

Introdução às Expressões Regulares



README.md

- \$(whoami)
- Motivação
- O que é?
- Pra que serve?
- História
- Metacaracteres
- Referências
- Onde praticar?



\$(whoami)

- Rafael Carreiro
- 22 anos
- Apaixonado por Linux

 carreirorco

 carreirorco

 carreirorco



Motivação

Dar o pontapé inicial e começar a utilizar Regex para resolver problemas do dia a dia.



O que é

“Um método formal de se especificar um texto utilizando padrões.”



O que é

“Um método formal de se especificar um texto utilizando padrões.”

[Rr]a[fp]h?ael



O que é

“Um método formal de se especificar um texto utilizando padrões.”

[Rr]a**[fp]**h?ael **[0-9][0-9]:[0-9][0-9]**



O que é

“Um método formal de se especificar um texto utilizando padrões.”

[Rr]a[fp]h?ael **[0-9][0-9]:[0-9][0-9]**

[0-9]{2}/[0-9]{2}/[0-9]{4}



Pra que serve?

“Buscar ou validar um padrão de texto que pode ser variável.”



Pra que serve?

- data
- horário
- número IP
- nome de pessoa
- endereço de e-mail
- endereço de Internet
- nome de usuário e senha
- declaração de uma função()
- dados que estão entre <tags></tags>
- número de telefone, RG, CPF, cartão de crédito
- ...



História

- Neurologistas
- Matemáticos
- Programadores (raiz)



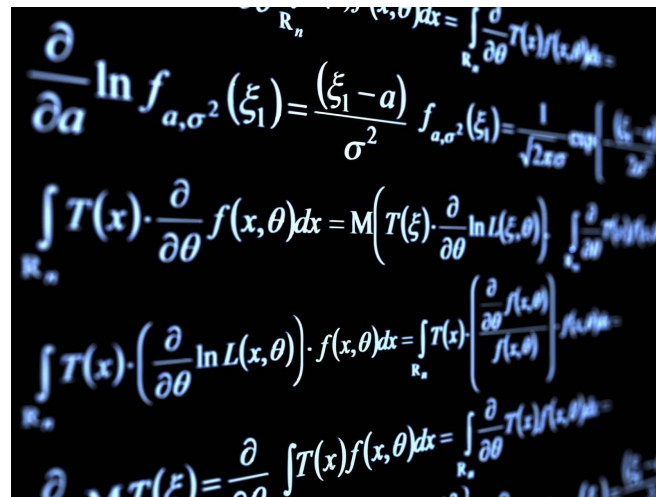
História

- **Neurologistas**
- Matemáticos
- Programadores (raiz)



História

- Neurologistas
- **Matemáticos**
- Programadores (raiz)


$$\frac{\partial}{\partial a} \ln f_{a, \sigma^2}(\xi_1) = \frac{(\xi_1 - a)}{\sigma^2} f_{a, \sigma^2}(\xi_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\xi_1 - a)^2}{2\sigma^2}\right\}$$
$$\int_{\mathbb{R}_n} T(x) \cdot \frac{\partial}{\partial \theta} f(x, \theta) dx = M\left(T(\xi) \cdot \frac{\partial}{\partial \theta} \ln L(\xi, \theta)\right) = \int_{\mathbb{R}_n} T(x) \cdot \frac{\partial}{\partial \theta} \ln L(x, \theta) \cdot f(x, \theta) dx = \int_{\mathbb{R}_n} T(x) \cdot \left(\frac{\partial}{\partial \theta} \ln L(x, \theta)\right) \cdot f(x, \theta) dx$$
$$\frac{\partial}{\partial \theta} \ln L(\xi, \theta) = \frac{\partial}{\partial \theta} \ln f(\xi, \theta) = \frac{\partial}{\partial \theta} \ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\xi - a)^2}{2\sigma^2}\right\} \right)$$

História

- Neurologistas
- Matemáticos
- **Programadores (raiz)**





PROGRAMADOR RAIZ

- Alterava código em produção
- Renomeava para `_old`
- Testava com print
- Tinha e-mail no BOL
- Subia o site com FTP linha de comando
- Usava XP SP2
- Usava VI
- Se reunia com a equipe para programar
- Usava metodologia gohorse, mas entregava



PROGRAMADOR NUTELLA

- Passa por 15 ambientes de homologação e não implanta
- Usa Git e ainda perde informações
- Gasta 90% do orçamento criando testes
- Tem e-mail no Gmail
- Sobe projeto na nuvem
- Usa MacOS
- Usa IDEs cheias de frescuras
- Tem a sua equipe toda no Slack
- Usa 3 métodos ágeis e atrasa a entrega

Metacaracteres



Metacaracteres

- Representantes . [] [^]
- Quantificadores ? * + { }
- Âncoras ^ \$ \b
- Outros \ | () \1 ... \9



Metacaracteres

- **Representantes** . [] [^]
- Quantificadores ? * + { }
- Âncoras ^ \$ \b
- Outros \ | () \1 ... \9



Metacaracteres Representantes

O ponto .

