Universidad ICESI

Regular Languages and Regular Expressions

Computation and Discrete Structures III

Juan Marcos Caicedo Mejía – jmcaicedo@icesi.edu.co

Department of Computing and Intelligent Systems
Barberi Faculty of Engineering, Design and Applied Sciences
ICESI University

Material adapted from Professor Oscar Bedoya

2026

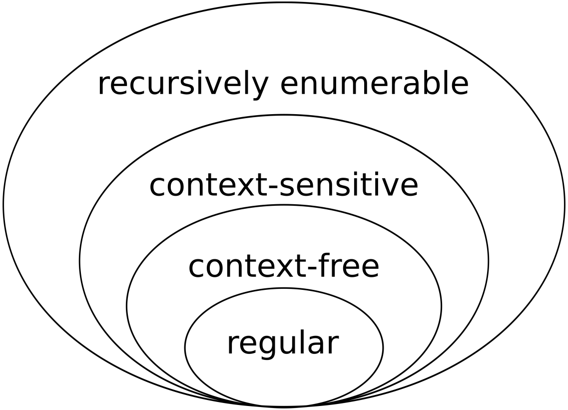
Contents

Contenido

Languages, Machine Types and Grammars

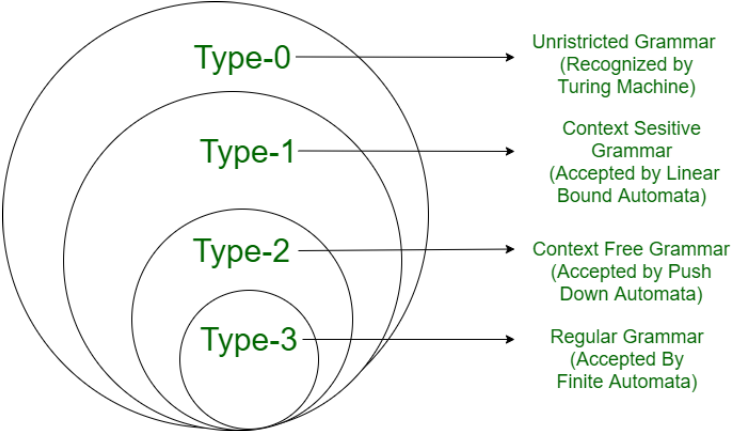
Type	Languages	Machine Type	Grammar Rules
0	Recursively enumerable	Turing machine	Unrestricted
1	Context-sensitive	Linear bounded automaton	$\alpha \to \beta, \quad \lvert\alpha\rvert \leq \lvert\beta\rvert$
2	Context-free	Pushdown automaton	$A \to \gamma$
3	Regular	Finite automaton	$A \to aB$ $A \to a$

Chomsky's Hierarchy



Source: `https://upload.wikimedia.org/wikipedia/commons/9/9a/Chomsky-hierarchy.svg`

Chomsky's Hierarchy

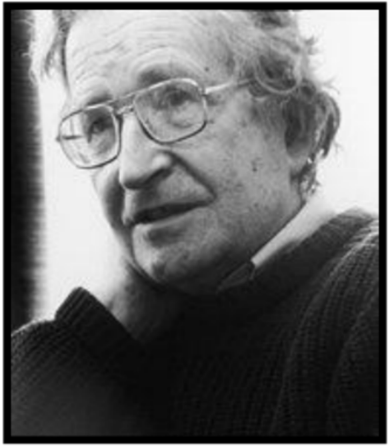


Source: `https://www.geeksforgeeks.org/theory-of-computation/chomsky-hierarchy-in-theory-of-computation/`

Noam Chomsky

Noam Chomsky

- Defined context-free grammars
- Creator of the Chomsky hierarchy (1956)
- Defined Chomsky Normal Form (1979)



(1928 -)

Regular languages

Regular languages

Given an alphabet Σ, regular languages over that alphabet are defined recursively as:

- $\emptyset$ is a regular language
- $\{\epsilon\}$ is a regular language
- For every symbol $a \in \Sigma$, $\{a\}$ is a regular language
- If A and B are regular languages, then
  - $A \cup B$, $A \cdot B$ and A^* are regular languages
- No other language is regular

Regular languages

Given $\Sigma = \{a, b\}$, the following statements are correct:

- $\emptyset$ and $\{\epsilon\}$ are regular languages
- $\{a\}$ and $\{b\}$ are regular languages
- $\{a, b\}$ is regular because it is the union of $\{a\}$ and $\{b\}$
- $\{ab\}$ is regular because it is the concatenation of $\{a\}$ and $\{b\}$
- $\{a, ab, b\}$ is regular because it is the union of two regular languages
- $\{a^n \mid n \geq 0\}$ is regular
- $\{a^m b^n \mid m \geq 0 \wedge n \geq 0\}$ is regular
- $\{(ab)^n \mid n \geq 0\}$ is regular

Regular languages

Given $\Sigma = \{a, b, c\}$, indicate if the following languages are regular:

- $\{a\}^*$
- $\{a\}^* \cup \{b\}^*$
- $\{a\}^* \cdot \{b\}^*$
- $\{a, bc\}^*$
- $\{a\} \cdot \{b, c, ab\}$
- $\{a^n b^n \mid n \geq 0\}$
- $\{a^l b^m c^n \mid l \geq 0, m \geq 0, n \geq 0\}$
- $\{a^n b^{2n} \mid n \geq 0\}$

Regular languages

Given $\Sigma = \{a, b, c\}$, indicate if the following languages are regular:

- $\{a\}^*$
- $\{a\}^* \cup \{b\}^*$
- $\{a\}^* \cdot \{b\}^*$
- $\{a, bc\}^*$
- $\{a\} \cdot \{b, c, ab\}$
- $\{a^n b^n \mid n \geq 0\}$, not regular
- $\{a^l b^m c^n \mid l \geq 0, m \geq 0, n \geq 0\}$
- $\{a^n b^{2n} \mid n \geq 0\}$, not regular

Regular Languages

$\{a^n b^n \mid n \geq 0\}$ is not regular

- $\{a^n \mid n \geq 0\} = \{\epsilon, a, aa, aaa, \ldots\}$
- $\{b^n \mid n \geq 0\} = \{\epsilon, b, bb, bbb, \ldots\}$

$aab \in \{\epsilon, a, aa, aaa, \ldots\} \cdot \{\epsilon, b, bb, bbb, \ldots\}$
but it does not satisfy $a^n b^n$

Regular languages

- Develop the language $L = \{abc, ab, a\}^+$
  $L = \{abc, ab, a, abcabc, abcab, abca, ...\}$

Regular languages

- Develop the language $L = \{abc, ab, a\}^+$
  $L = \{abc, ab, a, abcabc, abcab, abca, ...\}$
- Compare it with $\{abc, ab, a\}^*$

Regular languages

- Develop the language $L = \{abc, ab, a\}^+$
  $L = \{abc, ab, a, abcabc, abcab, abca, ...\}$
- Compare it with $\{abc, ab, a\}^*$
  $\{abc, ab, a\}^* = \{\epsilon, abc, ab, a, abcabc, abcab, abca, ...\}$
  $\{abc, ab, a\}^+ = \{abc, ab, a, abcabc, abcab, abca, ...\}$

Regular languages

- Develop the language $L = \{abc, ab, a\}^+$
  $L = \{abc, ab, a, abcabc, abcab, abca, ...\}$
- Compare it with $\{abc, ab, a\}^*$
  $\{abc, ab, a\}^* = \{\epsilon, abc, ab, a, abcabc, abcab, abca, ...\}$
  $\{abc, ab, a\}^+ = \{abc, ab, a, abcabc, abcab, abca, ...\}$
  $\{abc, ab, a\}^+ = \{abc, ab, a\}^* \cdot \{abc, ab, a\}$
- In general, $A^+ = A^* \cdot A$ holds

Regular languages

Indicate if the following languages are regular:

- $\{ab^na \mid n \geq 0\}$
- $\{a^nb^mc^{n+m} \mid n, m \geq 0\}$
- $\{a^nb^mc^n \mid n, m \geq 0\}$
- $\{(abc)^n \mid n \geq 0\}$
- $\{(abc)^n \cdot (cba)^n \mid n \geq 0\}$
- $\{(ab)^n \cdot (cd)^m \mid m, n \geq 0\}$
- $\{wcw \mid w \in \{a, b\}^*\}$
- $\{w \in \{a, b\}^* \mid |w| = 2k, \text{for } k \geq 0\}$
- $\{aa, ab, ba, bb\}^*$

Regular languages

Develop each of these regular languages:

- $\{a\}^*$
- $\{a\}^* \cup \{b\}^*$
- $\{a\}^* \cdot \{b\}^*$
- $\{a, bc\}^*$
- $\{abc, ab, a\}^+$
- $\{a\} \cdot \{b, c, ab\}$
- $\{(ab)^i \mid i \geq 0\}$
- $\{a^n b^m \mid n \geq 0, m \geq 0\}$
- $\{a^l b^m c^n \mid l \geq 0, m \geq 0, n \geq 0\}$

