Práctica 03: Expresiones regulares

Computabilidad y Algoritmia
Grado en Ingeniería Informática
Universidad de La Laguna

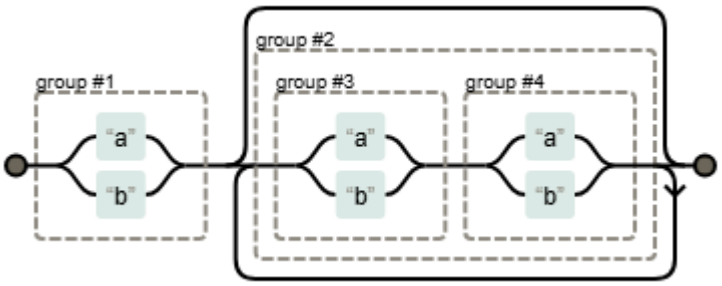
Alejandro Rodríguez Rojas
alu0101317038@ull.edu.es

Índice

1. Ejercicios sobre operadores básicos

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

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

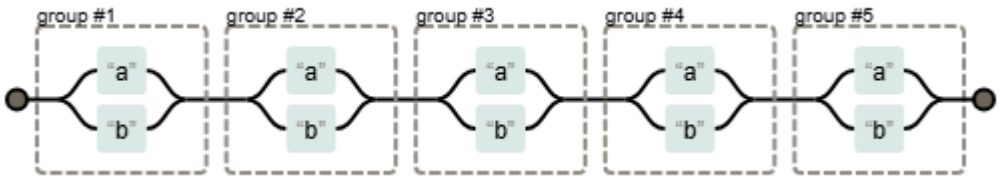


- Cadenas que pertenecen al lenguaje:
  - w1 = a
  - w2 = aba
  - w3 = babaa
  - w4 = aaaaaaa
  - w5 = bbabbba

- Cadenas que no pertenecen al lenguaje:
  - w6 = ba
  - w7 = abab
  - w8 = bbaabb
  - w9 = aa
  - w10 = bababa

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)

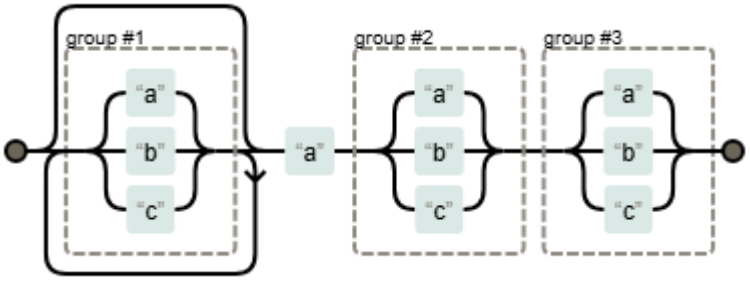


- Cadenas que pertenecen al lenguaje:
  - w1 = aaaaa
  - w2 = ababa
  - w3 = bbabb
  - w4 = bbbbb
  - w5 = aabba

- Cadenas que no pertenecen al lenguaje:
  - w6 = a
  - w7 = bb
  - w8 = bbb
  - w9 = abababababab
  - w10 = aabb

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)

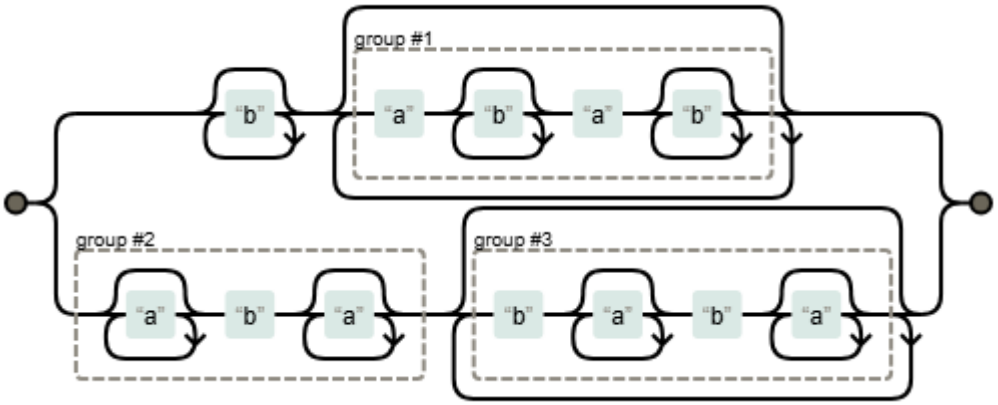


- Cadenas que pertenecen al lenguaje:
  - w1 = abc
  - w2 = abbabb
  - w3 = cabb
  - w4 = aaa
  - w5 = bacc

- Cadenas que no pertenecen al lenguaje:
  - w6 = ccc
  - w7 = bca
  - w8 = a
  - w9 = bbbca
  - w10 = aacccc

