Exame Preliminar de Doutorado em Banco de Dados

Duração: 3 horas

9 de agosto de 1999

Atenção:
Apesar da soma dos pontos das sete questões desta prova ser 14, tente responder corretamente o número máximo de questões, a fim de maximizar as suas chances de uma boa avaliação. A nota é o máximo das somas das notas de questões, tomado sobre os subconjuntos de questões cujo valor soma 10. Seja claro e preciso em suas respostas.
Boa sorte!

1. (Valor 2 pontos) Uma administradora de serviços médicos deseja controlar o cadastro de seus serviços e dos participantes desses serviços. Para isso, os dados de hospitais, clínicas, laboratórios, médicos, beneficiários (usuários dos serviços) devem ser armazenados. Existem alguns dados em comuns entre os elementos acima descritos (por exemplo endereço, telefone). Os hospitais e clínicas possuem nome, inscrição estadual. Os médicos e usuários possuem nome endereço, RG. Em particular, os médicos tambem possuem CRM (registro médico). Esses elementos são também classificados em físicos e jurídicos de acordo com o tipo do documento de identificação: CGC ou CPF. Hospitais, clínicas e laboratórios possuem CGC. Médicos e beneficiários possuem CPF. Um elemento não pode ser ao mesmo tempo físico e jurídico.

   Os beneficiários possuem uma particularidade: podem existir beneficiários dependentes, ou seja, estão vinculados a um beneficiário principal. Os dados são os mesmos de um beneficiário.

   Cada beneficiário utiliza serviços e os médicos executam serviços. Os serviços podem ser classificados em consulta, diagnose, cirúrgicos e terapêuticos. Cada serviço possui descrição, código, valor em reais.

   Para essa descrição faça o diagrama do modelo entidade relacionamento (estendido se for o caso) e o mapeie para o modelo relacional.

2. (Valor 3 pontos) Suponha que tenhamos um banco de dados para um agência de turismo especializada em pacotes ecológicos, consistindo dos seguintes atributos:

   - A : agente de viagem
   - E : escritório do agente
   - C : cliente
   - V : pacote de viagem
   - N : número de pessoas em um pacote de viagem de um cliente
   - P : preço do pacote de viagem

   As seguintes dependências funcionais são válidas: $V \longrightarrow P$, $C \longrightarrow A$, $CV \longrightarrow N$, $A \longrightarrow E$.

   (a) Ache uma chave para o esquema R = AECVNP.

   (b) Quantas chaves esta relação tem ? Prove sua afirmação.

   (c) Encontre uma decomposição sem perdas de R que esteja na forma normal de Boyce-Codd.

   (d) Encontre uma decomposição de R na terceira forma normal que seja sem perda e preserve dependências.

   (e) Suponha que R seja representado por CVN, CA, VP e CVE. Encontre uma cobertura mínima para a união das dependências funcionais projetadas nos esquemas.

   (f) Troque a dependência $V \longrightarrow P$ pela dependência multivalorada $V \longrightarrow\!\!\!\!\rightarrow P$. Encontre a base de dependências de C e de EV. Encontre uma decomposição de R na quarta forma normal.

3. (Valor 1 ponto) Descreva um procedimento que deve ser realizado após um crash do disco a fim de recuperar o banco de dados.

4. (Valor 2 pontos) O controle de concorrência pelo método de bloqueios bifásicos (*two-phase locking*) consiste em não permitir que um novo bloqueio seja solicitado por uma transação após o primeiro desbloqueio já ter sido feito. Mostre, através de exemplos, que a violação da regra acima leva à violação da condição de seriabilidade do escalonamento nos casos:

   (a) apenas um novo bloqueio de escrita foi solicitado (após um desbloqueio);

   (b) apenas um novo bloqueio de leitura foi solicitado (após um desbloqueio).

5. (Valor 2 pontos) Fazer um modelo entidade-relacionamento para a seguinte situação. Um banco financia operações de importação e exportação de diversas empresas e diversos produtos.

   Uma empresa é conhecida pelo banco por seu nome e CGC, se é nacional ou estrangeira, e pelo local da sede no país.

   Tanto as operações de importação quanto as de exportação devem ser armazenadas de acordo com o produto envolvido, o mês e ano da operação, o valor (dólar) da transação, a quantidade tramitada e sua unidade. Além disso, precisamos saber o país de origem/destino do produto e o porto de entrada/saída do mesmo.

   Para analisar as propostas de financiamento, o banco precisa ter dados sobre os produtos comercializados. Para isso, para uma série enorme de produtos, o banco deseja armazenar séries históricas anuais que contenham:

   - o valor da mão de obra na produção do produto
   - o faturamento de cada empresa cadastrada para aquele produto
   - a quantidade produzida de cada produto
   - o preço no mercado interno de cada produto.

   Infelizmente, há uma complicação com relação a codificação de produtos. Até recentemente, utilizava-se a Nomenclatura Brasileira de Mercado (NBM), mas com o advento do Mercosul esta foi reposta pela Nomenclatura Comum de Mercado (NCM), onde um produto NBM pode estar relacionado a vários NCM e vice-versa; ou seja não houve uma extensão de nomenclaturas. Por cima disso, ainda existe a codificação TEC de produtos, usada apenas para fins tarifários e ortogonal a NBM e NCM. Faça com que seu modelo aceite produtos codificados nestas três codificações.

6. (Valor 3 pontos) Considere uma hierarquia de generalização com a entidade G genérica e com entidades específicas E1 e E2. O critério de especialização c é sobreponível (isto é, um mesmo elemento pode ser especializado em E1 e E2) e parcial. Cada uma das entidades possuem seus respectivos atributos $a_1, a_2, a_3, \ldots, a_n$. Considere tambem a_1 como sendo o atributo identificador do conjunto de atributos.

   a) mapeie tal estrutura para o modelo relacional.

   b) faca uma consulta em SQL para saber a razão entre o número de ocorrências da Entidade E1 em relação à entidade G.

   c) faça uma inserção de uma tupla de E2 considerando a necessidade de verificar as consistências intrínsicas à hierarquia de generalização.

   d) se o critério c ao invés de ser atômico fosse um intervalo, como por exemplo intervalo de idade (0 -15, 16 - 30, etc), como seria o mapeamento para o modelo relacional ? Refaça a consulta do item b) acima com o critério em questão.

7. (Valor 1 ponto) Considere o seguinte esquema de dados:

LIVRO (Cod_livro, Título, Nome_do_autor)
EDIÇÃO (ISBN, Data, Preço, Editora, Cod_livro)

Assuma que Cod_livro identifica univocamente o livro e que não existem homônimos no atributo Nome_do_autor.

Considere também que ISBN identifica univocamente uma edição.

Expresse cada uma das consultas abaixo em álgebra relacional, se possivel.

a) Obter todos os nomes dos autores que escreveram mais de um livro.

c) Obter a editora que publicou o maior quantidade de edição de livros.

c) Obter o Cod_livro e Título do livro com a edição de menor preço e sua respectiva data.

d) Obter o nome do autor que publicou a maior quantidade de livros na editora selecionada no item b.