Regular languages

Develop each of these regular languages:

- $\{a\}^*$
- $\{a\}^* \cup \{b\}^*$, is $ab \in \{a\}^* \cup \{b\}^*$?
- $\{a\}^* \cdot \{b\}^*$
- $\{a, bc\}^*$
- $\{abc, ab, a\}^+$
- $\{a\} \cdot \{b, c, ab\}$
- $\{(ab)^i \mid i \geq 0\}$
- $\{a^n b^m \mid n \geq 0, m \geq 0\}$
- $\{a^l b^m c^n \mid l \geq 0, m \geq 0, n \geq 0\}$

Regular languages

Develop each of these regular languages:

- $\{a\}^*$
- $\{a\}^* \cup \{b\}^*$
- $\{a\}^* \cdot \{b\}^*$, is $bbb \in \{a\}^* \cdot \{b\}^*$?, is $baa \in \{a\}^* \cdot \{b\}^*$?
- $\{a, bc\}^*$
- $\{abc, ab, a\}^+$
- $\{a\} \cdot \{b, c, ab\}$
- $\{(ab)^i \mid i \geq 0\}$
- $\{a^n b^m \mid n \geq 0, m \geq 0\}$
- $\{a^l b^m c^n \mid l \geq 0, m \geq 0, n \geq 0\}$

Regular languages

Develop each of these regular languages:

- $\{a\}^*$
- $\{a\}^* \cup \{b\}^*$
- $\{a\}^* \cdot \{b\}^*$
- $\{a, bc\}^*$, is $bcbca \in \{a, bc\}^*$?, is $baaa \in \{a, bc\}^*$?
- $\{abc, ab, a\}^+$
- $\{a\} \cdot \{b, c, ab\}$
- $\{(ab)^i \mid i \geq 0\}$
- $\{a^n b^m \mid n \geq 0, m \geq 0\}$
- $\{a^l b^m c^n \mid l \geq 0, m \geq 0, n \geq 0\}$

Regular languages

Develop each of these regular languages:

- $\{a\}^* = \{\epsilon, a, aa, aaa, aaaa, ...\}$
- $\{a\}^* \cup \{b\}^* = \{\epsilon, a, aa, aaa, ...\} \cup \{\epsilon, b, bb, bbb, ...\} = \{\epsilon, a, b, aa, bb, aaa, bbb, ...\}$
- $\{a\}^* \cdot \{b\}^* = \{\epsilon, a, aa, aaa, ..., ab, aab, aaab, ..., b, b, bbb, ...\}$
- $\{a, bc\}^* = \{\epsilon, a, bc, aa, abc, bca, bcbca, aaa, ...\}$
- $\{abc, ab, a\}^+ = \{abc, ab, a, abcabc, abcab, abca, ...\}$
- $\{a\} \cdot \{b, c, ab\} = \{ab, ac, aab\}$
- $\{(ab)^i \mid i \geq 0\} = \{\epsilon, ab, abab, ababab, ...\}$
- $\{a^n b^m \mid n \geq 0, m \geq 0\} = \{\epsilon, a, b, ab, aab, abb, aaab, ...\}$
- $\{a^l b^m c^n \mid l \geq 0, m \geq 0, n \geq 0\} = \{\epsilon, a, b, c, ab, bc, abc, aa, aab, aac, ...\}$

Regular languages

Discuss the membership of the following strings given $L = \{a, bc\}^* \cup \{ad, d\}^*$

- Is $bcabc \in L$?
- Is $aabcad \in L$?
- Is $adbc \in L$?
- Is $adad \in L$?
- Is $adddd \in L$?

Regular languages

Discuss the membership of the following strings given $L = \{a, bc\}^* \cup \{ad, d\}^*$

- Is $bcabc \in L$? Yes
- Is $aabcad \in L$? No
- Is $adbc \in L$? No
- Is $adad \in L$? Yes
- Is $adddd \in L$? Yes