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*)*

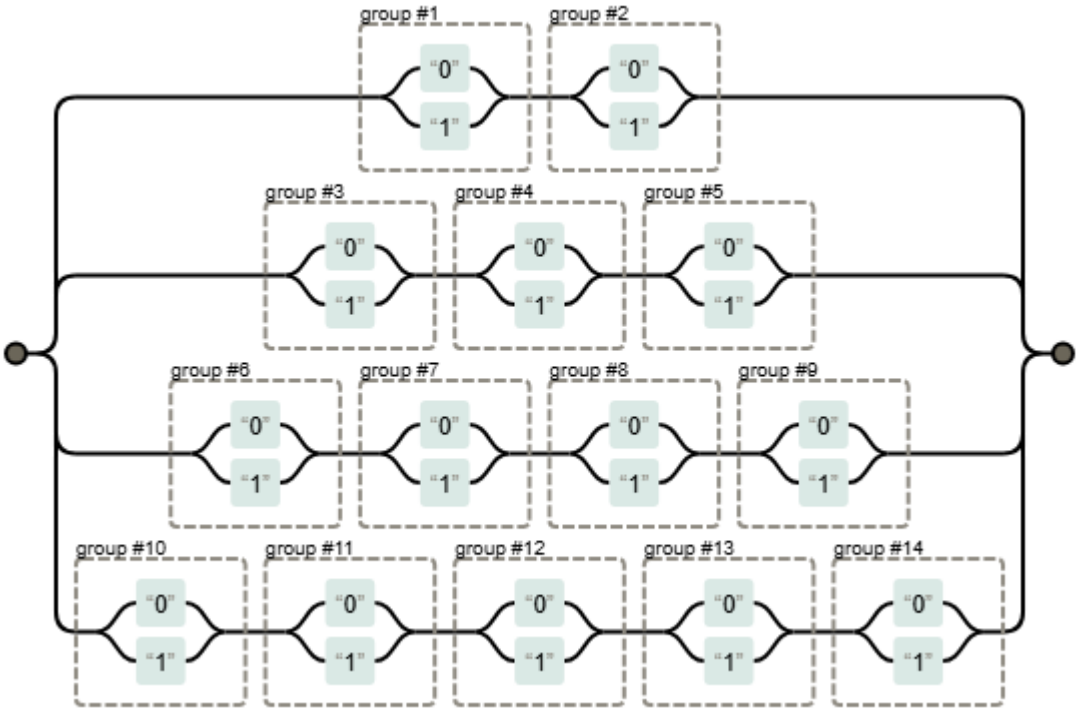


- Cadenas que pertenecen al lenguaje:
    - w1 = baa
    - w2 = aabb
    - w3 = bbba
    - w4 = aa
    - w5 = b

- Cadenas que no pertenecen al lenguaje:
    - w6 = a
    - w7 = bb
    - w8 = abb
    - w9 = aaaaabbbbbb
    - w10 = bbaaabb

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

- Expresión regular:
  (0|1)(0|1) | (0|1)(0|1)(0|1) | (0|1)(0|1)(0|1)(0|1) | (0|1)(0|1)(0|1)(0|1)(0|1)

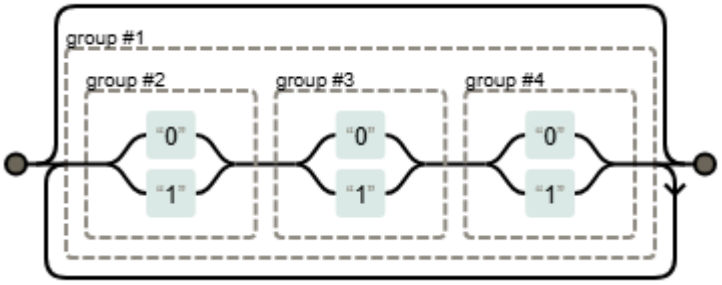


- Cadenas que pertenecen al lenguaje:
  - w1 = 01
  - w2 = 111
  - w3 = 0011
  - w4 = 001
  - w5 = 10011

- Cadenas que no pertenecen al lenguaje:
  - w6 = 1
  - w7 = 0
  - w8 = 1111110000
  - w9 = 0000000000
  - w10 = 101010101010110101010

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

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

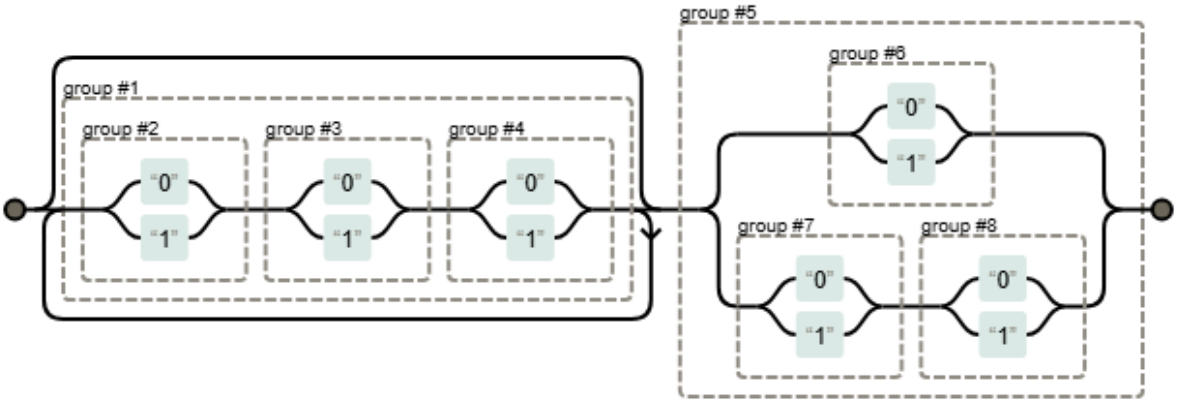


- Cadenas que pertenecen al lenguaje:
  - w1 = 000
  - w2 = 111000
  - w3 = 101010101
  - w4 = 001100111000
  - w5 = 110

