdots w_s$.
5. It remains to observe that
   - $r_{b+1} = \delta(r_b, w_{b+1}) = \delta(r_a, w_{b+1})$, and thus $w_1 \cdots w_a \cdot w_{b+1} \cdots w_s = x \cdot z = x \cdot y^0 \cdot z$ is accepted with the sequence of states $r_0, \ldots, r_a, r_{b+1}, \ldots, r_s$.
   - Any $x \cdot y^i \cdot z$ is accepted with the sequence

$$r_0, \ldots, r_a, (r_{a+1}, \ldots, r_b)^i, r_{b+1}, \ldots, r_s.$$

Pumping lemma for nonregularity

Pumping lemma

Let A be a regular language. Then there exists a number p such that any string $w \in A$ of length at least p, w can be written as $w = xyz$ such that ....

Recipe: assume that A is regular and p is an unknown (arbitrary) pumping length. Choose a good string s, and show that rewriting $s = xyz$ as required is impossible.

Pumping lemma for nonregularity

Pumping lemma

Let A be a regular language. Then there exists a number p such that any string $w \in A$ of length at least p, w can be written as $w = xyz$ such that ....

Recipe: assume that A is regular and p is an unknown (arbitrary) pumping length. Choose a good string s, and show that rewriting $s = xyz$ as required is impossible.
That is, we use the contraposition of Pumping lemma for proving nonregularity of A

Syntax

1. For every positive number p, ("$\forall p$")
2. there exists $w \in A$ of length at least p such that ("$\exists w \in A$")
3. for every split $w = xyz$ with $|y| \geq 1$ and $|xy| \leq p$ ("$\forall$ splits xyz")
4. there exists $i \geq 0$ with $xy^iz \notin A$ ("$\exists i$").

NONREGULARITY OF $B = \{0^n 1^n : n \geq 0\}$.

SYNTAX

1. For every positive number p, ("$\forall p$")
2. there exists $w \in A$ of length at least p such that ("$\exists w \in A$")
3. for every split $w = xyz$ with $|y| \geq 1$ and $|xy| \leq p$ ("$\forall$ splits xyz")
4. there exists $i \geq 0$ with $xy^i z \notin A$ ("$\exists i$").

$D = \{1^{n^2} : n \geq 0\}$

Syntax

1. For every positive number p, ("$\forall p$")
2. there exists $w \in A$ of length at least p such that ("$\exists w \in A$")
3. for every split $w = xyz$ with $|y| \geq 1$ and $|xy| \leq p$ ("$\forall$ splits xyz")
4. there exists $i \geq 0$ with $xy^iz \notin A$ ("$\exists i$").

$\{w : w$ has equal # of 0's and 1's$\}$

Syntax

1. For every positive number p, ("$\forall p$")
2. there exists $w \in A$ of length at least p such that ("$\exists w \in A$")
3. for every split $w = xyz$ with $|y| \geq 1$ and $|xy| \leq p$ ("$\forall$ splits xyz meeting the conditions")
4. there exists $i \geq 0$ with $xy^i z \notin A$ ("$\exists i$").

$D = \{0^i \cdot 1^j : i > j\}$

SYNTAX

1. For every positive number p, ("$\forall p$")
2. there exists $w \in A$ of length at least p such that ("$\exists w \in A$")
3. for every split $w = xyz$ with $|y| \geq 1$ and $|xy| \leq p$ ("$\forall$ splits xyz")
4. there exists $i \geq 0$ with $xy^iz \notin A$ ("$\exists i$").

$F = \{ww : w \in \{0,1\}^*\}$

Syntax

1. For every positive number p, ("$\forall p$")
2. there exists $w \in A$ of length at least p such that ("$\exists w \in A$")
3. for every split $w = xyz$ with $|y| \geq 1$ and $|xy| \leq p$ ("$\forall$ splits xyz")
4. there exists $i \geq 0$ with $xy^iz \notin A$ ("$\exists i$").

Myhill-Nerode Theorem

Fix an alphabet Σ and let L be a language over Σ.

Distiguishability of two strings by L

We say that $x, y \in \Sigma^*$ is indistinguishable by L if for all $z \in \Sigma^*$,

$$x \cdot z \in L \text{ if and only if } y \cdot z \in L,$$

written as $x \equiv_L z$.

Myhill-Nerode Theorem

Fix an alphabet Σ and let L be a language over Σ.

Myhill-Nerode Theorem

L is regular if and only if the number of equivalence classes of $\equiv_L$ is finite.

Moreover, the number of equivalence classes equals the number of states in a minimal (minimum) DFA.

Myhill-Nerode Theorem

Fix an alphabet Σ and let L be a language over Σ.

Myhill-Nerode Theorem

L is regular if and only if the number of equivalence classes of $\equiv_L$ is finite.

Moreover, the number of equivalence classes equals the number of states in a minimal (minimum) DFA.

- ($\rightarrow$) Define an equivalence relation $\sim_M$ from DFA M with $L(M) = L$. The Myhill-Nerode equivalence $\equiv_L$ is a OOOOOOOOOO of $\sim_M$.
- ($\leftarrow$) Build a DFA N from the equivalence classes of $\equiv_L$.