

# Lenguajes regulares

Pedro O. Pérez M., PhD.

Implementación de métodos computacionales  
Tecnológico de Monterrey

*pperezm@tec.mx*

02-2023

- ① Conjuntos regulares
  - Definición
  - Lenguajes regulares
  - Aplicaciones
  
- ② Expresiones regulares
  - Operadores
  - Ejemplos

Un conjunto de cadenas es regular si puede generarse a partir del conjunto vacío, el conjunto que contiene la cadena vacía y los conjuntos que contienen un solo elemento del alfabeto usando unión, concatenación y la cerradura Kleene (\*).

**Definición 1** Sea  $\Sigma$  un alfabeto. Los conjuntos regulares sobre  $\Sigma$  se definen recursivamente como sigue:

- ❶ Base:  $\emptyset$ ,  $\{\epsilon\}$  y  $\{a\}$  por cada  $a \in \Sigma$ , son conjuntos regulares sobre  $\Sigma$ .
- ❷ Recursivo: Sean  $X$  e  $Y$  conjuntos regulares sobre  $\Sigma$ , los conjuntos:  $X \cup Y$ ,  $XY$ ,  $X^*$  son conjuntos regulares sobre  $\Sigma$ .
- ❸ Cerradura:  $X$  es un conjunto regulares sobre  $\Sigma$  sólo si puede obtener de los elementos básicos mediante un número finito de aplicaciones del paso recursivo.

Un lenguaje se llama regular si está definido por un conjunto regular. Los siguientes ejemplos muestran cómo se pueden usar conjuntos regulares para describir las cadenas de un lenguaje.

- El conjunto de cadenas que comienzan y terminan con una  $a$  y contienen al menos una  $b$  es regular sobre  $\Sigma = a, b$ .

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

- El conjunto de cadenas que contienen exactamente dos  $b$ 's es regular sobre  $\Sigma = a, b$ .

$$\{a\}^*\{b\}\{a\}^*\{b\}\{a\}^*$$

- Una aplicación común de las expresiones regulares, quizás la más común para la mayoría de los usuarios de computadoras, es la especificación de patrones para buscar documentos y archivos.
- La principal diferencia entre el uso de expresiones regulares para la definición de lenguaje y para la búsqueda de texto es el alcance de la coincidencia deseada. Una cadena está en un lenguaje definido por una expresión regular si y solo si la cadena coincide con el patrón especificado por la expresión regular.

Sobre simplificando, podemos compararlo con las búsqueda fuera de línea (offline) y en línea (online).

- Por búsqueda fuera de línea queremos decir que se ejecuta un programa de búsqueda, la entrada al programa es un patrón y un mosaico, y la salida consiste en las líneas o el texto en el archivo que coincide con el patrón.
- Las herramientas de búsqueda en línea son proporcionadas por navegadores web, editores de texto y sistemas de procesamiento de texto. El objetivo es encontrar interactivamente el primero, el siguiente o encontrar secuencialmente todas las apariciones de subcadenas que coincidan con el patrón de búsqueda.

TABLE 2.2 Extended Regular Expression Operations

Operation	Symbol	Example	Regular Expression
concatenation		$ab$	$ab$
		$[a-c][AB]$	$aA \cup aB \cup bA \cup bB \cup cA \cup cB$
Kleene star	$*$	$[ab]^*$	$(a \cup b)^*$
disjunction	$ $	$[ab]^* A$	$(a \cup b)^* \cup A$
zero or more	$+$	$[ab]^+$	$(a \cup b)^+$
zero or one	$?$	$a?$	$(a \cup \lambda)$
one character	$.$	$a.a$	$a(a \cup b)a$ if $\Sigma = \{a, b\}$
$n$ -times	$\{n\}$	$a\{4\}$	$aaaa = a^4$
$n$ or more times	$\{n, \}$	$a\{4, \}$	$aaaaa^*$
$n$ to $m$ times	$\{n, m\}$	$a\{4, 6\}$	$aaaa \cup aaaaa \cup aaaaaa$

Para cada uno de los siguientes lenguajes definidos sobre el alfabeto  $\Sigma = 0, 1$ , diseña la expresión regular que lo reconozca:

- El conjunto de cadenas que terminan con 00.
- El conjunto de cadenas cuyo décimo símbolo, contando desde el final de la cadena, es un 1.
- El conjunto de cadenas que no contengan la cadena 101.
- El conjunto de número binarios divisibles por 4.



Para cada uno de los siguientes lenguajes, diseña la expresión regular que lo reconozca:

- El conjunto de cadenas de letras minúsculas que empiezan y terminan con  $a$ .
- El conjunto de cadenas de dígitos que no contengan ceros a la izquierda.
- El conjunto de cadenas de dígitos tales que todos los 2 preceden a todos los 9.
- Todas las cadenas de  $a$ 's y  $b$ 's que contienen un número impar de  $a$  o un número impar de  $b$ .
- Todas las cadenas de  $a$ 's y  $b$ 's que contienen exactamente tantas  $a$ 's como  $b$ 's.

Para cada uno de los siguientes lenguajes, diseña la expresión regular que lo reconozca:

- Todas las cadenas de letras minúsculas que comienzan o terminan con una  $a$ .
- Todas las cadenas de dígitos que representan números pares.
- Todas las cadenas de  $a$ 's y  $b$ 's que no contienen  $3b$ 's consecutivas.
- Todas las cadenas de  $a$ 's y  $b$ 's que contienen un número par de  $a$ 's y un número par de  $b$ 's.
- Todas las cadenas de  $0$ 's y  $1$ 's que contengan al menos tres  $1$  consecutivos.
- Todas las cadenas de  $0$ 's y  $1$ 's que comienzan y terminan con el mismo carácter.