- Cadenas que no pertenecen al lenguaje:
  - w6 = 1
  - w7 = 11
  - w8 = 1111
  - w9 = 0110011
  - w10 = 0111

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))

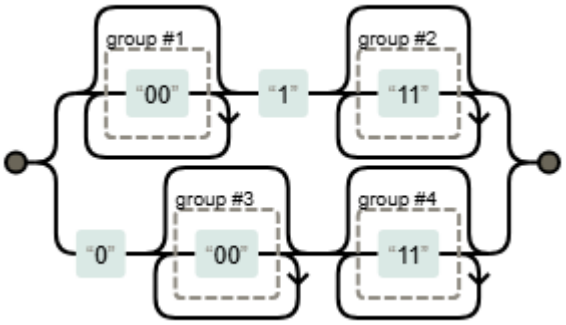


- Cadenas que pertenecen al lenguaje:
  - w1 = 0
  - w2 = 1
  - w3 = 11
  - w4 = 0011
  - w5 = 0000110

- Cadenas que no pertenecen al lenguaje:
  - w6 = 000
  - w7 = 101010
  - w8 = 111111111
  - w9 = 111
  - w10 = 101

1.8. Cadenas w sobre el alfabeto {0, 1} tal que w = 0^n1^m con n + m impar.

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

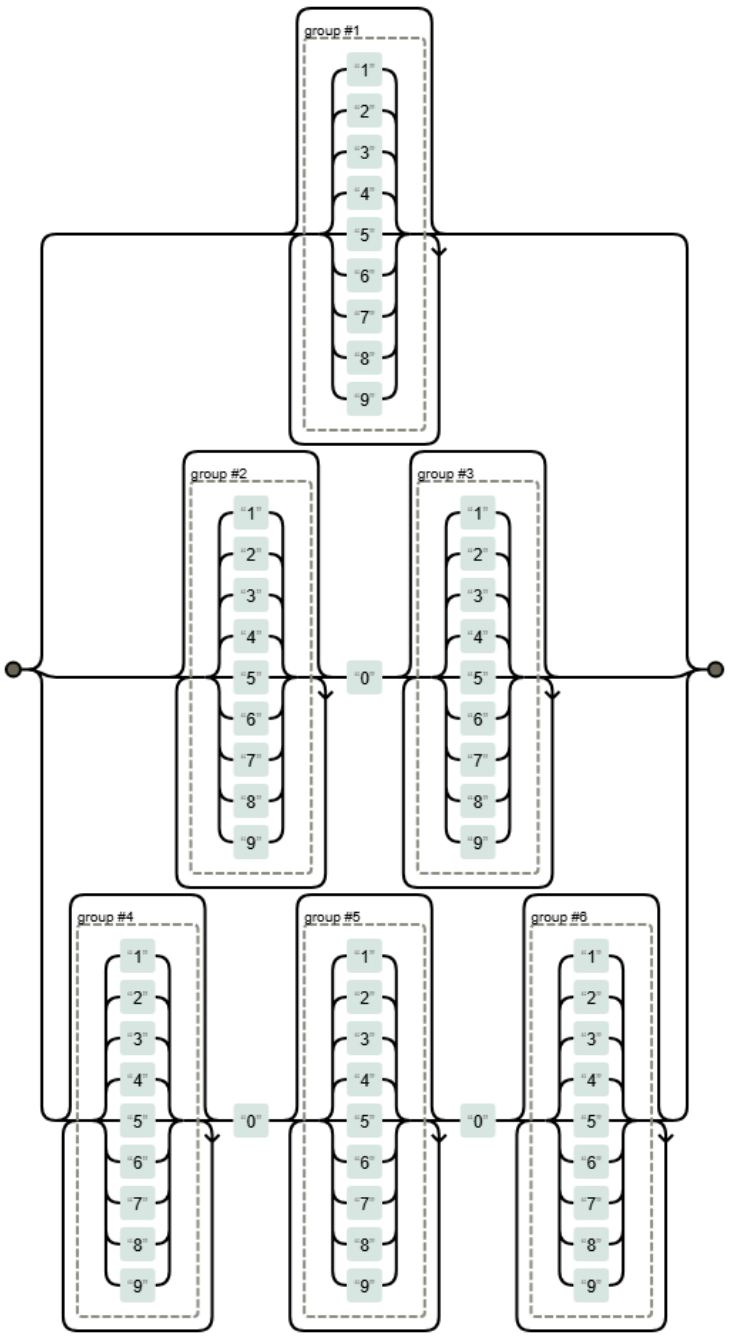


- Cadenas que pertenecen al lenguaje:
  - w1 = 0
  - w2 = 1
  - w3 = 0001111
  - w4 = 00111
  - w5 = 011

