

Práctica 03: Expresiones regulares

Computabilidad y Algoritmia

Cheuk Kelly Ng Pante (alu0101364544@ull.edu.es)

01/10/2024

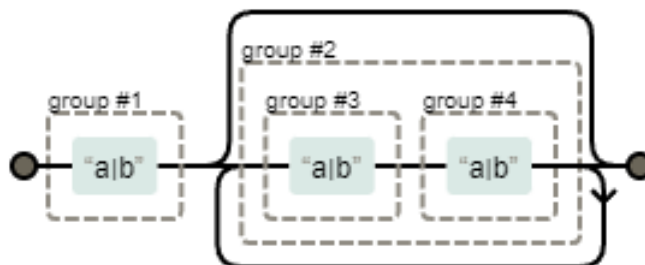
Índice general

1. Ejercicios sobre operadores básicos	1
1.1. Cadenas sobre el alfabeto $\{a, b\}$ con longitud impar.	1
1.2. Cadenas sobre el alfabeto $\{a, b\}$ con longitud igual a 5.	2
1.3. Cadenas sobre el alfabeto $\{a, b, c\}$ con una “a” en la antepenúltima posición.	3
1.4. Cadenas sobre el alfabeto $\{a, b\}$ con número de “a”s par o número de “b”s impar.	4
1.5. Cadenas w sobre el alfabeto $\{0, 1\}$ tales que $2 \leq w \leq 5$	5
1.6. Cadenas sobre el alfabeto $\{0, 1\}$ con longitud múltiplo de 3.	6
1.7. Cadenas sobre el alfabeto $\{0, 1\}$ con una longitud que no sea múltiplo de 3.	7
1.8. Cadenas w sobre el alfabeto $\{0, 1\}$ tal que $w = 0^n 1^m$ con $n + m$ impar.	8
1.9. Cadenas sobre el alfabeto $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ que tengan como máximo dos ceros.	9
1.10. Cadenas sobre el alfabeto $\{x, y, z\}$ que no contenga dos símbolos x consecutivos.	11
2. Ejercicios sobre operadores extendidos	12
2.1. Direcciones de correos electrónicos de estudiantes de la Universidad de La Laguna.	12
2.2. Palabras que terminen por una vocal.	12
2.3. Números enteros.	13
2.4. Texto que se encuentre entre paréntesis.	13
2.5. Código postal en España.	14
2.6. Palabras que contienen solo letras mayúsculas.	14
2.7. Número de teléfono en formato prefijo XXX-XXX-XXX , donde el prefijo del país puede indicarse empezando por 00 o bien con un símbolo +; por ejemplo, 0034 o +34 para España.	15
2.8. Fecha en formato DD/MM/AAAA.	15
2.9. Palabras de al menos 10 letras de longitud.	16
2.10. Palabras que terminen con “ing” o “ed”.	16
3. Modificación	17
3.1. Etiquetas HTML en un documento $< >$	17
3.2. Número decimal	17
3.3. Cadenas binarias con longitud impar que empiecen por 0 y terminen por 1	18

1. Ejercicios sobre operadores básicos

1.1. Cadenas sobre el alfabeto $\{a, b\}$ con longitud impar.

- **Expresión regular:** $(a|b)((a|b)(a|b))^*$



- Cadenas que pertenecen al lenguaje:

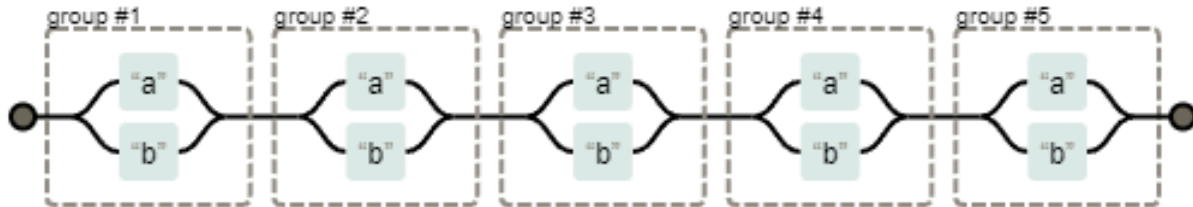
- $w_1 = a$
- $w_2 = aba$
- $w_3 = bbaba$
- $w_4 = aaaaaaaaa$
- $w_5 = abababababababa$

- Cadenas que no pertenecen al lenguaje:

- $w_6 = aa$
- $w_7 = abab$
- $w_8 = bbbbbb$
- $w_9 = baaaaaaaaab$
- $w_{10} = ababababababab$

1.2. Cadenas sobre el alfabeto $\{a, b\}$ con longitud igual a 5.

- Expresión regular: $(a|b)(a|b)(a|b)(a|b)(a|b)$



- Cadenas que pertenecen al lenguaje:

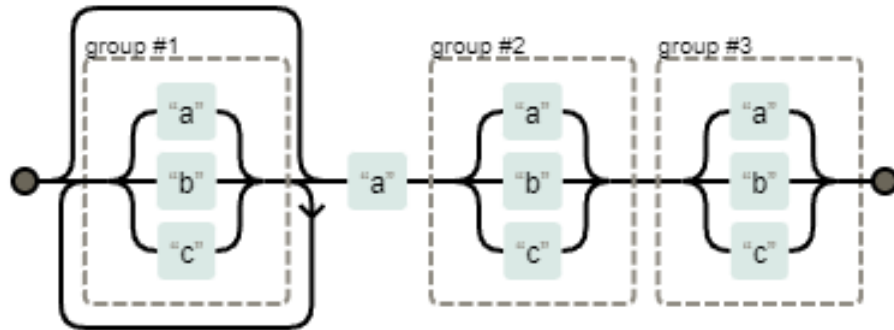
- $w_1 = aaaaa$
- $w_2 = bbbbb$
- $w_3 = ababa$
- $w_4 = babab$
- $w_5 = abbbb$