- Casa um caractere qualquer



Metacaracteres Representantes

O ponto .

- Casa um caractere qualquer

te.te  teste texte te1te te%te te-te te.te



Metacaracteres Representantes

O ponto .

- Casa um caractere qualquer

te.te  teste texte te1te te%te te-te te.te

n.o  nao não neo n1o nuo n.o



Metacaracteres Representantes

O ponto .

- Casa um caractere qualquer



te.te  teste texte te1te te%te te-te te.te

n.o  nao não neo n1o nuo n.o



Metacaracteres Representantes

A Lista []

- Somente os caracteres permitidos



Metacaracteres Representantes

A Lista []

- Somente os caracteres permitidos

n[aã]o  não nao



Metacaracteres Representantes

A Lista []

- Somente os caracteres permitidos

n[aã]o  não nao

[Nn][aã]o  não nao Nao Não



Metacaracteres Representantes

A Lista []

- Somente os caracteres permitidos

n[aã]o	→	não	nao		
[Nn][aã]o	→	não	nao	Nao	Não
[grp]ato	→	gato	rato	pato	



Metacaracteres Representantes

A Lista [] **INTERVALOS...**

- Com a lista podemos criar listas...

[0-9]  1 2 3 4 5 6 7 8 9 0

[a-z]  a b c d e f g h i j k l m n o p q r s t u v w x y z

[A-Z]  A B C D E F G H I J K L M N O P Q R S T U V W X Y Z



Metacaracteres Representantes

A Lista Negada [^]

- Caracteres proibidos



Metacaracteres Representantes

A Lista Negada [^]

- Caracteres proibidos

[^aeiou]



b c d f g h j k l m n p q r s t v w x y z



Metacaracteres Representantes

A Lista Negada [^]

- Caracteres proibidos

[^aeiou]



b c d f g h j k l m n p q r s t v w x y z

[^0-5]



6 7 8 9



Metacaracteres

- Representantes . [] [^]
- **Quantificadores** ? * + { }
- Âncoras ^ \$ \b
- Outros \ | () \1 ... \9



Metacaracteres Quantificadores

O Opcional ?

- Zero ou Um



Metacaracteres Quantificadores

O Opcional ?

- Zero ou Um

ab?c



abc

ac



Metacaracteres Quantificadores

O Opcional ?

- Zero ou Um

ab?c



abc ac

fala[!?.]?



fala fala! fala? fala.



Metacaracteres Quantificadores

O Opcional ?

- Zero ou Um

ab?c



abc ac

fala[!?.]?



fala fala! fala? fala.

[Rr]a[fp]h?ael



Rafael rafael Raphael raphael



Metacaracteres Quantificadores

O asterisco *

- Zero, um ou mais



ab*c  abc ac abbc abbbbbbbbc



Metacaracteres Quantificadores

O asterisco *

- Zero, um ou mais




ab*c		abc	ac	abbc	abbbbbbbbc
12*3		123	13	1223	122222223



Metacaracteres Quantificadores

O asterisco *

- Zero, um ou mais

ab*c		abc	ac	abbc	abbbbbbbbc
12*3		123	13	1223	122222223
Vaness*a		Vanessa		Vanesssssssssa	



Metacaracteres Quantificadores

O mais +

- Um ou mais



Metacaracteres Quantificadores

O mais +

- Um ou mais

Linux+



Linux

Linuxx

Linuxxxxxxx



Metacaracteres Quantificadores

O mais +

- Um ou mais

Linux+



Linux

Linuxx

Linuxxxxxxx

Melhor+



Melhor

Melhoor

Melhooooor



Metacaracteres Quantificadores

O mais +

- Um ou mais

Linux+



Linux

Linuxx

Linuxxxxxxx

Melhor+



Melhor

Melhoor

Melhooooor

Windows+



Windows

Windowss

Windowssss



Metacaracteres Quantificadores

O mais +

- Um ou mais

Linux+



Linux

Linuxx

Linuxxxxxxx

Melhor+



Melhor

Melhoor

Melhooooor

Windows+



Windows

Windowss

Windowsssss



Para nó+ssa alegria!