- Cadenas que no pertenecen al lenguaje:
  - w6 = 0011
  - w7 = 0111
  - w8 = 0101
  - w9 = 1111
  - w10 = 001111

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)*

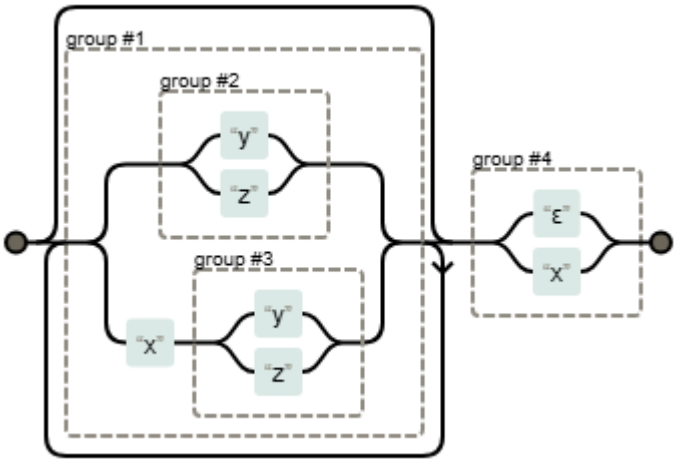


- Cadenas que pertenecen al lenguaje:
    - w1 = 12345
    - w2 = 54216023
    - w3 = 00
    - w4 = 10103234
    - w5 = 4443

- Cadenas que no pertenecen al lenguaje:
    - w6 = 000
    - w7 = 800709
    - w8 = 02220220
    - w9 = 01001
    - w10 = 003333300

1.10. Cadenas sobre el alfabeto {x, y, z} que no contengan dos símbolos x consecutivos.

- Expresión regular: ((y|z)|x(y|z))* (ε | x)



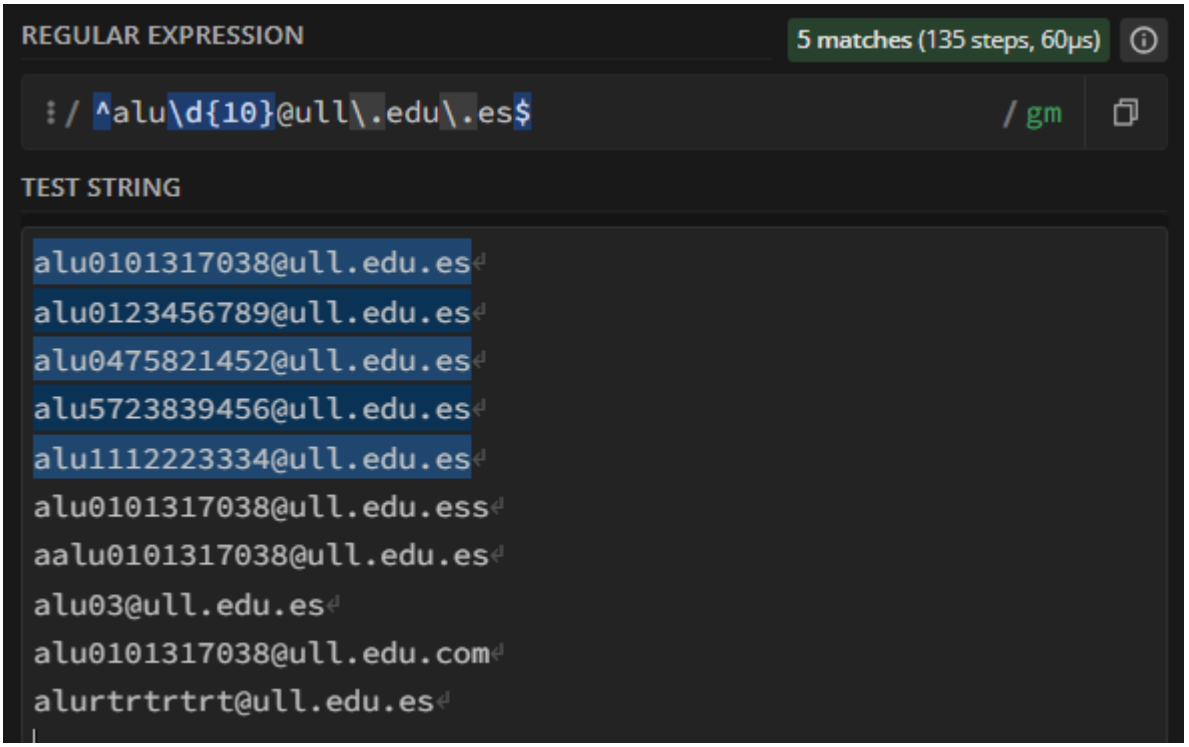
- Cadenas que pertenecen al lenguaje:
  - w1 = yxz
  - w2 = zxyxz
  - w3 = zzzxzzz
  - w4 = yxzx
  - w5 = zxz

- Cadenas que no pertenecen al lenguaje:
  - w6 = xx
  - w7 = zyxxy
  - w8 = zyyzxx
  - w9 = xxyzx
  - w10 = zxxz