- Cadenas que no pertenecen al lenguaje:

- $w_6 = aaaa$
- $w_7 = bbb$
- $w_8 = ababab$
- $w_9 = bababa$
- $w_{10} = abababab$

1.3. Cadenas sobre el alfabeto $\{a, b, c\}$ con una “a” en la antepenúltima posición.

- Expresión regular: $(a|b|c)^*a(a|b|c)(a|b|c)$



- Cadenas que pertenecen al lenguaje:

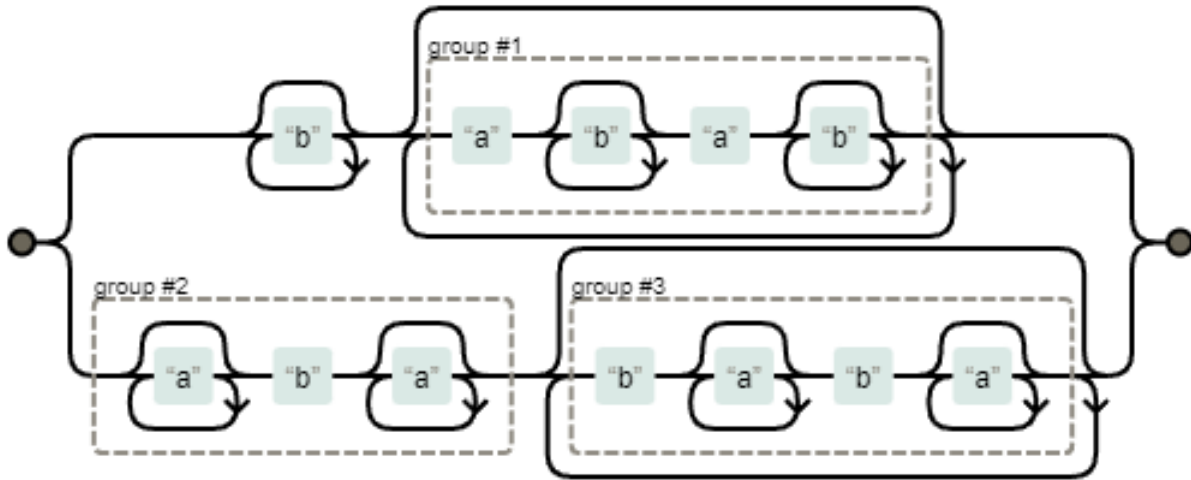
- $w_1 = abc$
- $w_2 = acbaabc$
- $w_3 = aaaaaaabb$
- $w_4 = abbaccbbccacaaa$
- $w_5 = abbbacbbbacacb$

- Cadenas que no pertenecen al lenguaje:

- $w_6 = a$
- $w_7 = bbb$
- $w_8 = aaaccabaabbc$
- $w_9 = abccdacc$
- $w_{10} = ababacccbabab$

1.4. Cadenas sobre el alfabeto $\{a, b\}$ con número de “a”s par o número de “b”s impar.

- Expresión regular: $b^*(ab^*ab^*)^*|(a^*ba^*)(ba^*ba^*)^*$



- Cadenas que pertenecen al lenguaje:

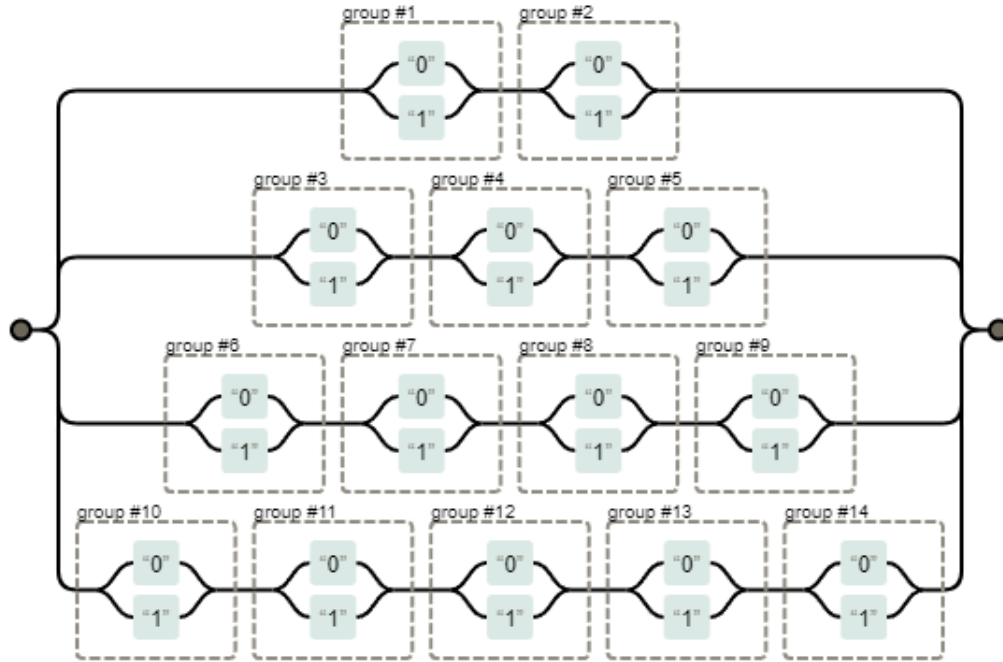
- $w_1 = baa$
- $w_2 = bb aa$
- $w_3 = bbaabbaa$
- $w_4 = bbaabbaabbaa$
- $w_5 = bbaabbaabbaabbaa$

