



**UNIVERSIDADE FEDERAL DE RORAIMA  
CENTRO DE CIÊNCIA E TECNOLOGIA  
CURSO DE CIÊNCIA DA COMPUTAÇÃO  
LINGUAGENS DE PROGRAMAÇÃO**

**Acadêmico:** Antonio de Oliveira Viana

**REGEX – REGULAR EXPRESSION**

Uma Expressão Regular é um método formal de se especificar um padrão de texto. É uma composição de símbolos, caracteres com funções especiais, chamados “metacaracteres” que, agrupados entre si e com caracteres literais, formam uma sequência, uma expressão. Expressões Regulares são notações para identificar padrões em textos. Essa expressão é testada em textos e retorna sucesso caso esse texto obedeça exatamente a todas as suas condições. Diz-se que o texto “casou” com a expressão. Definem um conjunto de um ou mais strings. Uma string simples é uma expressão regular que define uma string: ela mesma. Uma expressão regular complexa usa letras, números e caracteres especiais para definir muitas diferentes strings. Uma expressão regular casa com qualquer string que ela define.

A Expressão Regular serve para se dizer algo abrangente de forma específica. Definido o padrão de busca, tem-se uma lista (finita ou não) de possibilidades de casamento. Em um exemplo rápido, [rgp] ato pode casar “rato”, “gato” e “pato”.

Vários editores de texto e linguagens de programação têm suporte a Expressão Regular, então o tempo investido em seu aprendizado é recompensado pela larga variedade de aplicativos onde ele pode ser praticado. São suportadas por muitos programas em linha de comando e por muitas linguagens de programação para facilitar a resolução de problemas de manipulação de textos. Mas nem todas expressões regulares funcionam da mesma forma, podem variar de ferramentas e linguagens de programação.

Um dos principais programas para o uso das Expressões Regulares é o grep. Em essência, esse programa procura em arquivos de texto pela ocorrência de expressões regulares especificadas e imprime qualquer linha que casa na saída padrão. Grep possui a seguinte sintaxe grep [options] regex [file...], onde o regex é uma expressão regular.

Um caractere é considerado qualquer um que não seja nova linha (\n). Caracteres literais: caracteres que se representam (casam eles próprios). Caracteres especiais (metacaracteres): caracteres que são usados para representar casamentos mais complexos; são eles: expressão regular básica (^ \$ . [ ] \* \) e expressão regular estendida (? + ( ) { } |). Se precisar usar um caractere especial para representar a si próprio, deve-se escapá-lo. Um caractere chamado delimitador usualmente marca o início e o fim de uma expressão regular. O delimitador é sempre um caractere especial para a expressão regular que ele delimita (não representa a si próprio).