

FEUP

Licenciatura em Engenharia Informática e Computação
Tecnologia de Sistemas de Gestão de Bases de Dados
2001/2002

Exame de Avaliação

12 de Julho de 2002

NOME: _____

Observe por favor as seguintes instruções:

- Leia cuidadosamente o exame até ao fim por forma a escolher a sua estratégia.
- O exame tem a duração máxima de duas horas e meia (150 minutos).
- O exame é com consulta de todo o material próprio trazido para o efeito.
- Deve responder nos espaços fornecidos neste exame, podendo usar, em último recurso, o espaço das costas da folha.
- O exame tem 10 perguntas, com as pontuações indicadas, totalizando 100 pontos.

| Problema | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Total | NOTA |
|-------------|---|----|----|---|----|----|----|----|----|----|-------|------|
| Máx. Pontos | 5 | 15 | 15 | 5 | 10 | 10 | 10 | 10 | 10 | 10 | 100 | – |
| Pontos | | | | | | | | | | | | |

João Correia Lopes

1. Armazenamento de Dados: Ficheiros e Índices [5 pontos]

Na *buffer pool* mantida por um gestor de buffers de um SGBD pode ocorrer um fenómeno conhecido como *sequential flooding*.

Diga em que consiste este fenómeno e dê um exemplo ilustrativo de como ele pode acontecer na prática, de acordo como uma dada política de troca de páginas.

2. Indexação e Ordenação [15 pontos]

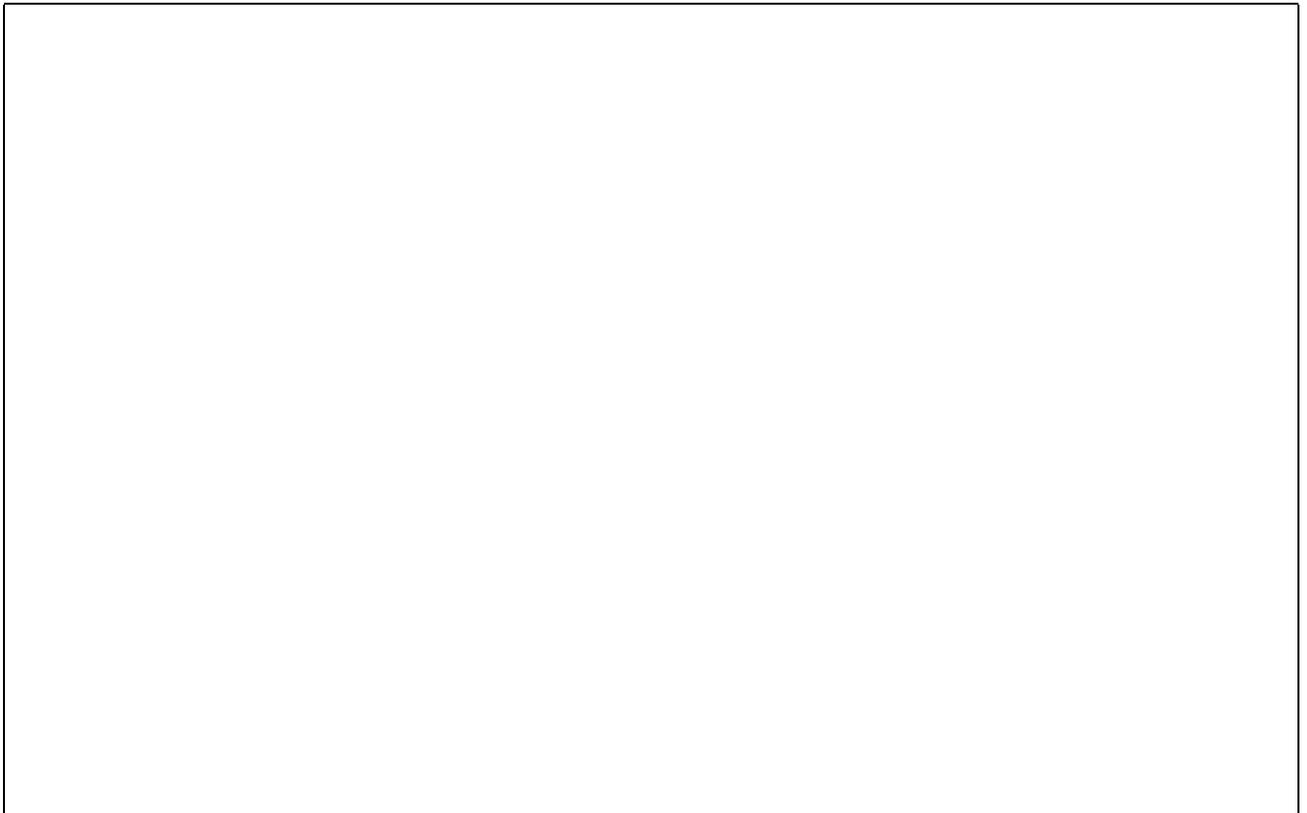
Considere que se pretende ordenar um ficheiro com 4500 registos de 48 bytes cada um, guardados em páginas de 512 bytes. A chave de ordenação (procura) tem 4 bytes, os identificadores de registo (*rid*) têm 8 bytes e os identificadores de página têm 4 bytes.