- Cadenas que no pertenecen al lenguaje:

- $w_6 = a$
- $w_7 = bb$
- $w_8 = abb$
- $w_9 = aaaaabbbbb$
- $w_{10} = bbaaabb$

1.5. Cadenas w sobre el alfabeto $\{0, 1\}$ tales que $2 \leq |w| \leq 5$.

- **Expresión regular:** $(0|1)(0|1) \mid (0|1)(0|1)(0|1) \mid (0|1)(0|1)(0|1)(0|1) \mid (0|1)(0|1)(0|1)(0|1)(0|1)$



- **Cadenas que pertenecen al lenguaje:**

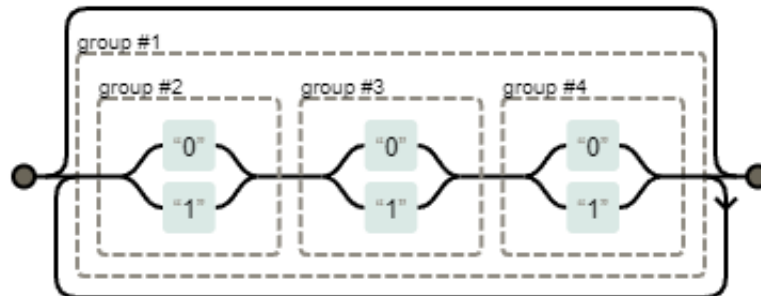
- $w_1 = 00$
- $w_2 = 11$
- $w_3 = 0011$
- $w_4 = 001$
- $w_5 = 10011$

- **Cadenas que no pertenecen al el lenguaje:**

- $w_6 = 1$
- $w_7 = 0$
- $w_8 = 11001100$
- $w_9 = 0000000$
- $w_{10} = 1000101111000$

1.6. Cadenas sobre el alfabeto $\{0, 1\}$ con longitud múltiplo de 3.

- Expresión regular: $((0|1)(0|1)(0|1))^*$



- Cadenas que pertenecen al lenguaje:

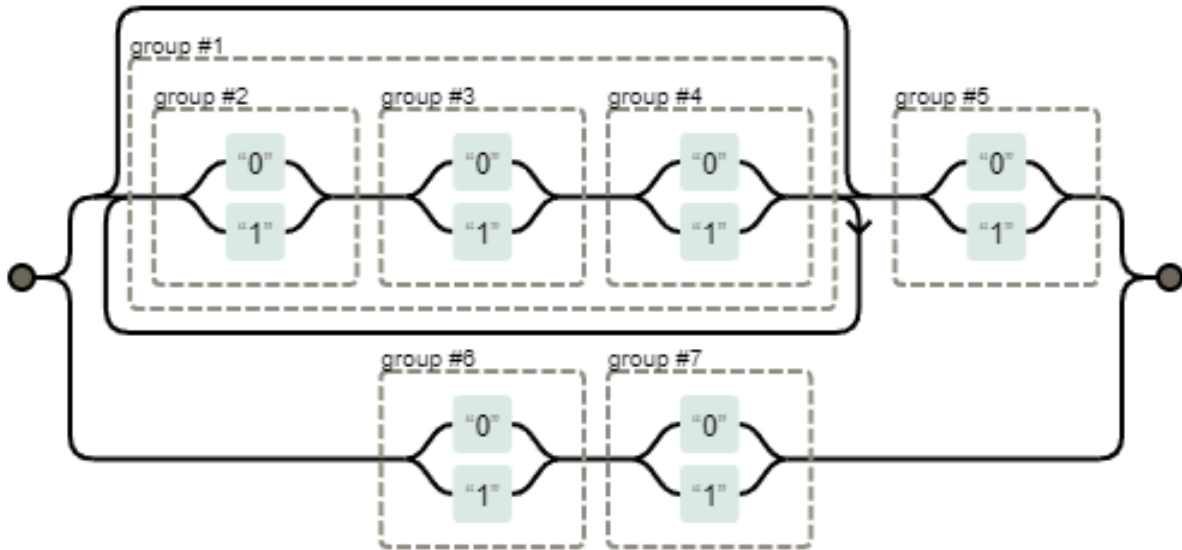
- $w_1 = 000$
- $w_2 = 111000$
- $w_3 = 101010$
- $w_4 = 001001001$
- $w_5 = 110110110110$

- Cadenas que no pertenecen al lenguaje:

- $w_6 = 0$
- $w_7 = 11$
- $w_8 = 1010$
- $w_9 = 00111$
- $w_{10} = 1001001$

1.7. Cadenas sobre el alfabeto $\{0, 1\}$ con una longitud que no sea múltiplo de 3.

- Expresión regular: $((0|1)(0|1)(0|1))^*(0|1)|(0|1)(0|1)$



- Cadenas que pertenecen al lenguaje:

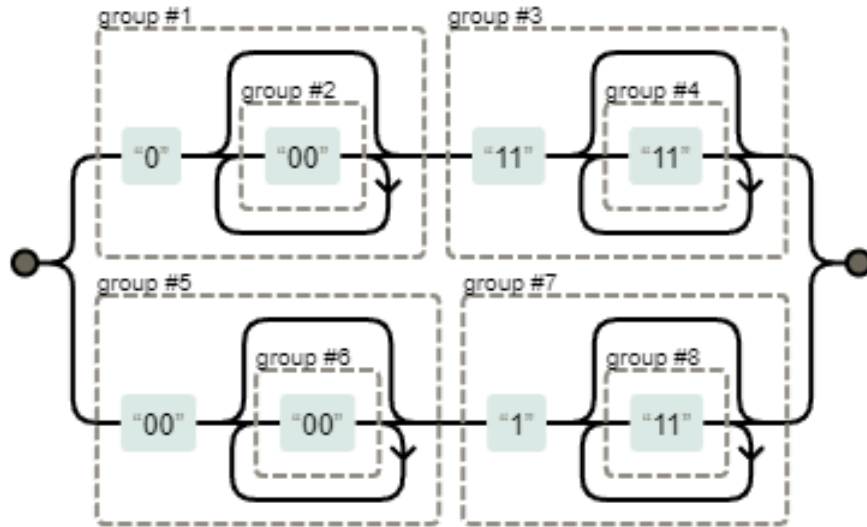
- $w_1 = 0$
- $w_2 = 1$
- $w_3 = 00$
- $w_4 = 1100$
- $w_5 = 0000110$

- Cadenas que no pertenecen al lenguaje:

- $w_6 = 000$
- $w_7 = 111000$
- $w_8 = 101010$
- $w_9 = 001001001$
- $w_{10} = 110110110110$

1.8. Cadenas w sobre el alfabeto $\{0,1\}$ tal que $w = 0^n 1^m$ con $n + m$ impar.

- Expresión regular: $(0(00)^*)(11(11)^*)|(00(00)^*)(1(11)^*)$



- Cadenas que pertenecen al lenguaje:

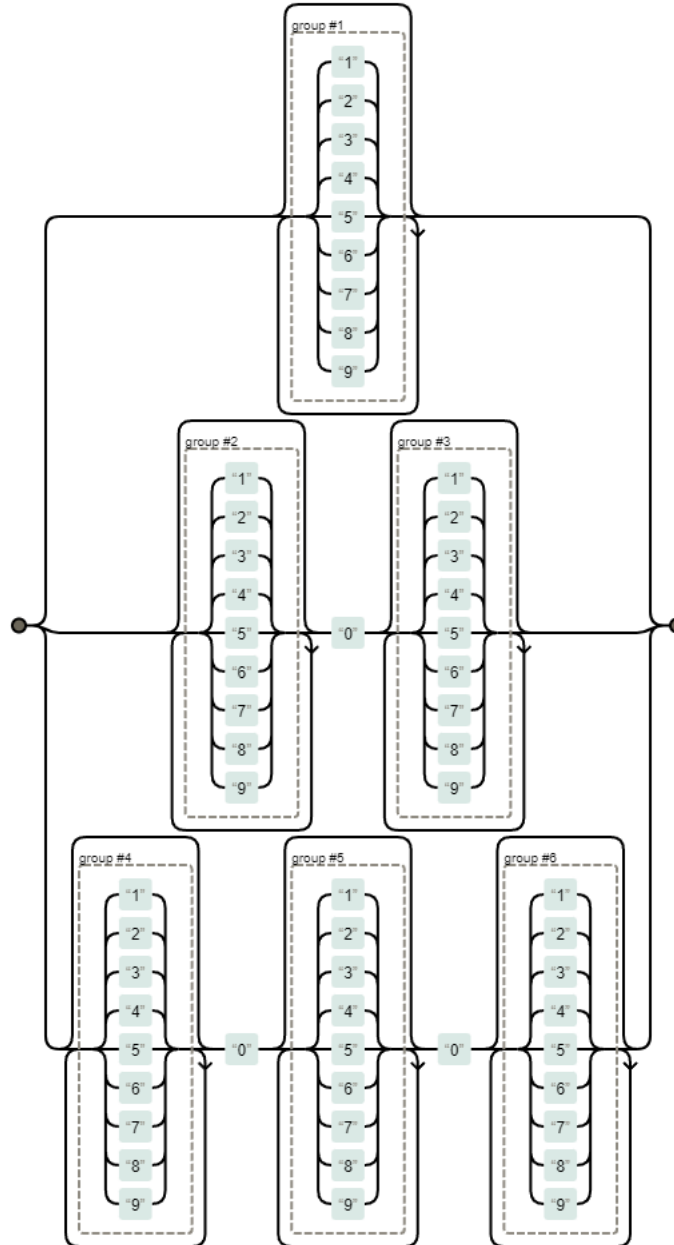
- $w_1 = 01$
- $w_2 = 000111$
- $w_3 = 0000111$
- $w_4 = 000001111$
- $w_5 = 00000011111$

- Cadenas que no pertenecen al lenguaje:

- $w_6 = 0$
- $w_7 = 1$
- $w_8 = 0000$
- $w_9 = 1111$
- $w_{10} = 000011$

1.9. Cadenas sobre el alfabeto $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ que tengan como máximo dos ceros.

- **Expresión regular:** $(1|2|3|4|5|6|7|8|9)^*(1|2|3|4|5|6|7|8|9)^*0(1|2|3|4|5|6|7|8|9)^*|(1|2|3|4|5|6|7|8|9)^*0(1|2|3|4|5|6|7|8|9)^*0(1|2|3|4|5|6|7|8|9)^*$



■ Cadenas que pertenecen al lenguaje:

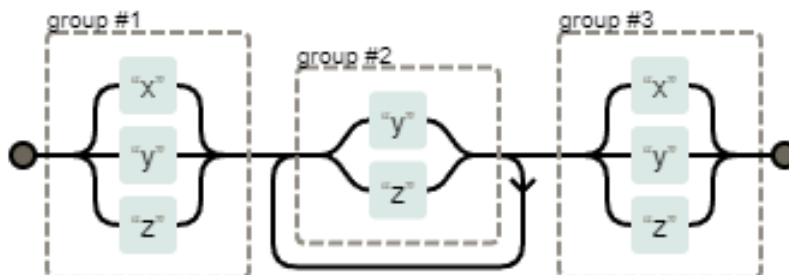
- $w_1 = 0$
- $w_2 = 00$
- $w_3 = 012310$
- $w_4 = 28720099172$
- $w_5 = 123987087871230$

■ Cadenas que no pertenecen al lenguaje:

- $w_6 = 000$
- $w_7 = 013208780$
- $w_8 = 023001003$
- $w_9 = 0289000000134$
- $w_{10} = 00788432100$

1.10. Cadenas sobre el alfabeto $\{x, y, z\}$ que no contenga dos símbolos **x** consecutivos.

- **Expresión regular:** $(x|y|z)(y|z)^+(x|y|z)$



- Cadenas que pertenecen al lenguaje:

- $w_1 = x$
- $w_2 = xyz$
- $w_3 = yzzzzzy$
- $w_4 = xzzzzzy$
- $w_5 = zy$

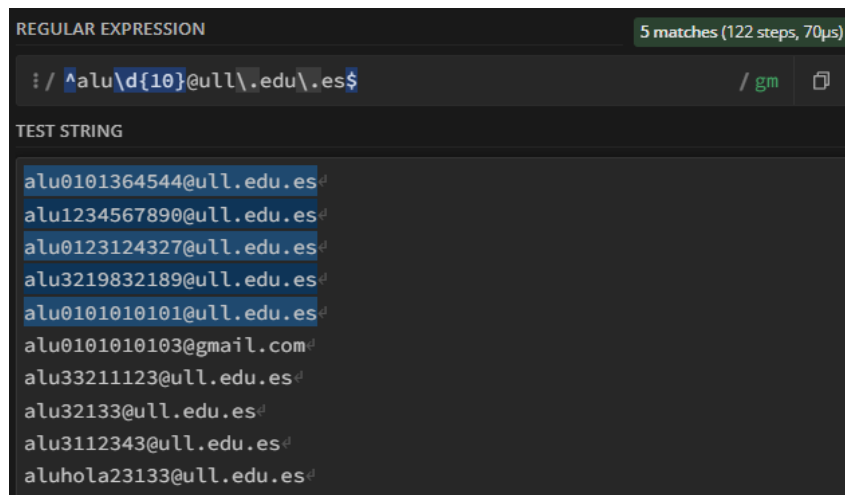
- Cadenas que no pertenecen al lenguaje:

- $w_6 = xx$
- $w_7 = xyx$
- $w_8 = zxx$
- $w_9 = yxx$
- $w_{10} = zxyx$

2. Ejercicios sobre operadores extendidos

2.1. Direcciones de correos electrónicos de estudiantes de la Universidad de La Laguna.

Expresión regular: `^alu\d{10}@ull\.edu\.es$`



REGULAR EXPRESSION 5 matches (122 steps, 70µs)

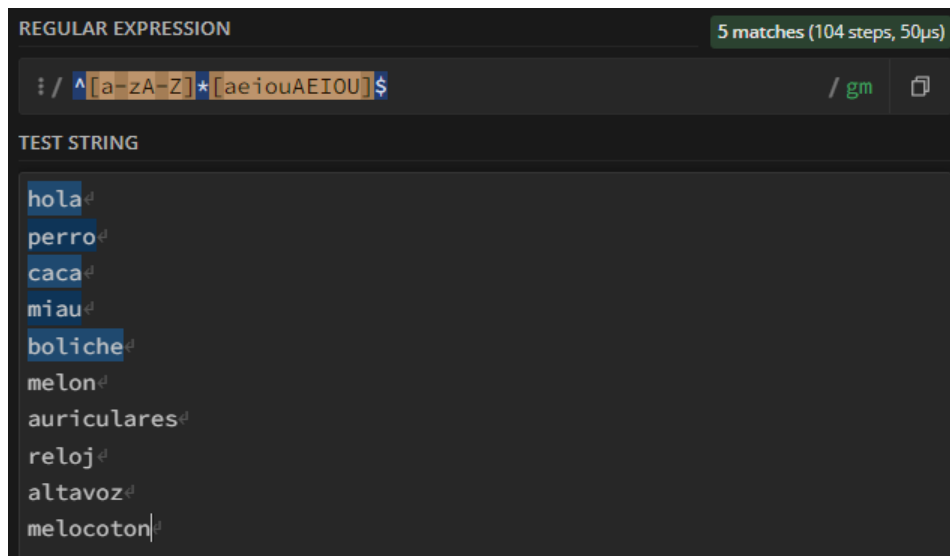
`:/^alu\d{10}@ull\.edu\.es$/gm`

TEST STRING

```
alu0101364544@ull.edu.esd
alu1234567890@ull.edu.esd
alu0123124327@ull.edu.esd
alu3219832189@ull.edu.esd
alu0101010101@ull.edu.esd
alu0101010103@gmail.comd
alu33211123@ull.edu.esd
alu32133@ull.edu.esd
alu3112343@ull.edu.esd
aluhola23133@ull.edu.esd
```

