

## Exercícios – Objetivo 103.7

1. Crie uma expressão regular (ER) para cada um dos padrões definidos abaixo. Você pode usar `grep` e o arquivo `teste.txt` para testar suas ERs.
  - (a) Linhas que contêm pata ou pato.
  - (b) Linhas que contêm a palavra pato.
  - (c) Linhas que contêm palavras começadas em pat e terminadas em o ou os.
  - (d) Linhas que terminam em eco.
  - (e) Linhas que começam com Pin.
  - (f) Linhas que começam com Pin e terminam em im ou ima.
  - (g) Linhas que terminam em um ponto de interrogação (?).
  - (h) Linhas que terminam em um ponto de interrogação (?) ou de exclamação (!).
  - (i) Linhas que contêm uma ou mais vírgulas.
  - (j) Linhas que contêm mais de uma vírgula, com palavras entre elas.
  - (k) Linhas que contêm um ou mais asteriscos (\*).
  - (l) Linhas que contêm apenas um asterisco.
  - (m) Linhas que contêm uma ou mais ocorrências de dois pontos consecutivos (.).
  - (n) Linhas que contêm operadores de autoincremento (++) E/OU autodecremento (--).
  - (o) Linhas que contêm operadores de autoincremento E autodecremento.
2. Em um sistema Linux, o diretório `/usr/share/dict` contém um ou mais dicionários usados por corretores ortográficos. O arquivo `words` contém o dicionário padrão, e outros arquivos (como `french` e `spanish`) contêm dicionários para diferentes idiomas. Letreco<sup>1</sup>, uma versão em português do Wordle<sup>2</sup>, é um jogo de adivinhação de palavras de cinco letras. O jogador tem seis tentativas para adivinhar uma palavra. Para cada tentativa, o jogo indica quais letras existem na palavra e estão na posição correta (em verde), quais letras existem na palavra mas estão na posição incorreta (em amarelo), e quais letras não fazem parte da palavra (em vermelho). O jogo aceita apenas letras no conjunto a-z: cedilhas e acentos são ignorados (“c” é usado no lugar de “ç”, e “a” é usado no lugar de “á”). Use o dicionário do sistema e ferramentas como `grep`, `sed` e `tr` para aumentar suas chances de adivinhar a palavra correta em um desses jogos. **DICA:** primeiro gere um arquivo contendo apenas palavras que podem aparecer no jogo, e depois vá criando, de forma iterativa, expressões regulares para selecionar nesse arquivo as palavras que casam com as dicas fornecidas pelo jogo. Você também pode usar um pipeline com múltiplos greps para dividir as tarefas de inclusão e exclusão de palavras. Por exemplo,

```
$ cat letreco.txt | grep '^s' | grep a | grep -v '[ob]'
```

seleciona palavras que começam por s, contêm a, e não contêm b ou o.

---

<sup>1</sup><https://www.gabtoschi.com/letreco/>

<sup>2</sup><https://www.nytimes.com/games/wordle/index.html>

3. Escreva uma expressão regular que case com expressões aritméticas simples na forma

*número op número = número*

Um *número* pode ter uma quantidade arbitrária de dígitos. Os operadores possíveis são os quatro básicos (+, -, \*, /). Considere que não haja espaços entre os termos. Um exemplo de expressão válida seria 12+45=37. Teste sua ER com `grep` e o arquivo `expr-arit-1.txt`.

4. Modifique a ER do exercício 3 para que números e operadores possam (ou não) ser separados por espaços ou Tab. Teste sua ER com `grep` e o arquivo `expr-arit-2.txt`.
5. Use `grep` e/ou `sed` para remover as *tags* HTML de um arquivo e exibi-lo sem linhas em branco (vazias ou contendo apenas espaços/Tabs).
6. O arquivo `agenda.txt` contém uma lista simples de números de telefone. Infelizmente, ela contém números de celular (que iniciam por 8 ou 9) em dois formatos, com 8 e 9 dígitos. Use `sed` para acrescentar o dígito 9 na frente dos números de celular que ainda estão no formato de 8 dígitos.