2. Ejercicios sobre operadores extendidos

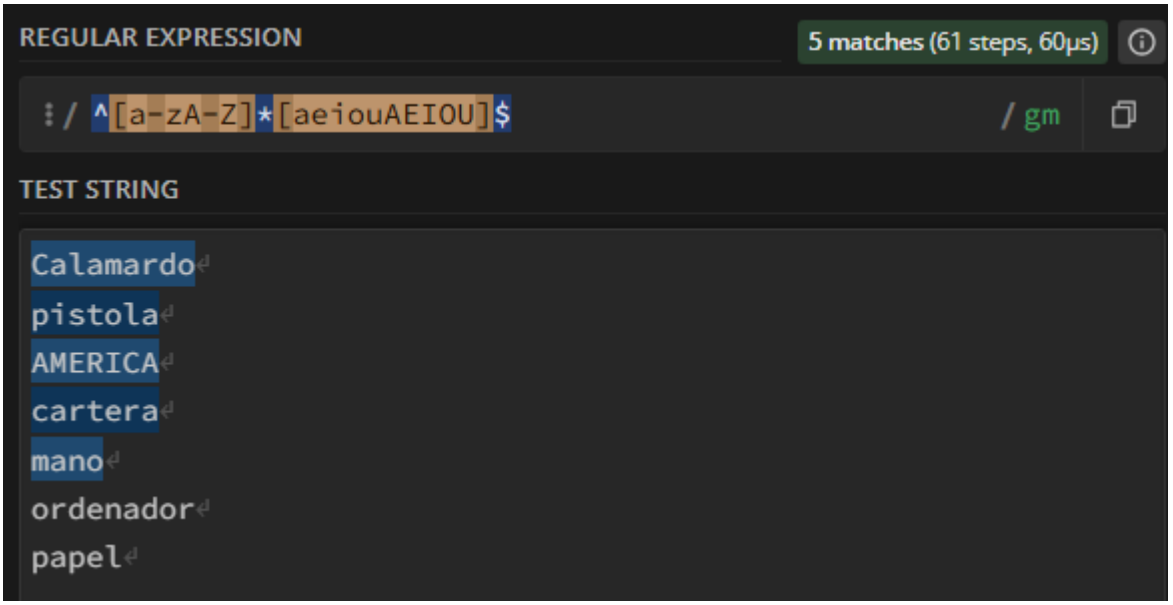
2.1. Direcciones de correos electrónicos de estudiantes de la Universidad de La Laguna.

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



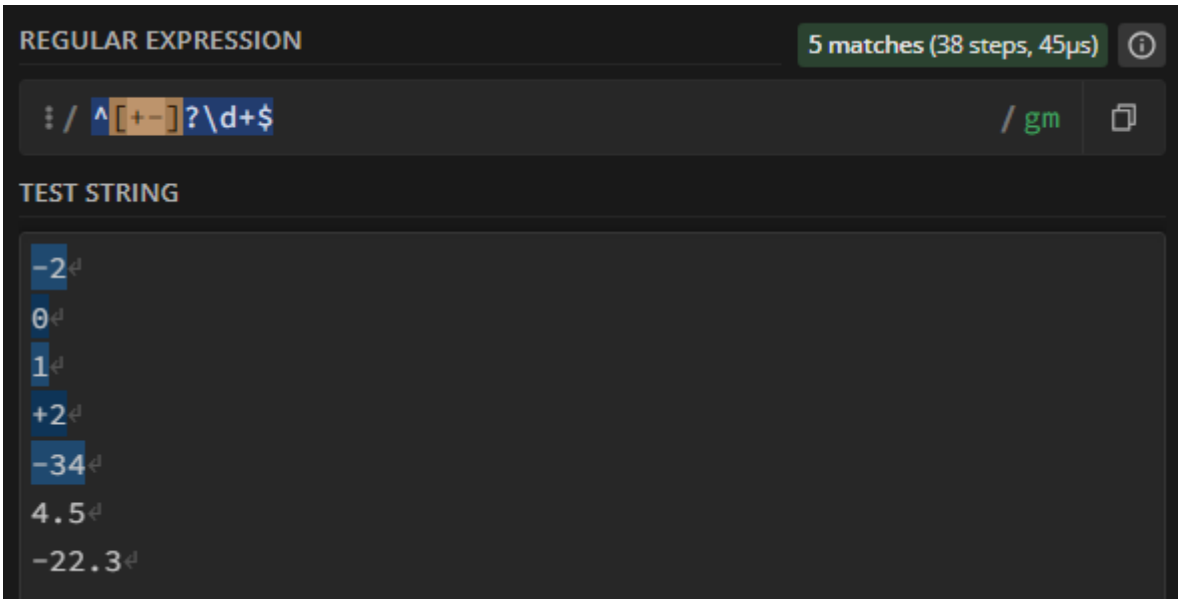
2.2. Palabras que terminen por una vocal.

