

Processamento de Linguagens (3º Ano MIEI)

Trabalho Prático 1a - GAWK

Relatório de Desenvolvimento

Fábio Luís Baião da Silva
(A75662)

João da Cunha Coelho
(A74859)

Luís Miguel Moreira Fernandes
(A74748)

13 de Março de 2017

Resumo

Este projeto visa provar que as *Expressões Regulares* podem ser muito úteis, mostrando que é possível utilizá-las na rotina do dia-a-dia.

A Via Verde é uma coleção de sistemas eletrónicos de portagens usado em Portugal desde 1991. O suporte para este projeto é um extrato mensal de um dos seus utilizadores, que irá ser usado para responder a umas *queries* usando **GAWK**, um dos processadores de texto mais proeminentes no GNU/Linux.

Conteúdo

1	Introdução	2
2	Estrutura do Ficheiro XML	3
3	Apresentação da Solução de Leitura e Parser	4
3.1	Row Separator (RS)	4
3.2	Field Separator (FS)	4
4	Apresentação das Queries e Resoluções	5
5	Conclusão	6

Capítulo 1

Introdução

Introdução do Trabalho a.k.a. Coisas para o Coelho apaneleirar.

Capítulo 2

Estrutura do Ficheiro XML

De seguida é sucintamente explicada a estrutura do ficheiro XML em análise.

Cabeçalho com informação sobre o mês a que se refere o extrato e os dados do cliente:

```
<EXTRACTO id="011114056/08/2015">
<MES_EMISSAO>Ago-2015</MES_EMISSAO>
<CLIENTE id="514714936">
  <NIF>987653210</NIF>
  <NOME>PEDRO MANUEL RANGEL SANTOS HENRIQUES</NOME>
  <MORADA>RUA XXX</MORADA>
  <LOCALIDADE>BRAGA</LOCALIDADE>
  <CODIGO_POSTAL>4715-012 BRAGA</CODIGO_POSTAL>
</CLIENTE>
<IDENTIFICADOR id="28876820811">
<MATRICULA>00-LJ-11</MATRICULA>
<REF_PAGAMENTO>1234567</REF_PAGAMENTO>
```

Todas as transações de serviços oferecidos pela ViaVerde (portagens, parques de estacionamento, etc), sendo que cada transação tem campos como: data e hora de entrada e saída, local de entrada e saída, valor, etc:

```
<TRANSACCAO>
  <DATA_ENTRADA>26-07-2015</DATA_ENTRADA>
  <HORA_ENTRADA>11:33</HORA_ENTRADA>
  <ENTRADA>Povoa N-S</ENTRADA>
  <DATA_SAIDA>26-07-2015</DATA_SAIDA>
  <HORA_SAIDA>11:42</HORA_SAIDA>
  <SAIDA>Angeiras N-S</SAIDA>
  <IMPORTANCIA>2,00</IMPORTANCIA>
  <VALOR_DESCONTO>0,00</VALOR_DESCONTO>
  <TAXA_IVA>23</TAXA_IVA>
  <OPERADOR>I. de Portugal (N1)</OPERADOR>
  <TIPO>Portagens</TIPO>
  <DATA_DEBITO>05-08-2015</DATA_DEBITO>
  <CARTAO>6749036</CARTAO>
</TRANSACCAO>
```

Rodapé com valores totais:

```
...
<TOTAL>77,40</TOTAL>
</IDENTIFICADOR>
<TOTAL>77,40</TOTAL>
<TOTAL_IVA>14,49</TOTAL_IVA>
</EXTRACTO>
```

Capítulo 3

Apresentação da Solução de Leitura e Parser

```
RS = "<TRANSACCAO>";  
FS = "[<>]";
```

3.1 Row Separator (RS)

A opção passou por separar os *records* pela tag <TRANSACAO>.

Com o *Record Separator* definido desta forma sabe-se que o primeiro *record* corresponde ao cabeçalho, enquanto que os restantes irão corresponder às transações a analisar.

3.2 Field Separator (FS)

Aqui foi tomada a decisão de escolher como *field separator* os caracteres <e >. Cada transação é definida por vários segmentos com o seguinte formato:

```
<...>...</...>\n
```

Com o *field separator* que foi anunciado anteriormente, os dados relevantes para análise encontrar-se-ão sempre no segundo (tag) e terceiro (valor) campo de cada linha. Desta forma, para se obter o valor de determinada tag, basta iterar pelos campos de cada linha, obtendo-se a seguinte função:

```
int main() {  
    return 0;  
}
```

Capítulo 4

Apresentação das Queries e Resoluções

a) Calcular o número de 'entradas' em cada dia do mês

```
print "Data \t\t N° de entradas"

for (i in dias) {
    print i "\t\t" dias[i];
}
```

b) Escrever a lista de locais de saída

```
print "Local \t\t N° de saídas"
for (i in locais) {
    print i "\t\t" locais[i];
}
```

c) Calcular o total gasto no mês

```
total = 0;
iva = 0;
for (i in imp) {
    total += imp[i] * (ivas[i]/100 + 1) - desc[i];
}
print total;
```

d) Calcular o total gasto no mês apenas em parques

```
totalP = 0;
totalPt = 0;
for (i in tipos){
    if (tipos[i] ~ /Parque/){
        totalP += imp[i];
    }
    if (tipos[i] ~ /Portagens/){
        totalPt += imp[i];
    }
}
print "Total gasto em parques: " totalP;
print "Total gasto em portagens: " totalPt;
```

Capítulo 5

Conclusão

CONCLUSÃO DO TRABALHO (mais paneleirices para o João Coelho se entreter).
Trabalho futuro.