

Integração de Dados

Duração: 90 minutos

Licenciatura em Engenharia Informática: 2º ano - 2º semestre

2016/2017

TESTE N° 1

Responda às questões deste teste no espaço reservado. Use o ficheiro XML fornecido em anexo.

Nome: _____ N° aluno: _____

1. [15%] Sistemas de Integração de Dados

Analise a situação descrita de seguida e responda às questões colocadas.

A editora de livros **TodosALer** expandiu o seu negócio adquirindo uma livraria de vendas exclusivas on-line **BooksforAll**. No processo de fusão surgiu a necessidade de integração dos dados das bases de dados das duas empresas. A editora **TodosALer** tem três bases de dados principais:

- LivrosPublicados(isbn, titulo, autorId, preco, direitosAutor)
- Escritores(autorId, nome, contacto, data_nasc)
- Clientes(ClienteId, nome, morada, contacto)

A livraria **BooksforAll** tem duas bases de dados principais:

- Livros (isbn, titulo, Nomeautor, preçoVenda, desconto)
- Clientes(ClienteId, nome, apelido, contacto, data_nasc)

No processo de junção tornou-se fundamental fazer pesquisas sobre os clientes e sobre os livros sem obter dados redundantes ou inconsistentes, uma vez que as bases de dados originais possuem informação repetida sobre livros e clientes.

- Pesquisa 1: dado o id de cliente, quais os dados (nome, morada, telefone) desse cliente
- Pesquisa 2: dado o isbn de um livro qual o título, o nome do autor e o preço desse livro

Identifique e detalhe os componentes de um possível sistema de integração de dados capaz de responder às pesquisas anteriores.



2. [20%] Expressões Regulares

- a) Assinale as cadeias encontradas pelas seguintes expressões regulares (respostas incorrectas serão penalizadas)

`\b(ab)+[ab]+\b`

A -- abab
B -- abbb
C -- aabbab
D -- ab
E -- babbb
F -- bbaa
G -- abaaaa

`\b[ab][acd]?(^ab)+\b`

A -- andre
B -- alic
C -- balde
D -- acida
E -- bonita
F -- bolos
G -- adoro

`\b[ana]*[bela]?(^top)\b`

A -- ana
B -- bela
C -- top
D -- alo
E -- anel
F -- annaaaa
G -- bis

- b) Escreva uma expressão regular que encontre **palavras** com caracteres alfabéticos. As palavras devem ser começadas e terminadas pelas vogais **a** ou **u** (**maiúscula ou minúscula**) e com tamanho entre 2 e 7 caracteres. Alguns exemplos: **ursa, Alma, Unida, ALuA, UIVA, uva,**

- c) Escreva uma expressão regular que encontre **frases** interrogativas de tamanho entre 10 e 20 caracteres. Nas frases assuma que começam por qualquer carácter maiúsculo, seguidos de caracteres minúsculos, espaçamentos ou dígitos e terminando com o sinal de interrogação.

- d) Escreva uma expressão regular que encontre as **palavras** que comecem por **ga, go, gi** ou **gr** seguidas de qualquer carácter minúsculo.

Na frase seguinte, as palavras a sublinhado mostram exemplos do que a ER deve validar.

o gato preto e grande gosta de gaivotas brancas gigantes

- e) Analise o seguinte ficheiro de texto que armazena horas de entrada num servidor. O primeiro campo é um identificador, o segundo campo é o dia e o mês de acesso e o terceiro campo a hora.

```
log01 - 12/01 - 09:00
log02 - 05/01 - 06:10
log02 - 05/01 - 06:66
log03 - 07/01 - 10:00
log99 - 19/01 - 08:57
log123 - 21/01 - 09:01
log1099 - 30/01 - 08:32
log012 - 04/02 - 07:45
```

Recorrendo a grupos, construa uma expressão regular que encontre os dias do mês de Janeiro em que os logs foram efectuados antes das 9:00. A ER regular deve encontrar apenas horas válidas, No ficheiro anterior há duas horas inválidas que não devem ser consideradas (06:66 e 01:99).

3. [15%] JSON/XML

Analise o XML da esquerda e proponha uma representação no formato JSON para guardar a mesma informação.

```
<loja>
  <produto id="p001">
    <nome>Bolachas</nome>
    <preco>12</preco>
    <stock status="ok">25</stock>
  </produto>
  <produto id="p002">
    <nome>Cereais</nome>
    <preco>2.5</preco>
    <stock status="nulo">0</stock>
  </produto>
  ....
</loja>
```

4. XPath [25%]

a) Analise o documento XML fornecido em anexo. Escreva as expressões XPath que permitam encontrar a seguinte informação:

- i. Nomes (elemento **nome**) dos produtos de mercearia

- ii. Lista de produtos de toda a loja com stocks abaixo de 10 unidades e sem desconto.

- iii. Qual o **tipo** do produto dos frescos mais caro?

- iv. Quantos produtos de **papelaria** são da marca Ambar?

- v. Lista de produtos dos **frescos** que se seguem ao “Leite de Soja”

vi. Quantas marcas de **cadernos** existem no ficheiro?

vii. Listagem dos **nomes** (texto) de todos os tipos de Leite do ficheiro

viii. Qual a **média** de preços dos produtos de mercearia?

ix. **Nome** e **preço** do penúltimo produto dos frescos do ficheiro.

x. Elementos **preço** e **validade** (texto) do último produto da lista de frescos do ficheiro. A resposta deve fazer utilização obrigatória do eixo de navegação **following-sibling**.

b) Indique o que devolvem as seguintes pesquisas XPath. Transcreva do ficheiro a informação que é devolvida.

i. `//produto[contains(@tipo,"cereais")]/nome`

ii. `/loja/frescos/following-sibling::papelaria/produto[last()-2]/following::marca`

iii. `/loja//produto[last()]/preceding-sibling::*[1]/nome`

5. [10%] DTD

a) Escreva o DTD que valide o elemento **mercearia**

b) Escreva o DTD que valide o elemento **desc.**

c) Escreva o DTD que valide o elemento **loja.**

d) Escreva o DTD que valide o elemento **validade.**

e) Escreva o DTD que valide o elemento **stock.**

f) Escreva o DTD que valide o atributo **desconto** uma enumeração de 2 valores (sim, não) e opcional.

g) Escreva o DTD que valide o atributo **quant** (obrigatório).

6. [15%] XSD

Escreva apenas o XSD solicitado. Use a instrução **ref** para referenciar os elementos e atributos necessários, assumindo que se encontram definidos.

a) Escreva o XSD que valide o elemento **validade** como data

b) Escreva o XSD que valide o atributo **quant** como inteiro

c) Escreva o XSD que valide o elemento **frescos**

d) Escreva o XSD que valide o elemento **preco**

h) Escreva o XSD que valide o atributo **desconto** uma enumeração de 2 valores (sim, não).

- i) Escreva um **tipo de dados** XSD que permita validar os elementos **mercearia**, **frescos** e **papelaria**

- j) Usando o tipo de dados anterior, escreva o XSD para validar os três elementos **mercearia**, **frescos** e **papelaria**