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



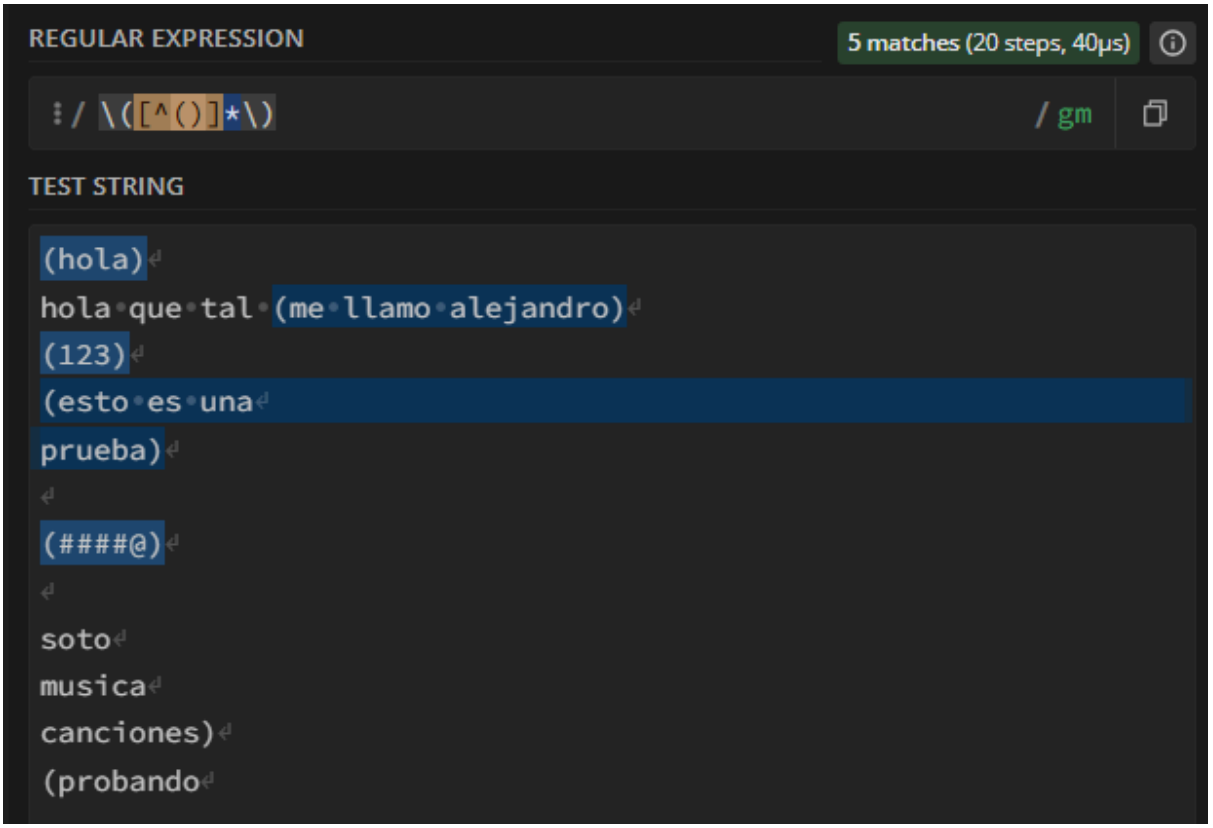
2.3. Números enteros.

- Expresión regular: ^[+-]?\d+$



2.4. Texto que se encuentre entre paréntesis.

- Expresión regular: \(([^()]*)\)

