

Licenciatura em Engenharia Informática – DEI/ISEP  
**Linguagens de Programação 2012/2013**  
Aula Prática-Laboratorial

## Ficha PL 2

### Expressões Regulares

#### Objectivos:

- Introdução ao reconhecimento de padrões e expressões regulares;
- Exercícios com expressões regulares com vista à consolidação dos conhecimentos introduzidos na Ficha TP 1;
- Implementação de analisadores léxicos em FLEX

## 2. EXPRESSÕES REGULARES

As expressões regulares (ER) constituem um método formal para especificar um padrão de texto. As expressões regulares são regras bem definidas que permitem gerar um conjunto de sequências de símbolos (palavras). Traduz um conjunto de padrões, possivelmente complicados e difíceis pela sua dimensão de enumerar, numa expressão de dimensão curta e relativamente fácil de interpretar. Uma ER constrói-se com base num conjunto de meta-caracteres ( \*, [], ^, ., -, ?, entre outros) que são interpretados de forma especial, escritos numa dada sequência e/ou intercalados com sequências de caracteres sem significado especial que não seja a sua verificação.

### 2.1 Exercícios propostos

1. Descreva informalmente (por palavras) as linguagens representadas pelas seguintes expressões:
  - a)  $0(0|1)^*0$
  - b)  $(01)^*$
  - c)  $(0|1)^*0(0|1)(0|1)$
  - d)  $0^*10^*10^*10^*$
  - e)  $(0|1)^*111(0|1)^*$
2. Considerando o alfabeto  $\Sigma = \{A, B, C, \dots, Z, a, b, c, \dots, z\}$ , represente as seguintes linguagens utilizando expressões regulares:
  - a)  $L(A) = \{u \in \Sigma^* : u \text{ tem pelo menos dois } a\text{'s juntos}\}$
  - b)  $L(A) = \{u \in \Sigma^* : u \text{ tem comprimento inferior a } 4\}$
  - c)  $L(A) = \{u \in \Sigma^* : u \text{ começa por uma maiúscula}\}$
  - d)  $L(A) = \{u \in \Sigma^* : u \text{ tem um número par de maiúsculas, um número ímpar de minúsculas e tem comprimento múltiplo de } 4\}$

3. Diga se são verdadeiras ou falsas as seguintes afirmações:
  - a)  $(0|1)^* \Leftrightarrow (0^*|1^*)$
  - b)  $(0|1)^* \Leftrightarrow (0^*1^*)^*$
4. Considerando o alfabeto  $\Sigma = \{0,1\}$ , represente as seguintes linguagens utilizando expressões regulares:
  - a)  $L(A) = \{u \in \Sigma^* : u \text{ é um número binário múltiplo de } 4\}$
  - b)  $L(A) = \{u \in \Sigma^* : 111 \text{ não é factor de } u\}$
  - c)  $L(A) = \{u \in \Sigma^* : u \text{ tem um número par de } 1\text{'s}\}$
  - d)  $L(A) = \{u \in \Sigma^* : u \text{ é vazia ou tem dígitos todos iguais, sendo de comprimento par as sequências de } 0\text{'s, e ímpar as sequências de } 1\text{'s}\}$
  - e)  $L(A) = \{u \in \Sigma^* : u \text{ é múltiplo de } 4 \text{ mas não de } 8\}$
  - f)  $L(A) = \{u \in \Sigma^* : u \text{ tem comprimento } 3\}$
5. Para a classificação dos alunos de uma escola, cada nota encontra-se codificada através do par: NT, em que N é um valor inteiro entre 0 e 20 e T pode ser a letra F, ou E ou o par de letras RF. Uma nota não pode conter um zero na casa das dezenas. Uma sequência de notas, terá de conter pelo menos uma nota. Escrever uma expressão regular que valide uma sequência ilimitada de notas.
6. Considerando o alfabeto  $\Sigma = \{0,1,2,3,4,5,6,7,8,9\}$ , represente as seguintes linguagens utilizando expressões regulares:
  - a)  $L(A) = \{u \in \Sigma^* : u \text{ é múltiplo de } 5\}$
  - b)  $L(A) = \{u \in \Sigma^* : u \text{ é um número ímpar}\}$
  - c)  $L(A) = \{u \in \Sigma^* : u \text{ é um inteiro maior ou igual a } 1000\}$
7. Escreva uma expressão regular para as seguintes linguagens:
  - a) String no alfabeto  $\{a, b, c, d\}$  em que o primeiro "b" (se existir) é precedido de um "a";
  - b) Todas as strings no alfabeto  $\{a, b, c, \dots, z\}$  que contêm as cinco vogais em sequência e por ordem alfabética pelo menos 1 vez.  
**xaeiouz** - válido  
**xaaeiouz** - válido  
**xaeiioz** - inválido  
**xayeioz** - inválido  
**xayaeiouwez** - válido  
**xaeiouyaeiouz** - válido
8. As quantias monetárias são expressas de diferentes formas. Escreva as expressões regulares que permitam realizar o reconhecimento das moedas apresentadas na tabela 2.1 e um programa em FLEX, que implemente as expressões regulares definidas.

**Tabela 2.1:** Representação de moedas

Moeda	Exemplo
Euro	e12,23; e1,00; e2,35; 23,50EUR
Libra	£12.50; £22.12; £22.99
Dólar	\$25.13; \$5.00; \$0.30
Escudo	12\$50; 25\$00; 150\$00; 0\$50

## 2.2 Exercícios complementares

- Considerando o alfabeto  $\Sigma = \{0,1\}$ , represente as seguintes linguagens utilizando expressões regulares:
  - $L(A) = \{u \in \Sigma^* : 10 \text{ é prefixo de } u\}$
  - $L(A) = \{u \in \Sigma^* : u \text{ é um número binário que ou não é par ou é múltiplo de } 8\}$
  - $L(A) = \{u \in \Sigma^* : u \text{ é múltiplo de } 8 \text{ mas não de } 4\}$
  - $L(A) = \{u \in \Sigma^* : u \text{ tem no máximo comprimento } 3 \text{ e é uma capicua}\}$
- Considerando o alfabeto  $\Sigma = \{0,1,2,3,4,5,6,7,8,9\}$ , represente as seguintes linguagens utilizando expressões regulares:
  - $L(A) = \{u \in \Sigma^* : u \text{ é uma sequência numérica crescente}\}$
  - $L(A) = \{u \in \Sigma^* : u \text{ começa e termina no mesmo algarismo}\}$
- Considere o extracto seguinte de um ficheiro que contém informação georreferenciada sobre um conjunto de cidades.

```
# Cidades
41°9'N      8°38'W      11-02-2007    9:10:34      Porto
38°42'N      9°11'W      11-02-2007    9:10:42      Lisboa
51°30'25"N   0°07'39"W   11-02-2007    9:10:43      Londres
22°54'30"S   43°11'47"W  11-02-2007    9:10:43      Rio
55°45'8"N    37°37'56"E  11-02-2007    9:10:45      Moscovo
#Trip description
```

Escreva as expressões que permitem a identificação dos campos apresentados na tabela 2.2. Considere também a necessidade de reconhecer as linhas de comentário (iniciadas com o carácter #).

**Tabela 2.2:** Campo de informação geo-referenciada

Campo	Exemplo
Latitude	41°9'N
	22°54'30"S
Longitude	9°11'W
	37°37'56"E
Data	11-02-2007
Hora	9:10:34
Cidade	Porto

- Escreva um programa em FLEX, que implemente as expressões regulares definidas na alínea anterior.