2.2. Palabras que terminen por una vocal.

Expresión regular: `^[a-zA-Z]*[aeiouAEIOU]$`



REGULAR EXPRESSION 5 matches (104 steps, 50µs)

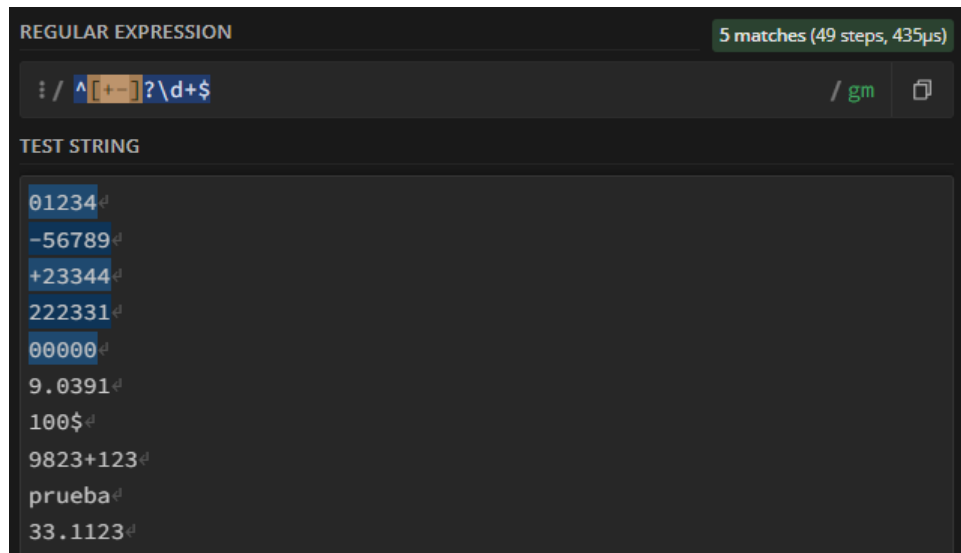
`:/^[a-zA-Z]*[aeiouAEIOU]$/gm`

TEST STRING

```
holad
perrod
cacad
miaud
boliched
melond
auricularesd
relojd
altavozd
melocotond
```

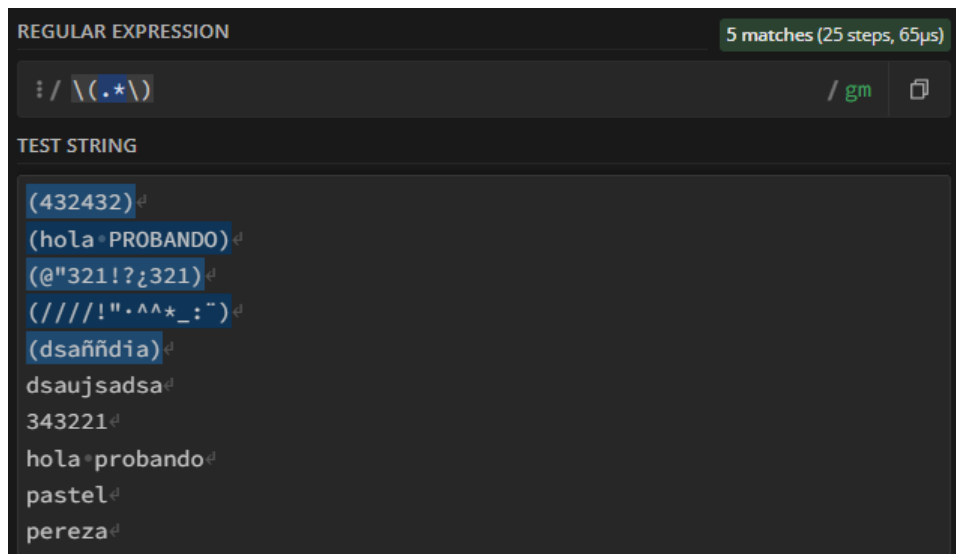
2.3. Números enteros.

Expresión regular: `^[+-]?\d+$`



2.4. Texto que se encuentre entre paréntesis.

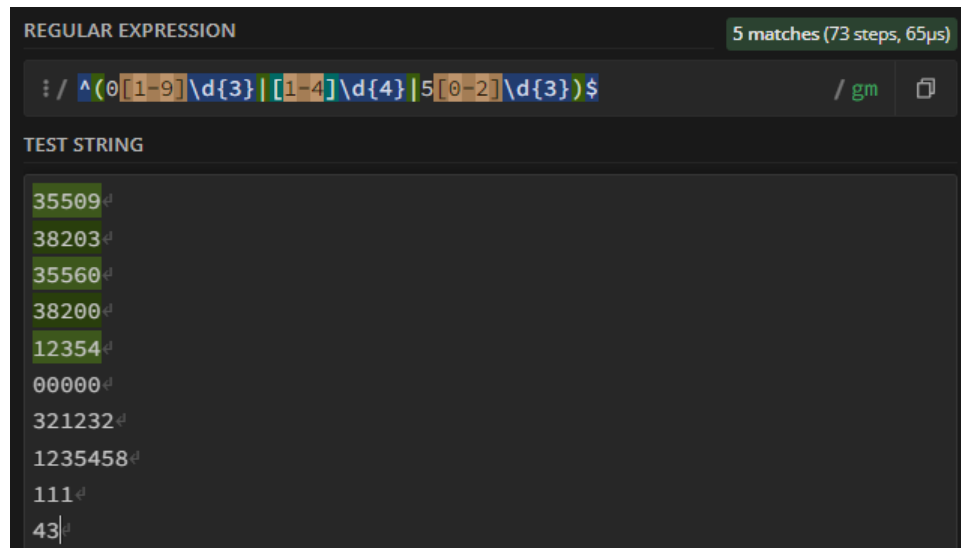
Expresión regular: `\(.*\)`



2.5. Código postal en España.

Expresión regular:

`^(0[1-9]\d{3}|\[1-4\]\d{4}|\5[0-2]\d{3})$`



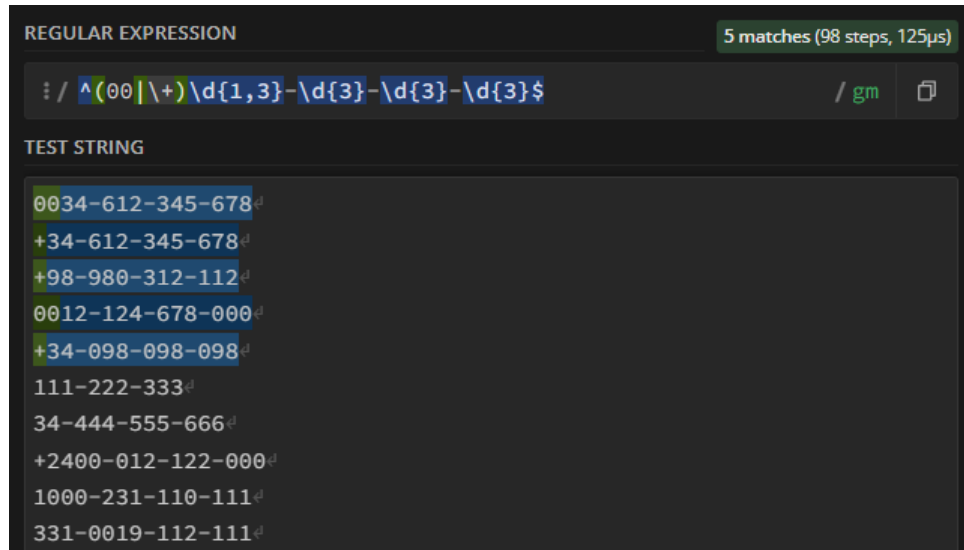
2.6. Palabras que contienen solo letras mayúsculas.

Expresión regular: `^(?=.*[A-Z])[a-zA-Z]+$`



2.7. Número de teléfono en formato prefijo XXX-XXX-XXX, donde el prefijo del país puede indicarse empezando por 00 o bien con un símbolo +; por ejemplo, 0034 o +34 para España.

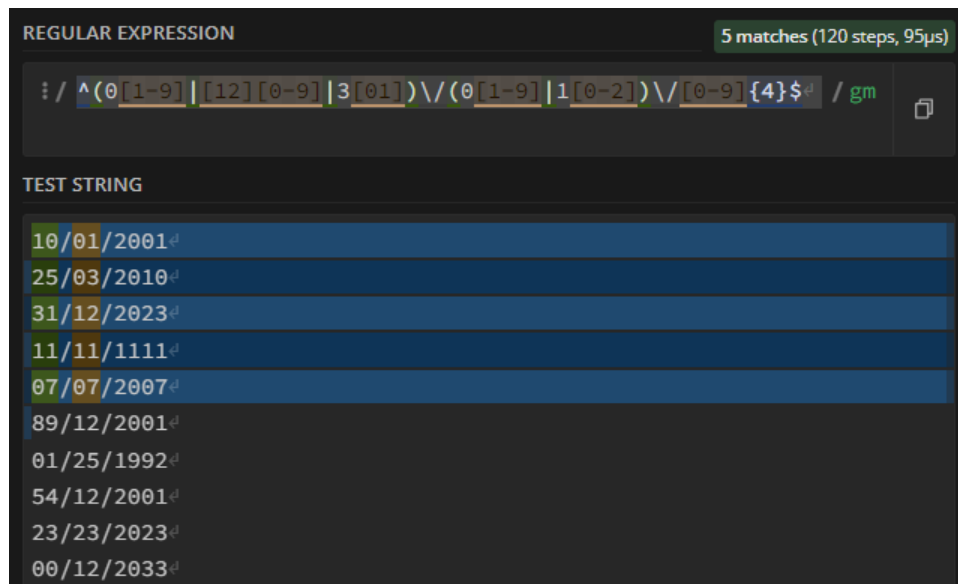
Expresión regular: `^(00|\+)\d{1,3}-\d{3}-\d{3}-\d{3}$`



2.8. Fecha en formato DD/MM/AAAA.

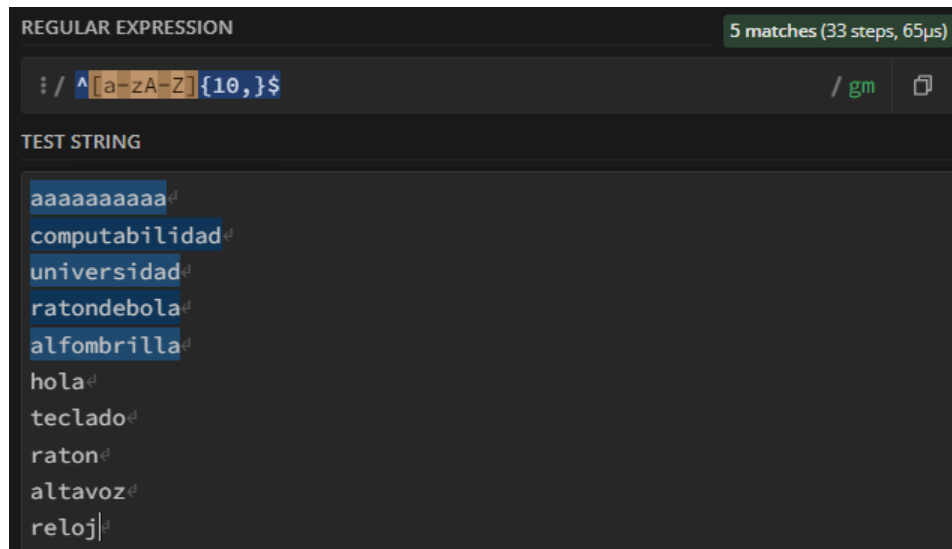
Expresión regular:

`^(0[1-9] | [12] [0-9] | 3[01])\/(0[1-9] | 1[0-2])\[0-9]{4}$`



2.9. Palabras de al menos 10 letras de longitud.

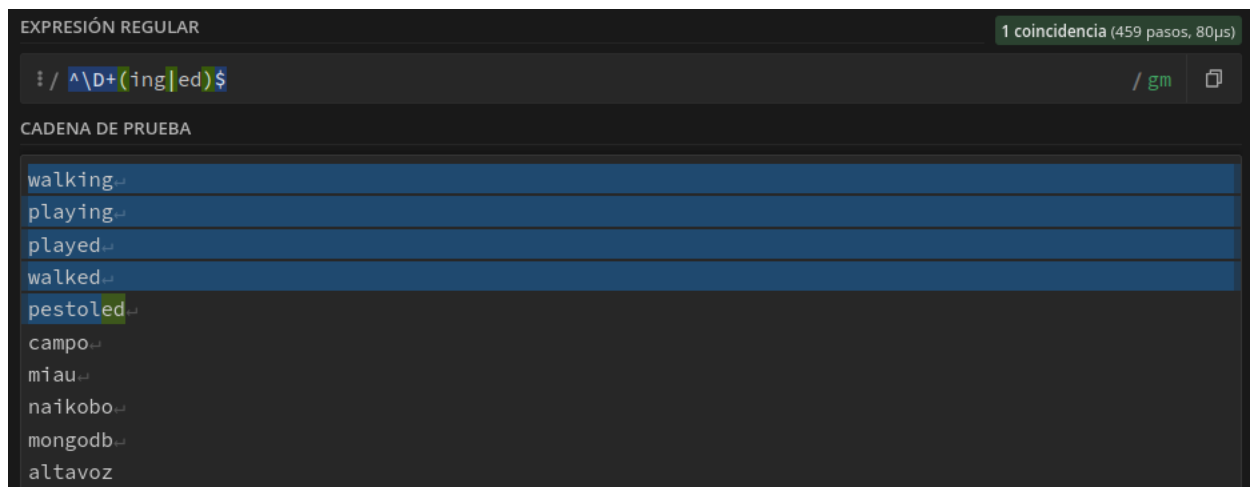
Expresión regular: `^[a-zA-Z]{10,}$`



2.10. Palabras que terminen con “ing” o “ed”.

Expresión regular:

`^\D+(ing|ed)$`



3. Modificación

3.1. Etiquetas HTML en un documento < >

Expresión regular:

`<\s*\/?\s*[a-zA-Z][a-zA-Z0-9]*\s*[>]*>`

EXPRESIÓN REGULAR 11 coincidencias (110 pasos, 100µs)

:/ <\s*\/?\s*[a-zA-Z][a-zA-Z0-9]*\s*[>]*> / gm

CADENA DE PRUEBA

```
<html>
<head>
  <title>Esta es mi primera pagina</title>
</head>
<body>
  <h1>Esto es un encabezado</h1>
  <p>Y esto es un parrafo, donde podemos escribir todo el rollo que se nos ocurra.
</body>
</html>
```

3.2. Número decimal

Expresión regular:

`^[+-]?(\d+(\.\d+)?|\.\d+)$`

EXPRESIÓN REGULAR 5 coincidencias (102 pasos, 80µs)

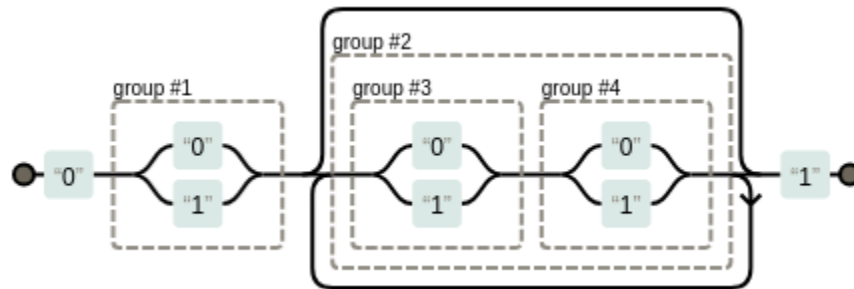
:/ ^[+-]?(\d+(\.\d+)?|\.\d+)\$ / gm

CADENA DE PRUEBA

```
-89.9
11.9
+45.1
.9999
-35
abc
89.99.99
+
-0.9182.11
123.
```

3.3. Cadenas binarias con longitud impar que empiecen por 0 y terminen por 1

Expresión regular: $0(0|1)((0|1)(0|1))^*1$



EXPRESIÓN REGULAR 6 coincidencias (217 pasos, 120µs)

:/ 0(0|1)((0|1)(0|1))*1 / gm

CADENA DE PRUEBA

00001
011
011111101100011
00000011101
01011
0
01
0001
111
10