

# Teste de Base de Dados

(Duração: 2 horas)

Universidade do Algarve

10 de Dezembro de 2001

- Escreva o seu nome, nº de aluno e curso em todas as folhas.
- Não é permitido falar com os colegas durante o exame. Se o fizerem, terão a prova anulada.
- Caso opte por desistir, escreva “Desisto” nesta página, assine e entregue o enunciado ao docente.
- Não pode chamar o docente para lhe ajudar a interpretar o enunciado. A interpretação do enunciado faz parte da avaliação.
- O teste tem 7 perguntas e a cotação de cada uma aparece entre parêntesis.

**(3 valores) Pergunta 1.** Pretende-se desenhar uma base de dados para implementar um search engine para a web. Faça o diagrama entidade-associação para modelar o problema. Não se esqueça de indicar as chaves e outras restrições que possam haver. O search engine será bastante simplificado e apenas deve ter em atenção os seguintes pontos:

1. uma página é identificada por uma URL  
(ex: `http://w3.ualg.pt/~flobo/bd/aula1.html`).
2. uma URL pode ser decomposta em 3 partes: protocolo (ex: `ftp`, `http`), domínio (ex: `w3.ualg.pt`), e pathname (ex: `~flobo/bd/aula1.html`). Para efeitos deste exercício, assume que todos os URLs têm o protocolo `http`.
3. uma página pode ter links que apontam para outras páginas.
4. uma página tem várias palavras.
5. a mesma palavra pode aparecer várias vezes na mesma página. Deves ter uma maneira de distinguir as várias ocorrências de cada palavra.
6. cada palavra tem atributos de estilo (normal, bold, ou itálico). Além disso, cada palavra pode aparecer dentro ou fora de tags especiais (para efeitos deste exercício, considera apenas os tags `<title>`, `<h1>`, e `<h2>`).

**(0,5 valores) Pergunta 2.** Ao realizar a pergunta anterior, é quase garantido que criou um conjunto de entidades para representar páginas web. Para chave desse conjunto de entidades pode-se utilizar a URL, mas também pode-se atribuir um código a cada página (por exemplo, um número inteiro único). Qual das duas estratégias é que acha mais adequada? Justifique a resposta.

**(3 valores) Pergunta 3.** Converta o modelo entidade-associação para o modelo relacional, e crie as respectivas tabelas e restrições utilizando os mecanismos oferecidos pela linguagem SQL.

**(8 valores) Pergunta 4.** Esta alínea é baseada no modelo relacional obtido na pergunta 3. Escreva código em SQL para responder às seguintes perguntas.

- Qual a(s) página(s) que têm a palavra Jardel num tag <h1> ou <h2>?
- Qual a(s) página(s) do domínio pt que são mais referenciadas? Nota: self-links (páginas que apontam para si próprias) devem ser ignorados.
- Crie uma restrição adicional que não permita que uma página tenha mais do que um <title>.

**(0,5 valores) Pergunta 5.** Acha que os verdadeiros search engines (ex: Google, Altavista) são implementados utilizando um sistema de gestão de base de dados relacional? Justifique a resposta.

**(2 valores) Pergunta 6.** Considere as relações Filmes e Participa:

```
Filmes(filme,ano,duracao,estudio)
```

```
----- ----
```

```
Participa(filme,ano,actor)
```

```
----- --- -----
```

Pretende-se saber quais os estúdios que já produziram filmes em que tenha participado o Tom Cruise. De seguida, apresenta-se uma possível tentativa de resposta, quer em álgebra relacional, quer em SQL.

$$\pi_{estudio} \left( \sigma_{actor = 'Tom Cruise'} (Participa \bowtie Filmes) \right)$$

```
SELECT estudio
FROM Participa NATURAL JOIN Filmes
WHERE actor='Tom Cruise';
```

1. Poderá existir alguma diferença entre o resultado obtido pela query feita em álgebra relacional e pela query feita em SQL? Justifique a resposta.
2. Está alguma coisa mal feita na query em álgebra relacional e/ou na query em SQL. Em caso afirmativo, corrija o que está mal.

**(3 valores) Pergunta 7** Ontem fiz uma maratona a corrigir os trabalhos dos alunos de Programação I. Os trabalhos eram individuais e tinham uma nota de zero a vinte. Para dispensarem de exame, os alunos tinham de obter nota superior a 6 no trabalho. Infelizmente, as notas foram desastrosas e houve mesmo uma grande percentagem de alunos que cometeu fraude (copiaram). O resultado da correcção foi uma relação/tabela de nome AlunosP1 contendo 5 atributos:

```
AlunosP1( numero, curso, nota, maisque6, copiou )
-----
```

- \* 'numero' é chave da relação.
- \* 'curso' pode ser IG, EI, ou ESC.
- \* 'nota' é um inteiro de 0 a 20.
- \* 'maisque6' é um boolean (TRUE se a nota > 6, FALSE caso contrário).
- \* 'copiou' também é um boolean (TRUE se copiou, FALSE se não copiou).  
Para os alunos que copiaram a nota foi zero (e um processo disciplinar para o Conselho Pedagógico e Directivo da FCT/Ualg).

1. Existem dependências funcionais completamente não triviais na relação AlunosP1?  
Em caso afirmativo, indique quais são. Justifique a resposta.
2. A relação AlunosP1 está na forma BCNF? Em caso negativo, decomponha AlunosP1 em várias relações de modo a que todas elas fiquem na forma BCNF. Justifique a resposta.