



INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
RIO GRANDE DO NORTE

Introdução a Sistemas Abertos

Expressões Regulares

Curingas

- Curingas (ou referência global) é um recurso usado para especificar um ou mais arquivos ou diretórios do sistema de uma só vez. Este é um recurso permite que você faça a filtragem do que será listado, copiado, apagado, etc.
 - ▶ ***** - Faz referência a um nome completo/restante de um arquivo/diretório.
 - *ato = pato, rato, grato, Prato, retrato, ...
 - ▶ **?** - Faz referência a uma letra naquela posição.
 - ?ato = pato, rato, gato, ...

Curingas

- Lista: [...]
 - ▶ Mais específica que o ponto, a lista determina quais caracteres ou símbolos podem ser casados
 - ▶ Só pode ser casado um caracter por vez dentro de cada lista
 - [pgr]ato = prato, gato ou rato
 - n[aã]o = nao ou não
 - [a-z][0-9] - Faz referência a caracteres de a até z seguido de um caracter de 0 até 9.
 - [a,z][1,0] - Faz a referência aos caracteres a ou z seguido de um caracter 1 ou 0 naquela posição.
 - [a-z,1,0] - Faz referência a intervalo de caracteres de a até z ou 1 ou 0 naquela posição.

Expressões Regulares

- Uma expressão regular (**Regex**), é uma construção que utiliza pequenas ferramentas, feita para obter uma determinada sequência de caracteres de um texto.
- Vários programas e linguagens de programação tem suporte à **Regex**, dentre eles o **grep**, **egrep**, **find**, **sed**, **awk**, **lex**, **perl**, e também alguns editores como o **vim** e **emacs**.

Metacaracteres

- Metacaracteres são caracteres com funções específicas, que informam padrões e posições impossíveis de serem especificadas com caracteres normais.
- Circunflexo ^ - Representa o início de uma linha

```
$ grep ^root /etc/passwd  
root:x:0:0:root:/root:/bin/bash
```

- Cifrão \$ - Representa o fim de uma linha.

```
$ grep bash$ /etc/passwd  
root:x:0:0:root:/root:/bin/bash  
mysql:x:27:27:MySQL Server:/var/lib/mysql:/bin/bash  
carlos:x:500:500:carlos:/home/carlos:/bin/bash  
ana:x:501:501:Ana Paula Moreira:/home/ana:/bin/bash  
acs:x:502:502:Antonio Carlos Silva:/home/acs:/bin/bash
```

Metacaracteres

- Lista [] - Lista de Caracteres que podem ser casados.

```
$ grep '[Cc]arlos' /etc/passwd  
carlos:x:500:500:carlos:/home/carlos:/bin/bash  
acs:x:502:502:Antonio Carlos Silva:/home/acs:/bin/bash
```
- Lista negada [^] - Qualquer caractere exceto os da lista.

```
$ grep '^[^aeiou]' /etc/passwd
```
- Ponto . - Qualquer caractere na posição.

```
$ grep '^.....$' /etc/passwd  
news:x:9:13:news:/etc/news:
```

Metacaracteres

- Chaves { } - Indica a quantidade de repetições de um caractere.

a{2} chaves Casa a letra 'a' duas vezes

a{2,4} chaves Casa a letra 'a' de duas a quatro vezes

a{2,} chaves Casa a letra 'a' no mínimo duas vezes

```
$ egrep '^.{27}$' passwd  
news:x:9:13:news:/etc/news:
```

Metacaracteres

- Outros metacaracteres que podem ser usados são o asterisco, o mais e a interrogação (* + ?). Eles definem quantidades e funcionam como as chaves, porém com uma sintaxe mais prática:

Meta	Nome	Equivalente	Descrição
?	opcional	{0,1}	Pode aparecer ou não (opcional)
*	asterisco	{0,}	Pode aparecer em qualquer quantidade
+	mais	{1,}	Deve aparecer no mínimo uma vez

Metacaracteres

- Ou | - Para fazer o OR lógico, onde se procura por uma coisa ou outra, deve-se usar o pipe “|” e delimitar as opções com os parênteses “()”:

```
$ egrep '^ (ana|carlos|acs):' /etc/passwd  
carlos:x:500:500:carlos:/home/carlos:/bin/bash  
ana:x:501:501:Ana Paula Moreira:/home/ana:/bin/bash  
acs:x:502:502:Antonio Carlos Silva:/home/acs:/bin/bash
```

Exercícios

- Criar expressões regulares para encontrar:
 - data
 - horário
 - número IP
 - endereço de e-mail
 - dados que estão entre `<tags></tags>`
 - número de telefone, RG, CPF, cartão de crédito