- a) Supondo que existe um índice em *B+ tree* para o campo de ordenação, calcule o custo de retirar os registos pela ordem desejada para o caso (1) em que o índice é agrupado e (2) o índice não é agrupado.

- b) Supondo que estão disponíveis 4 páginas de buffer, calcule o número de corridas ordenadas produzidas pelo 1º passo de um algoritmo de ordenação externa, calcule o número de passos necessários para ordenar o ficheiro e o custo total da ordenação em termos de I/O.



- c) Calcule o número de registos do maior ficheiro que consegue ordenar com as 4 páginas de buffer em apenas 2 passos; refaça os cálculos para o caso de existirem 257 páginas de buffer.



3. Optimizaç o de interroga es [15 pontos]

Considere o seguinte esquema de rela o de empregados de uma dada empresa:

Emp(nome, cargo, cidade, dept)

em que os dom nios s o *strings* do mesmo tamanho, e a seguinte interroga o em SQL:

```
SELECT nome
FROM Emp
WHERE cargo="Capataz" AND dept="Obras";
```

Considere ainda que apenas 10% dos registos satisfaz a condi o de selec o cargo="Capataz", que apenas 10% dos registos satisfaz dept="Obras", que 5% dos registos satisfaz a conjun o das duas condi es, que a rela o cont m 10 000 p ginas e que existem 10 p ginas de buffer dispon veis.

- a) Supondo a exist ncia apenas de um  ndice aglomerado do tipo *B+ tree* no campo cargo, descreva o melhor plano para responder   pergunta e calcule o respectivo custo em I/O.

- b) Supondo a exist ncia apenas de um  ndice aglomerado do tipo *B+ tree* nos campos <cargo,dept>, descreva o melhor plano para responder   pergunta e calcule o respectivo custo em I/O.

4. Limitações do Modelo Relacional [5 pontos]

Para resolver limitações do Modelo Relacional foram propostos novos modelos de dados, nomeadamente o Modelo Relacional-Objecto e o Modelo Orientado aos Objectos. Nos últimos 15 anos tem-se assistido a um aumento sistemático do número de novas instalações Relacional-Objecto e de migração de instalações existentes para este modelo, em contraste com o número reduzido de instalações de SGBDs Orientadas aos Objectos.

Identifique e descreva brevemente o que, na sua opinião, tem contribuído para esta situação.

5. SQL3, ADTs e Colecções [10 pontos]

Suponha que pretende guardar numa base de dados informação sobre os empregados de uma dada empresa.

Para cada empregado, identificado pelo número de contribuinte (`nctb`), deve guardar-se `anos` (os anos que o empregado trabalhou para a empresa) e `nome`. Há empregados “a prazo” e normais e o salário é calculado, de maneira diferente para os dois tipos, por um método `calcSalario` que aceita `anos` como parâmetro. Para cada empregado normal deve ser guardado o nome e a idade de cada filho. Para cada departamento, identificado pelo código (`cod`), deve guardar-se `nome` e `emps` (conjunto de empregados que trabalham no departamento).

Considere ainda válidas as seguintes restrições de integridade:

R1: Nenhum empregado a prazo trabalha na empresa mais do que 3 anos.

R2: Qualquer empregado normal ganha mais do que qualquer um dos empregados a prazo.

R3: Nenhum empregado com mais do que três filhos ganha menos do que 1000.