Regular languages

Interpret regular language

Interpret the type of words that belong to the following language

$$L = \{a\}^* \cup \{b\}^*$$

Regular languages

Interpret regular language

Interpret the type of words that belong to the following language

$$L = \{a\}^* \cup \{b\}^*$$

Interpretation

Strings that have only a's or only b's. These symbols do not appear mixed

Regular languages

Interpret regular language

Interpret the type of words that belong to the following language

$$L = \{a\}^* \cdot \{b\}^*$$

Regular languages

Interpret regular language

Interpret the type of words that belong to the following language

$$L = \{a\}^* \cdot \{b\}^*$$

Interpretation

Strings that have zero or more a’s followed by zero or more b’s

Regular languages

Given a regular language in natural language, define language structure

Given $\Sigma = \{a, b\}$ define language A of all words that have exactly one a

Regular languages

Given a regular language in natural language, define language structure

Given $\Sigma = \{a, b\}$ define language A of all words that have exactly one a

Language structure

$$A = \{b\}^* \cdot \{a\} \cdot \{b\}^*$$

Regular languages

Given a regular language in natural language, define language structure

Given $\Sigma = \{a, b\}$ define language B of all words that begin with b

Regular languages

Given a regular language in natural language, define language structure

Given $\Sigma = \{a, b\}$ define language B of all words that begin with b

Language structure

$$B = \{b\} \cdot (\{a\} \cup \{b\})^*$$

Regular languages

Given a regular language in natural language, define language structure

Given $\Sigma = \{a, b\}$ define language C of all words that contain the string *ba*

Regular languages

Given a regular language in natural language, define language structure

Given $\Sigma = \{a, b\}$ define language C of all words that contain the string *ba*

Language structure

$$C = (\{a\} \cup \{b\})^* \cdot \{ba\} \cdot (\{a\} \cup \{b\})^*$$

Contenido

Regular expression

A regular expression is a simplified way of representing a regular language

Regular Language	Regular Expression
$\{ab\}$	ab
$\{a\}^*$	a^*
$\{a\}^+$	a^+
$\{a\} \cup \{b\}$	$a \cup b$

Regular expressions

Some regular expressions:

- b^*
- $b(a \cup b)^*$
- $(a \cup b)^* ba(a \cup b)^*$

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that begin with *b* and end with *a*

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that begin with b and end with a

Regular expression

$$b(a \cup b)^* a$$

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have exactly two a's

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have exactly two a's

Regular expression

$$b^*ab^*ab^*$$

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have an even number of a's

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have an even number of a's

Regular expression

$$(b^*ab^*ab^*)^*$$

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have even length

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have even length

Regular expression

$$(aa \cup ab \cup ba \cup bb)^*$$

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have odd length

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have odd length

Regular expression

$$(aa \cup ab \cup ba \cup bb)^*(a \cup b)$$

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have at least one b

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ that have at least one b

Regular expression

$$(a \cup b)^* b (a \cup b)^*$$

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ where the second-to-last symbol is an a

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ where the second-to-last symbol is an a

Regular expression

$$(a \cup b)^* a (a \cup b)$$

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ where the third-to-last symbol is an *a*

Regular expressions

Natural language description of a regular expression

Indicate the **regular expression** that denotes the language of all words over $\Sigma = \{a, b\}$ where the third-to-last symbol is an a

Regular expression

$$(a \cup b)^* a (a \cup b)(a \cup b)$$

Exercises for the reader

Indicate the **regular expression** for each of the following languages:

- All strings over $\{a, b\}$ that begin with aa and end with bb
- All strings over $\{a, b\}$ that have exactly three b's
- All strings over $\{a, b\}$ that begin and end with different symbols

Equivalent regular expressions

1. $r \cup s = s \cup r$
2. $r \cup \emptyset = r = \emptyset \cup r$
3. $r \cup r = r$
4. $(r \cup s) \cup t = r \cup (s \cup t)$
5. $r \cdot \epsilon = \epsilon \cdot r = r$
6. $r \cdot \emptyset = \emptyset \cdot r = \emptyset$
7. $(rs)t = r(st)$
8. $r(s \cup t) = rs \cup rt$
9. $r^* = (r^*)^* = r^*r^* = (\epsilon \cup r)^*$
10. $(r \cup s)^* = (r^* \cup s^*)^* = (r^*s^*)^*$

Equivalent regular expressions

11. $r(sr)^* = (rs)^*r$
12. $(r^*s)^* = \epsilon \cup (r \cup s)^*s$
13. $(rs^*)^* = \epsilon \cup r(r \cup s)^*$
14. $s(r \cup \epsilon)^*(r \cup \epsilon) \cup s = sr^*$
15. $rr^* = r^*r$
16. $(r \cup \epsilon)^* = r^*$

References

Kozen, D. C. (2007)
Automata and computability
Springer Science & Business Media. Lectures 3–4.