**PARA
NOOOOOOSSA
ALEGRIA!**



Metacaracteres Quantificadores

As Chaves { }

- De x até y



Metacaracteres Quantificadores

As Chaves { }

- De x até y



Metacaracteres Quantificadores

As Chaves { }

- De x até y, definimos quantas vezes...

a{1,3}  a aa aaa



Metacaracteres Quantificadores

As Chaves { }

- De x até y, definimos quantas vezes...

a{1,3}



a

aa

aaa

0{2,4}



00

000

0000



Metacaracteres

- Representantes . [] [^]
- Quantificadores ? * + { }
- **Âncoras** ^ \$ \b
- Outros \ | () \1 ... \9



Metacaracteres Âncoras

O circunflexo ^

- Início da linha



Metacaracteres Âncoras

O circunflexo ^

- Início da linha

[^][0-9]



Linhas que começam por números



Metacaracteres Âncoras

O circunflexo ^

- Início da linha

^[0-9]



Linhas que começam por números

^[a-z]



Linhas que começam por letras minúsculas



Metacaracteres Âncoras

O circunflexo ^

- Início da linha

^[0-9]



Linhas que começam por números

^[a-z]



Linhas que começam por letras minúsculas

^[A-Z]



Linhas que começam por letras maiúsculas



Metacaracteres Âncoras

O cifrão **\$**

- Final da linha



Metacaracteres Âncoras

O cifrão \$

- Final da linha

`[0-9]$`



Linhas que terminam com números



Metacaracteres Âncoras

O cifrão \$

- Final da linha

`[0-9]$`



Linhas que terminam com números

`[a-z]$`



Linhas que terminam com letras minúsculas



Metacaracteres Âncoras

O cifrão \$

- Final da linha

[0-9]\$



Linhas que terminam com números

[a-z]\$



Linhas que terminam com letras minúsculas

[A-Z]\$



Linhas que terminam com letras maiúsculas



Metacaracteres Âncoras

A borda **\b**

- Onde começa e/ou termina a palavra



Metacaracteres Âncoras

A borda **\b**

- Onde começa e/ou termina a palavra

dia



dia

diafragma

melodia

bom-dia



Metacaracteres Âncoras

A borda **\b**

- Onde começa e/ou termina a palavra

dia



dia

diafragma

melodia

bom-dia

\bdia



dia

diafragma




bom-dia



Metacaracteres Âncoras

A borda **\b**

- Onde começa e/ou termina a palavra

dia		dia	diafragma	melodia	bom-dia
\b dia		dia	diafragma		bom-dia
dia \b		dia		melodia	bom-dia



Metacaracteres

- Representantes . [] [^]
- Quantificadores ? * + { }
- Âncoras ^ \$ \b
- **Outros** \ | () \1 ... \9



Metacaracteres Outros

O escape \

- Tira o poder do metacaractere



Metacaracteres Outros

O escape \

- Tira o poder do metacaractere

\.



Tira o poder do ponto .



Metacaracteres Outros

O escape \

- Tira o poder do metacaractere

\.



Tira o poder do ponto .

\(\\)






Tira o poder dos parenteses ()



Metacaracteres Outros

O escape \

- Tira o poder do metacaractere

<code>\.</code>		Tira o poder do ponto .
<code>\(\\)</code>		Tira o poder dos parenteses ()
<code>\\</code>		Tira o poder dele mesmo



Metacaracteres Outros

O pipe |

- Alternativo, também conhecido como **OU**



Metacaracteres Outros

O pipe |

- Alternativo, também conhecido como **OU**

um|outro  **Ou** um **ou** outro



Metacaracteres Outros

O pipe |

- Alternativo, também conhecido como **OU**

um|outro  **Ou** um **ou** outro

[gpr]ato  gato | pato | rato



Metacaracteres Outros

O grupo ()



Metacaracteres Outros

O grupo ()

(ha)+



ha

haha

hahaha



Metacaracteres Outros

O grupo ()

(ha)+



ha

haha

hahaha

boa-(tarde|noite)



boa-tarde

boa-noite



Metacaracteres Outros

O grupo ()

(ha)+	→	ha	haha	hahaha
boa-(tarde noite)	→	boa-tarde	boa-noite	
(super hiper)mercado	→	supermercado	hipermercado	mercado



Metacaracteres Outros

Os retrovisores **\1** , **\2**, ... , **\9**



Metacaracteres Outros

Os retrovisores **\1** , **\2**, ... , **\9**

(quero)-\1



quero-quero



Metacaracteres Outros

Os retrovisores \1 , \2, ... , \9

(quero)-\1



quero-quero

(mega)(mente) \1 \2



mega mente



Metacaracteres Outros

Os retrovisores \1 , \2, ... , \9

(quero)-\1



quero-quero

(mega)(mente) \1 \2



mega mente

(mega)(mente) \2 \1



mega mente



Referências

<http://aurelio.net/regex/guia/>

<http://piazinho.com.br/>

<http://unicobit.com.br/programacao/expressoes-regulares-parte-01/>

<https://tableless.com.br/o-basico-sobre-expressoes-regulares/>



Onde praticar?

<http://www.regexpal.com/>

<https://regex101.com/>

<http://www.gethifi.com/tools/regex>

<http://piazinho.com.br/ed5/exemplos.html#1>



\$