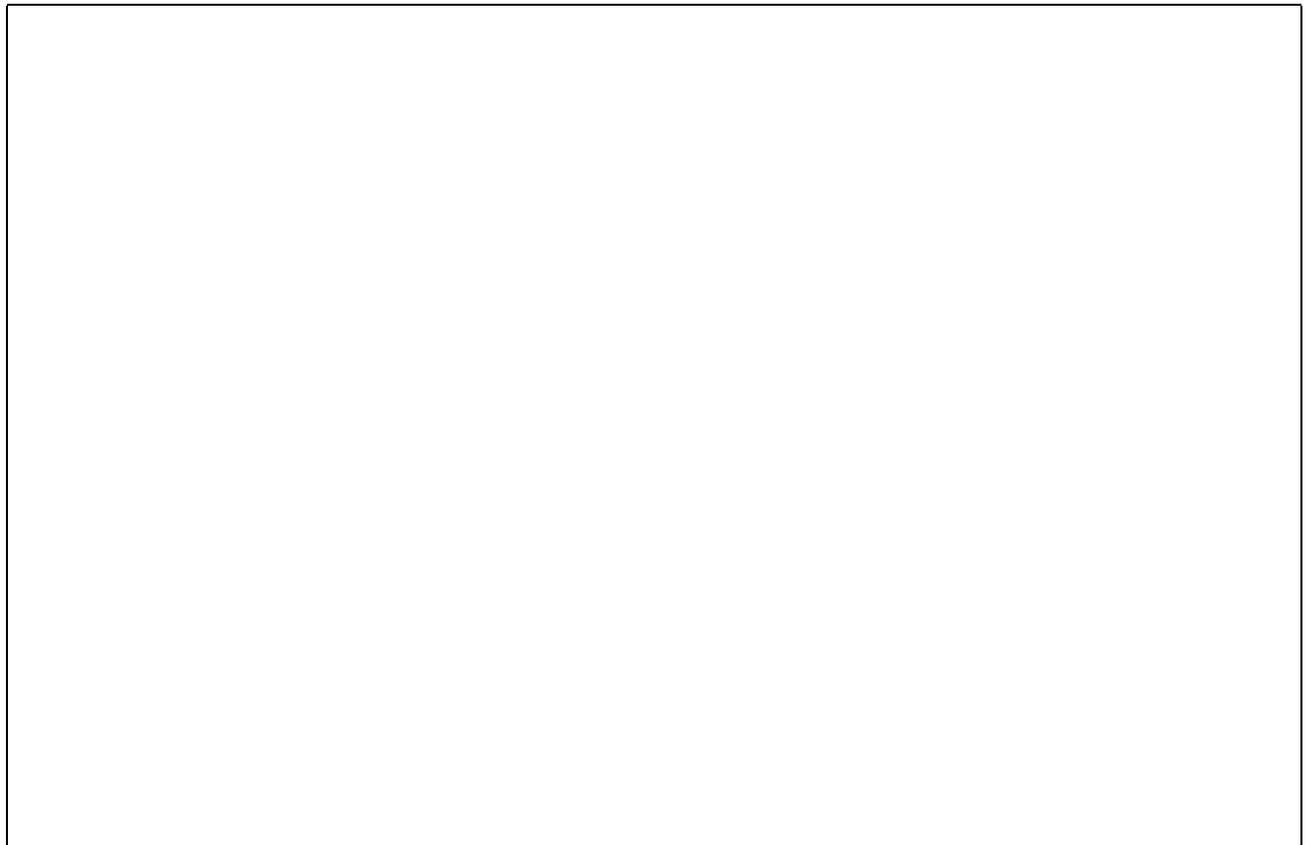
Complete o esquema relacional-objecto SQL3, por exemplo usando a notação apresentada nas aulas, considerando os requisitos enumerados para a aplicação referida, sem esquecer a primeira restrição de integridade (R1).

```
// Dominio para empregado Normal ou a Prazo
CREATE DOMAIN TipoDeEmp AS CHAR(1) DEFAULT 'N' CHECK (tipo IN ('N','P'));

// tipo para empregado
CREATE TYPE TipoEmp (
    nctb      CHAR(8) UNIQUE NOT NULL,
    nome      CHAR(52),
    anos      INTEGER,
    tipo      TipoDeEmp,
    filhos    LIST(REF(TipoFilho)),
    PUBLIC FUNCTION CalcSalario(anos: integer): REAL
);

// tipo para departamento
CREATE TYPE TipoDept (
    cod       INTEGER UNIQUE NOT NULL,
    nome      CHAR(52),
    emps     LIST(REF(TipoEmp))
);

// tipo para filho
CREATE TYPE TipoFilho (
    nome      CHAR(52),
    nasce     DATE
);
```



6. Módulos Persistentes em SQL3 [10 pontos]

Considere novamente a base de dados do problema 5.

Apresente o código SQL de um módulo persistente de servidor com a função e o procedimento seguintes:

```
maisUmAno(); // incrementa em um o número de anos dos empregados que têm
// mais do que dois filhos
//
quantosAPrazo(dep, sal): integer; // devolve o numero de empregados a prazo do
// departamento dep que ganham mais do que o salário sal
```

7. Restrições de Integridade e Gatilhos [10 pontos]

Considere novamente a base de dados do problema 5.

- a) Escreva uma asserção em SQL3 para impor a restrição R2.

- b) Escreva um ou mais gatilhos em SQL3 para impor a restrição R3 de forma incremental ao ser actualizada a informação sobre empregados.

8. Estrutura Lógica de Documentos XML [10 pontos]

Considere o seguinte DTD para documentos XML:

```
<!DOCTYPE Empresa [
  <!ELEMENT EMP (DEPARTAMENTO*)>
  <!ELEMENT DEPARTAMENTO (LOCAL+, EMPREGADO*)>
  <!ATTLIST DEPARTAMENTO Cod ID #REQUIRED Nome CDATA>
  <!ELEMENT LOCAL EMPTY>
  <!ATTLIST LOCAL Cod ID #REQUIRED Nome CDATA>
  <!ELEMENT EMPREGADO (FILHO | FILHA)*>
  <!ATTLIST EMPREGADO Cod ID #REQUIRED Nome CDATA #REQUIRED Telefone CDATA
    Locais IDREFS #REQUIRED>
  <!ELEMENT FILHO (#PCDATA)>
  <!ELEMENT FILHA (#PCDATA)>
]>
```

Verifique se o documento XML seguinte é bem formado e se está conforme com o DTD apresentado (isto é, se é válido); no caso de não estar, assinale os pontos onde isso se verifica.

```
<?XML VERSION="1.0" STANDALONE="no"?>
<!DOCTYPE Empresa SYSTEM "../DTDs/empresa.dtd">
<EMP>
  <DEPARTAMENTO Cod="D1" Nome="Vendas">
    <EMPREGADO Cod="E1" Nome="João Lopes" Telefone="123456789">
      <EMPREGADO Cod="E2" Nome="Joaquim Lopes" Telefone="123456789">
    </DEPARTAMENTO>
  <DEPARTAMENTO Cod="D2" Nome="Produção">
    <LOCAL Cod="L1" Nome="Porto"/>
    <EMPREGADO Cod="E1" Nome="João Lopes" Telefone="123456789" Local="L1">
    <EMPREGADO Cod="E2" Nome="Joaquim Lopes" Telefone="123456789" Local="L1">
    <LOCAL Cod="L2" Nome="Lisboa">
    <EMPREGADO Cod="E3" Nome="João Lopes" Locais="L1 L2 E1">
      <FILHA>"Maria Lopes"</FILHA>
      <FILHO>"Pedro Lopes"</FILHO>
    </EMPREGADO>
  </DEPARTAMENTO>
  <DEPARTAMENTO Cod="D3">
    <LOCAL Cod="L3" Nome="Porto"></LOCAL>
  </DEPARTAMENTO>
</EMP>
```

9. Transformação e apresentação de XML [10 pontos]

Considere novamente o DTD apresentado no problema 8 e instâncias de documentos XML de acordo com esse DTD.

Apresente um conjunto de regras de transformação XSLT que permitam passar para HTML, para ser mostrado num navegador Web, o nome de todos os departamentos e, para cada um, o nome dos empregados que têm filhos e o nome dos respectivos filhos.

10. XML Schemas [10 pontos]

Para o DTD apresentado no problema 8 e considerando apenas os elementos LOCAL e EMPREGADO, apresente somente a parte correspondente de um XML Schema equivalente ao DTD (isto é, que permita as mesmas instâncias de documentos XML)

FIM.