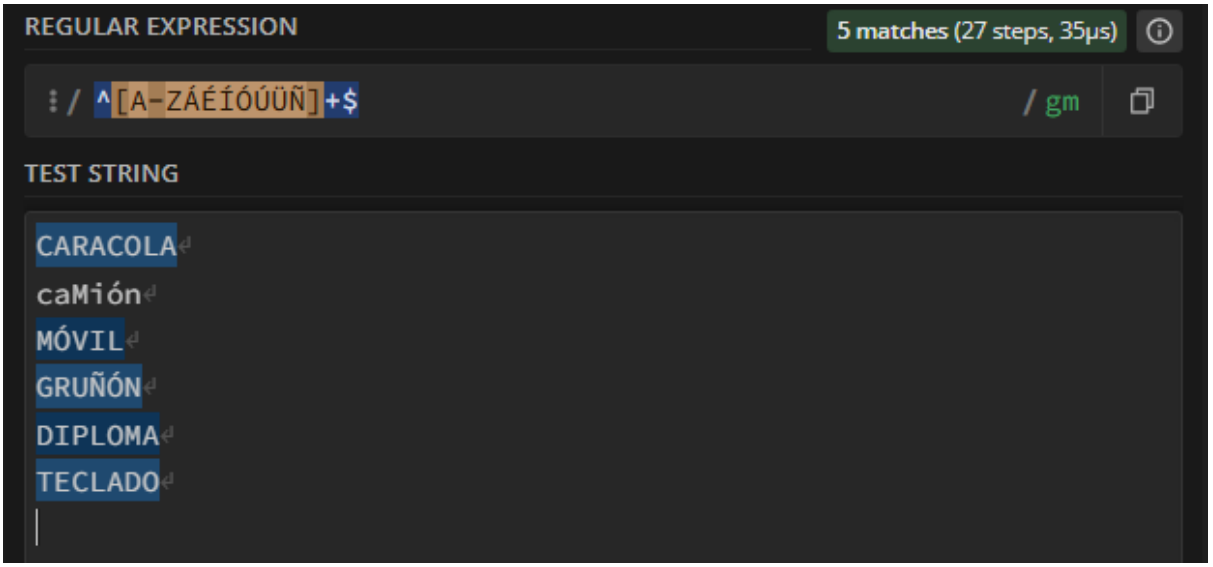


2.5. Código postal en España.

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

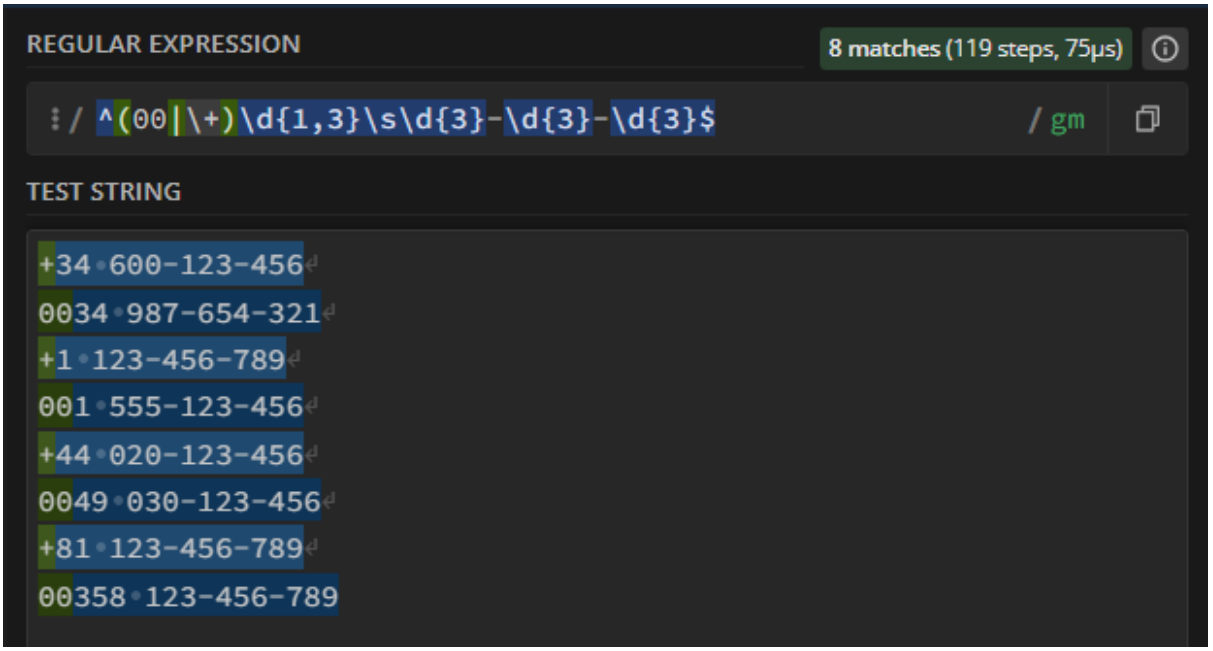
2.6. Palabras que contienen sólo letras mayúsculas.

- Expresión regular: ^[A-ZÁÉÍÓÚÜÑ]+$



2.7. Números 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}\s\d{3}-\d{3}-\d{3}$



2.8. Fecha en formato DD/MM/AAAA

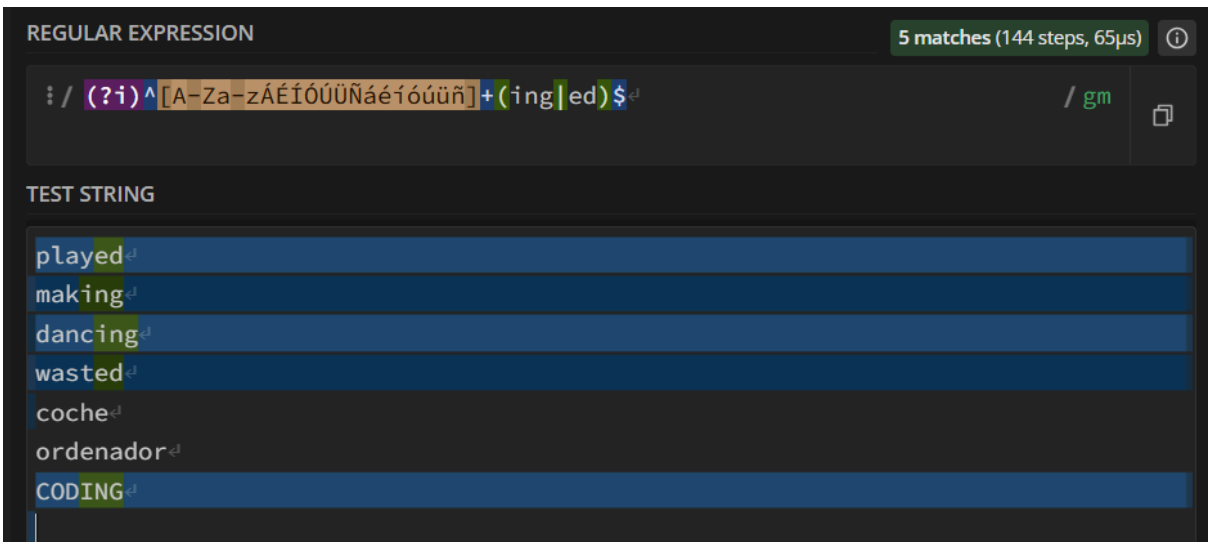
- Expresión regular: ^(0[1-9]|[12][0-9]|3[01])\/(0[1-9]|1[0-2])\/[0-9]{4}$

