

**21053 - Fundamentos de Bases de Dados**  
**2012-2013**  
**e-fólio C**  
**Resolução e Critérios de Correção**

**PARA A RESOLUÇÃO DO E-FÓLIO, ACONSELHA-SE QUE LEIA  
ATENTAMENTE O SEGUINTE:**

- 1) O e-fólio é constituído por 3 perguntas. A cotação global é de 3 valores.
- 2) O e-fólio deve ser entregue num único ficheiro PDF, não zipado, com fundo branco, com perguntas numeradas e sem necessidade de rodar o texto para o ler. Penalização de 1 a 3 valores.
- 3) Não são aceites e-fólios manuscritos, i.e. tem penalização de 100%.
- 4) O nome do ficheiro deve seguir a normal “eFolioC” + <nº estudante> + <nome estudante com o máximo de 3 palavras>. Penalização de 1 a 3 valores.
- 5) Na primeira página do e-fólio deve constar o nome completo do estudante bem como o seu número. Penalização de 1 a 3 valores.
- 6) Durante a realização do e-fólio, os estudantes devem concentrar-se na resolução do seu trabalho individual, não sendo permitida a colocação de perguntas ao professor ou entre colegas.
- 7) A interpretação das perguntas também faz parte da sua resolução, se encontrar alguma ambiguidade deve indicar claramente como foi resolvida.
- 8) A legibilidade, a objectividade e a clareza nas respostas serão valorizadas, pelo que, a falta destas qualidades serão penalizadas.

A informação da avaliação do estudante está contida no **vetor das cotações**:

Questão: 1 2 3.1 3.2 3.3

C: 5 10 5 5 5 décimas

1) (1/2 valor) Escreva um texto, com 200 a 400 palavras, sobre o tema “SQL versus NoSQL”.

A “Structured Query Language”, SQL, é a linguagem declarativa de consulta em bases de dados relacionais. O SQL foi desenvolvido originalmente no início dos anos 1970 nos laboratórios da IBM em São José, EUA, no sistema R, que tinha por objetivo demonstrar a viabilidade da implementação do modelo relacional proposto por E. F. Codd.

O SQL tem tido várias atualizações. O SQL-1999 disponibiliza consultas recursivas e gatilhos (triggers). O SQL-2003 introduz características para o XML. As linguagens de consultas como o XQuery e o SPARQL são as variantes para as bases de dados XML e RDF. O SQL é utilizado pelos maiores construtores de bases de dados relacionais, tais como DB2 da IBM, Oracle, MySQL ou MS-SQL Server.

O termo NoSQL significa "Not Only SQL" e pronuncia-se “noseeque”. Carlo Strozzi cunhou o termo NoSQL em 1998 quando da divulgação da sua base de dados relacional em “open-source”.

A operação de junção de base de dados relacionais pode degradar o desempenho do sistema e especialmente com tabelas de milhões de linhas. Por essa razão a Google e a Amazon desenvolveram as suas próprias bases de dados. Atualmente o termo NoSQL é cada vez mais conotado com uma classe de bases de dados não-relacionais e com grande escalabilidade. Existe um vasto conjunto de soluções NoSQL: Wide Column Store, Document Store, Key Value / Tuple Store, Graph Databases, Object Databases, Grid & Cloud Database Solutions, XML Databases e Multivalued Database.

Para ambientes de informática de gestão corrente e estruturados as bases de dados relacionais continuam a ser as mais adequadas. Para volumes de dados de grandes dimensões, pouco estruturados ou com complexidades para além da gestão corrente, as bases de dados NoSQL poderão ser as escolhidas.

Critérios de correção:

- discutir os SQL 0,2 valores
- discutir os NoSQL 0,2 valores
- comparar 0,1 valores

2) (1 valor) Dado o conjunto de dependências funcionais

$F = \{a \rightarrow abce, ac \rightarrow acde, b \rightarrow c, bc \rightarrow cd, bd \rightarrow bcd\}$

encontre a cobertura canónica (não redundante e reduzido à esquerda) e a redução à direita, utilizando as rotinas da bibliografia de David Maier do Capítulo 5.

Resposta:

$a \rightarrow abce$   
 $ac \rightarrow acde$   
 $b \rightarrow c$   
 $bc \rightarrow cd$   
 $bd \rightarrow bcd$

No Redundant FD  $X \rightarrow Y$

Member( $a \rightarrow abce$ )=0;  $abce \not\subseteq \text{Closure}=a$ ;  
Member( $ac \rightarrow acde$ )=1;  $acde \subseteq \text{Closure}=abcde$ ;  
Member( $b \rightarrow c$ )=0;  $c \not\subseteq \text{Closure}=b$ ;  
Member( $bc \rightarrow cd$ )=0;  $cd \not\subseteq \text{Closure}=bc$ ;  
Member( $bd \rightarrow bcd$ )=1;  $bcd \subseteq \text{Closure}=bcd$ ;

$a \rightarrow abce$   
 $b \rightarrow c$   
 $bc \rightarrow cd$

Left Reduced FD  $(X-A) \rightarrow Y$

(X-a) Member( $\rightarrow abce$ )=0;  $abce \not\subseteq \text{Closure}=\text{=}$ ;  
(X-b) Member( $\rightarrow c$ )=0;  $c \not\subseteq \text{Closure}=\text{=}$ ;  
(X-b) Member( $c \rightarrow cd$ )=0;  $cd \not\subseteq \text{Closure}=c$ ;  
(X-c) Member( $b \rightarrow cd$ )=1;  $cd \subseteq \text{Closure}=bcd$ ;

$a \rightarrow abce$   
 $b \rightarrow c$   
 $b \rightarrow cd$

Right Reduced FD  $X \rightarrow (Y-A)$

(Y-a) Member( $a \rightarrow a$ )=1;  $a \subseteq \text{Closure}=abcde$ ;  
(Y-c) Member( $a \rightarrow c$ )=1;  $c \subseteq \text{Closure}=abcde$ ;  
(Y-) Member( $a \rightarrow$ )=0;  $\_ \not\subseteq \text{Closure}=abcde$ ;  
(Y-) Member( $a \rightarrow$ )=0;  $\_ \not\subseteq \text{Closure}=abcde$ ;  
(Y-c) Member( $b \rightarrow c$ )=1;  $c \subseteq \text{Closure}=bcd$ ;  
(Y-c) Member( $b \rightarrow c$ )=0;  $c \not\subseteq \text{Closure}=bd$ ;  
(Y-d) Member( $b \rightarrow d$ )=0;  $d \not\subseteq \text{Closure}=bc$ ;

$a \rightarrow be$   
 $b \rightarrow$   
 $b \rightarrow cd$

Pack:  $a \rightarrow be, b \rightarrow cd$

Critérios de correção:

- FD não redundantes 0,5 valores
- FD reduzidas à esquerda e direita 0,5 valores
- Penalização de 50% a 100% para respostas sem justificação

### 3) Projeto de Bases de Dados

Um empresa, com mais de 300 funcionários, utiliza o e-mail para responder às diversas solicitações que lhe são colocadas do exterior e do interior. A empresa tem as seguintes características:

- A empresa tem feito grandes esforços para passar de uma empresa “brick and mortar” para uma empresa totalmente “on-line”.
- Foi criada um longa lista de “quem faz o quê”, onde estão atribuídas as tarefas para cada funcionário.
- Contudo, as queixas dos clientes mantêm-se. Eles referem que são constantemente reenviados de serviço para serviço sem obter uma resposta concreta.

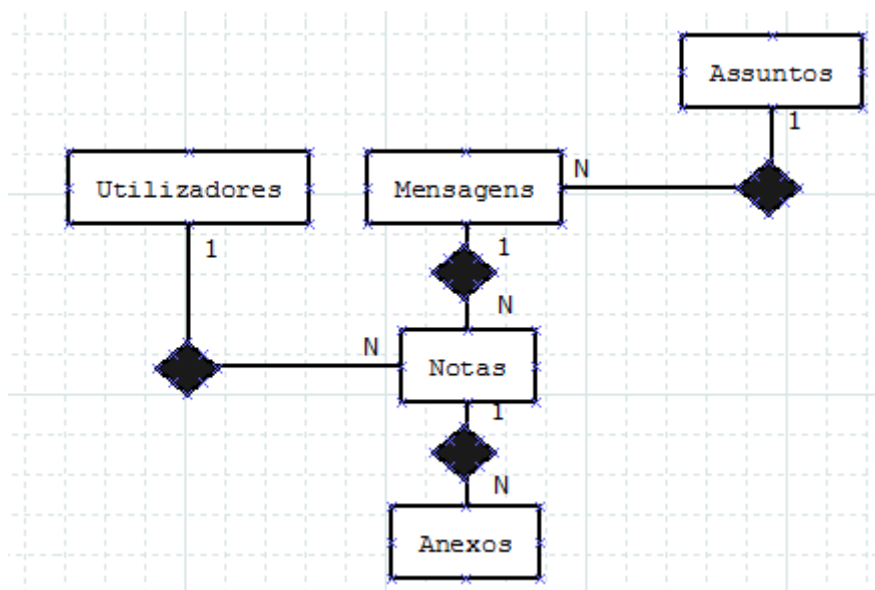
Depois de várias reuniões os analistas concluíram de a associação (“matching”) entre as funções dos empregados e as solicitações dos clientes é muito difícil de estabelecer. Foi referido o seguinte:

- Visto que é impossível controlar os múltiplos fluxos de e-mails, e de estabelecer o referido “matching” a empresa decidiu banir a utilização do e-mail interno e criar uma nova aplicação informática.
- Na organização foi criado uma nova estrutura para triagem e controlo das solicitações, mantendo-se os diferentes serviços operacionais.
- O novo serviço de Triagem&Controlo para cada mensagem que entra é associado um serviço que trata do assunto. Em cada serviço existe um funcionário que distribuí as mensagens pelos colegas. Quando o funcionário não consegue responder à solicitação devolve a mensagem ao serviço de Triagem&Controlo que encontra uma forma de resolver o problema.
- A empresa vai recorrer à utilização de formulários on-line como forma entrada das solicitações e banir progressivamente os e-mails.
- No novo repositório de mensagens, que substituí as várias caixas de e-mail, cada mensagem deve associar anexos, deve associar mensagens auxiliares (ou notas), deve referir os utilizadores intervenientes e deve estar classificada por assunto.

3.1) (1/2 valor) Comece por identificar as diferentes entidades do Modelo Entidade-Relação. De seguida, identifique os relacionamentos entre as entidades do Modelo Entidade-Relação. Dê um nome ao relacionamento. Classifique cada relacionamento quanto à cardinalidade (1:1, 1:N, N:N). De seguida, desenhe o diagrama do Modelo Entidade-Relação. Identifique os relacionamentos e os atributos das entidades. Em alternativa, pode apresentar um modelo de classes do UML (Resposta: 1 página)

Entidades: Mensagens, Notas, Assuntos, Utilizadores e Anexos

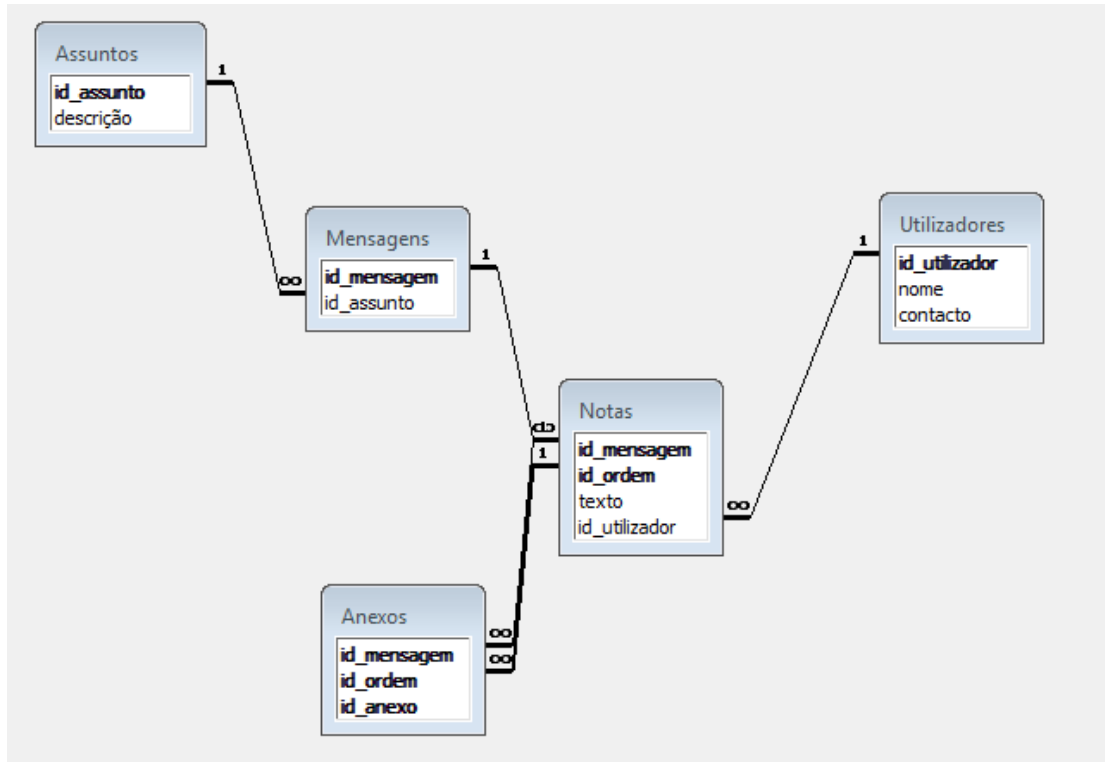
Relacionamentos: todos são 1:N e com nome “tem”



CrITÉrios de correção:

- o estudante deve considerar essencialmente o último parágrafo “No novo repositório de mensagens, que substituí as várias caixas de e-mail, cada mensagem deve associar anexos, deve associar mensagens auxiliares (ou notas), deve referir os utilizadores intervenientes e deve estar classificada por assunto.”
- o estudante deve identificar as entidades: Mensagens, Notas, Assuntos, Utilizadores, Anexos
- Penalização de 0,3 valores para falta das referidas entidades ou pela existência de entidades desadequadas como o serviço de Triagem&Controlo;

3.2) (1/2 valor) Desenhe a base de dados relacional correspondente ao modelo anterior, em que nas ligações de 1:N, a tabela com uma única linha é desenhada em cima e da tabela com várias linhas é desenhada por baixo. A base de dados não deve exceder as 7 tabelas. Se utilizar ligações que denotem transitividade, deve justificá-la convenientemente. (Resposta: 1 página)



Critérios de correção:

- identificar as tabelas: Mensagens, Notas, Assuntos, Utilizadores, Anexos
- Penalização de 0,3 valores para falta de referidas tabelas ou pela existência de tabelas desadequadas;
- Penalização de 0,3 valores para relações transitivas não justificadas;

3.3) (1/2 valor) Identifique três factores críticos de sucesso (FCS) e escreva as respectivas consultas em SQL. (Resposta: 1 página)

FCS 1) Quantas mensagens (notas) tem cada utilizador?

```
SELECT Utilizadores.id_utilizador, COUNT(Notas.texto)
FROM Utilizadores, Notas
WHERE Utilizadores.id_utilizador = Notas.id_utilizador
GROUP BY Utilizadores.id_utilizador;
```

FCS 2) Quantas mensagens (notas) tem cada assunto?

```
SELECT Assuntos.id_assunto, COUNT(Notas.texto)
FROM Assuntos, Mensagens, Notas
WHERE Assuntos.id_assunto = Mensagens.id_assunto
AND Mensagens.id_mensagem = Notas.id_mensagem
GROUP BY Assuntos.id_assunto;
```

FCS 3) Quantas notas tem cada mensagem? Crie mapa por ordem descendente.

```
SELECT Mensagens.id_mensagem, COUNT(Notas.texto)
FROM Mensagens, Notas
WHERE Mensagens.id_mensagem = Notas.id_mensagem
GROUP BY Mensagens.id_mensagem
ORDER BY COUNT(Notas.texto) DESC;
```

**CrITÉrios de correção:**

- 1/3 para cada FCS