Теория автоматов и формальных языков

Заблоцкий Данил

19 марта 2024 г.

Оглавление

Глава 1

Введение

Лекция 1: Регулярные языки и регулярные выражения

от 12 фев 8:45

Определение 1.1 (Алфавит, слово)**.** Σ – *алфавит* (как правило, конечный),

$$\Sigma = \{a_1, \dots, a_n\}.$$

Слово над алфавитом Σ – конечный упорядоченный набор символов из Σ.

Σ^* – все слова над Σ.

Пример. $\Sigma = \{$а,б,$\dots$,я$\}$, слова: яблоко, абвежр.

ε – пустое слово.

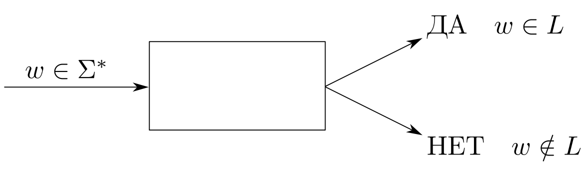
Определение 1.2 (Язык)**.** *Язык* L – подмножество слов над Σ,

$$L \subseteq \Sigma^*.$$

Пример. $\Sigma_2 = \{0, 1\}, \quad \Sigma_2^* = \{\underset{\varepsilon}{\emptyset}, 1, \emptyset\emptyset, \emptyset 1, 1\emptyset, \dots\}$

Замечание (Как «конечно» описать бесконечный язык?)**.**

1. Конечный набор правил построения языков.
2. Алгоритм-распознаватель:



Замечание (Конструкции)**.**

1. L_1, L_2 – языки,

$$L_1 \cup L_2 \text{ – объединение,}$$

$$L_1 \cap L_2 \text{ – пересечение,}$$
$$\overline{L_1} = \Sigma^* \setminus L_1 \text{ – дополнение.}$$

2. Конструкция: $w_1, w_2 \longrightarrow w_1 w_2$,

$$L_1, L_2 \longrightarrow L_1 L_2 = \{w_1 w_2, \quad w_1 \in L_1,\ w_2 \in L_2\}.$$

Пример.

$$\begin{array}{l} L_1 = \{a\} \\ L_2 = \{b\} \end{array} \quad L_1 L_2 = \{ab\},\ \varepsilon b = b.$$

$$\begin{array}{rc} L_1 L_2 = & \{ab\} \\ & \not\parallel \\ L_2 L_1 = & \{ba\} \end{array}$$

Пример.

$$\begin{array}{l} L_1 = \{a, b\} \\ L_2 = \{a, b\} \end{array}, \quad \underbrace{L_1 L_1}_{\text{степень}} = L_1 L_2 = L_1^2 = \{aa, ab, ba, bb\},$$

$$L_1^2, L_1^3 = L_1^2 L_1 = \left\{ \begin{array}{l} aaa, aba, baa, bba, \\ aab, abb, bab, bbb \end{array} \right\}$$

3. Итерация (Ж. Клани)

$$L \longrightarrow L^* = \bigcup_{n=\emptyset}^{\infty} L^n, \quad \boxed{L^1 \subset L},$$

$$L^0 = \{\varepsilon\}.$$

Пример. $\Sigma = \{a, b\}$,

$$L = \{a, b\}, \quad L^* = \left\{ \begin{array}{l} \varepsilon, a, b, aa, ab, ba, bb, \\ aaa, \ldots, bbb, \ldots \end{array} \right\}.$$

> **Определение 1.3** (Регулярная языка, регулярные языки).
>
> 1. $\emptyset, \{\varepsilon\}, \{a_i\},\ a_i \in \Sigma$ – *регулярная языка.*
> 2. L_1, L_2 – *регулярные языки,*
>
> $$L_1 \cup L_2,\ L_1 L_2,\ L_1^* \text{ – тоже регулярные языки.}$$

Пример. $L = \left(\{a\} \underbrace{\left(\{a\} \cup \{b\}\right)^*}_{\substack{\parallel \\ \text{все слова из } a,b}} \{b\} \right)^*$ – все слова, начинающиеся на a и заканчивающиеся на b, например:

$$\underbrace{abbab}\underbrace{ab}\underbrace{aab}.$$

Определение 1.4 (Регулярное выражение).

1. $\emptyset, \{\varepsilon\}, \{a_i\},\ a_i \in \Sigma$ – *регулярные выражения.*
2. Если R_1, R_2 – регулярные выражение, то

$$R_1 + R_2,\ R_1R_2,\ (R_1)^* \text{ – тоже регулярные выражения.}$$

Пример. Любой конечный язык – регулярный.

Пример. $\Sigma = \{a, b\},\ L_1 = \{$все слова из a, b четной длины (включая ε)$\}$:

$$\left\{\begin{array}{l} \varepsilon, aa, ab, ba, bb, \\ aaaa, \dots, bbbb, \dots \end{array}\right\} = \big((a+b)(a+b)\big)^*.$$

$\Sigma = \{a, b\},\ L_2 = \{w :\ $ в w четное число a и $b\}$

$$L_2 \subset L_1 \quad \left(\begin{array}{c} ab \in L_1 \\ ab \notin L_2 \end{array}\right)$$

$$\big((aa+bb)^*(ab+ba)(aa+bb)^*(ab+ba)(aa+bb)^*\big)^*$$

$L_3 = \{w :\ $ в w число a четно$\}$

$$b \dots bab \dots bab \dots b = \big(b^*ab^*ab^*\big)^*$$

Пример (Нерегулярный язык). $L = \{$симметричные слова из $a, b\}$

$$abba,\ aba,\ aa.$$

Литература

[1] Ахо, Ульман – «Теория синтаксического анализа, перевода и компиляции»

[2] Ахо, Сеги, Ульман – «Компиляторы. Принципы, технологии, инструменты»

[3] Серебряков, Галочкин, Гончар, Фуручян – «Теория и реализация языков программирования»