2.9. Palabras de al menos 10 letras de longitud.

- Expresión regular: ^[A-Za-zÁÉÍÓÚÜÑáéíóúüñ]{10,}$

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

- Expresión regular: (?i)^[A-Za-zÁÉÍÓÚÜÑáéíóúüñ]+(ing|ed)$

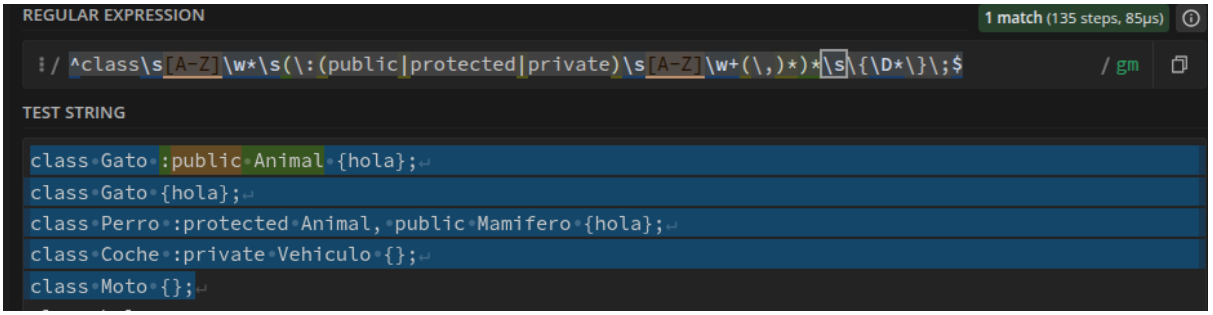


3. Modificación

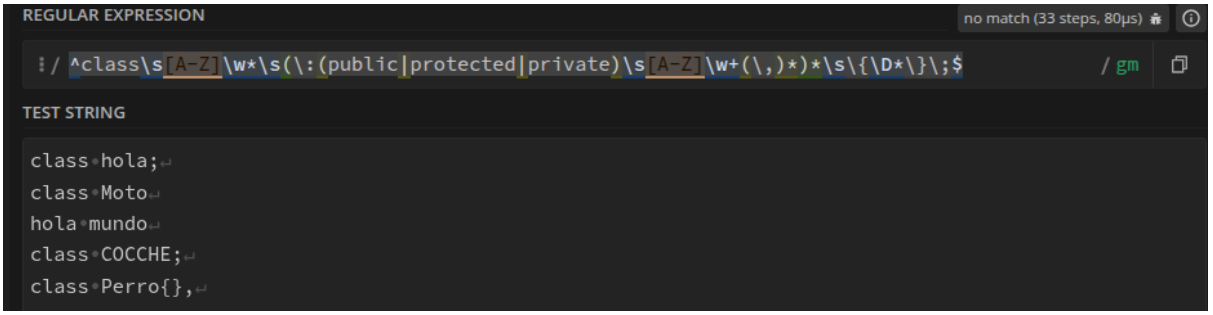
3.1. Sentencias class en C++

- Expresión regular:
  ^class\s[A-Z]\w*\s(\:(public|protected|private)\s[A-Z]\w+(\,)*)*\s\{\D*\}\;$

- Cadenas que si acepta:



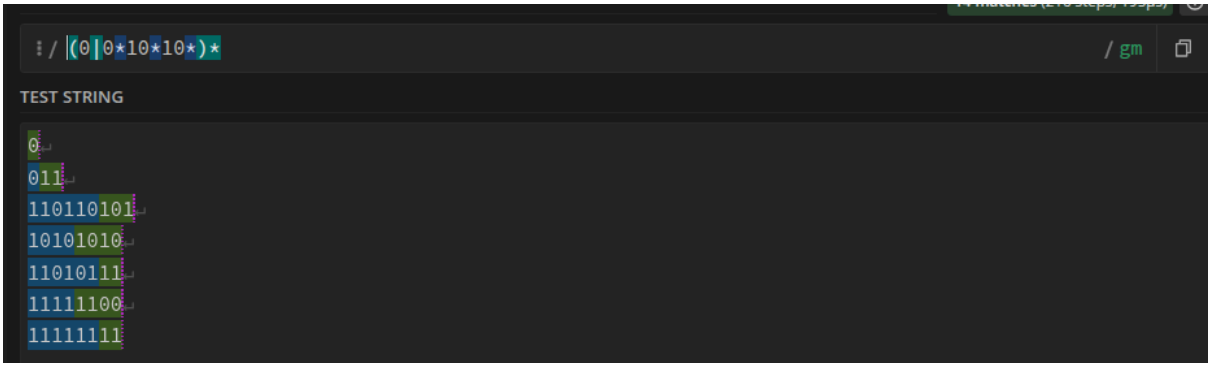
- Cadenas que no acepta:



3.2. Lenguaje binario con número par de 1's

- Expresión regular: (0|0*10*10*)*

- Cadenas que si acepta:



- Cadenas que no acepta: