

Curso de Tecnologia em Sistemas de Computação
Disciplina Banco de Dados
AD1 1º semestre de 2007.

Nome –

Observações:

1. Prova com consulta.
-

Questão 1

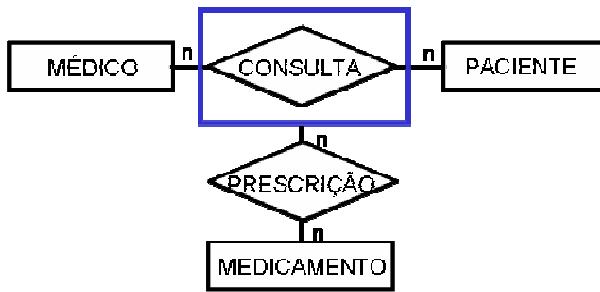
Explique a diferença entre entidade fraca e entidade associativa. Como são representadas no diagrama ER? Dê um exemplo de cada um destes tipos de entidade.

Resposta: Entidade fraca é uma entidade que tem sua existência condicionada à existência de uma outra entidade, da qual depende. É representada no diagrama ER por um traço mais forte no relacionamento que liga a entidade fraca à entidade da qual depende. O traço é colocado do lado da entidade fraca. Já uma entidade associativa é um relacionamento que atua como se fosse uma entidade, podendo estar relacionado a outras entidades. No diagrama ER, uma entidade associativa é representada por um losango dentro de um retângulo.

Exemplo de entidade fraca:



Exemplo de entidade associativa:



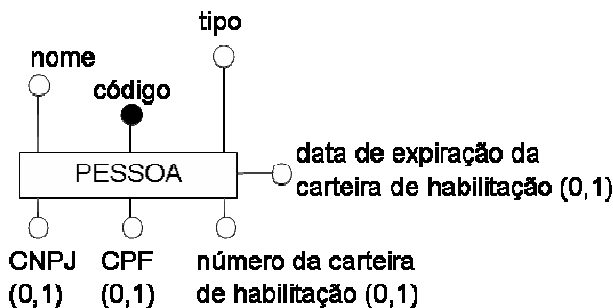
Questão 2

Por que pode-se dizer que atributos multivalorados são indesejáveis em diagramas ER?

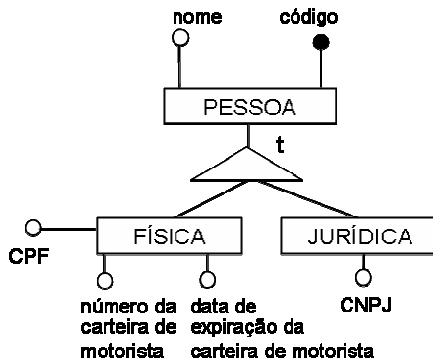
Resposta: Porque atributos multivalorados não têm implementação direta nos SGBDs que seguem o padrão SQL/2. Além disso, tais atributos podem induzir a erros de modelagem, escondendo entidades e/ou relacionamentos que seriam facilmente identificados se o atributo multivalorado tivesse sido modelado como uma entidade relacionada.

Questão 3

Analise o diagrama abaixo e discuta uma opção para representá-lo sem o uso de atributos opcionais.

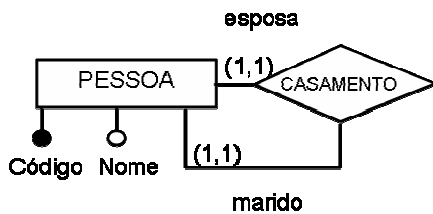


Resposta: Uma opção melhor para representar o diagrama acima é o uso de especialização/generalização, como mostrado abaixo.



Questão 4

Analise o diagrama ER abaixo.



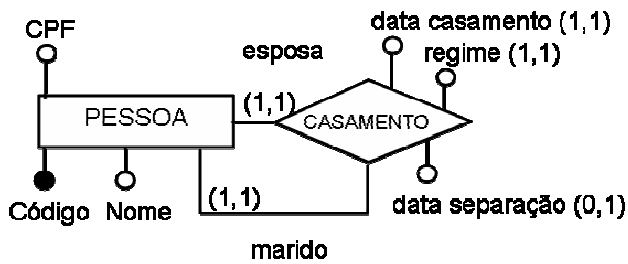
Suponha que deseja-se incluir os seguintes atributos:

- um atributo data para indicar a data do casamento
- um atributo para indicar o CPF da pessoa
- um atributo para indicar a data de separação do casal
- um atributo para indicar o regime do casamento (comunhão total de bens, comunhão parcial de bens, etc.)

Para cada um dos atributos acima, indique onde ele deve ser colocado no diagrama e justifique sua decisão. Redesenhe o diagrama com os novos atributos. Indique a cardinalidade de cada um dos atributos.

Resposta:

- os atributos data do casamento, data de separação e regime devem ser incluídos no relacionamento casamento, já que eles não são informações específicas de pessoa, e sim do casamento em si (cada ocorrência de um par pessoa/pessoa, no relacionamento de casamento possui estas informações).
- o atributo CPF deve ser incluído na entidade pessoa, já que é informação específica de pessoa (cada ocorrência de pessoa possui um CPF)



Questão 5

A *Musical Records* decidiu armazenar informações sobre os músicos que gravam seus álbuns, assim como outros dados da companhia em um banco de dados. A companhia decidiu contratar você como o projetista do banco de dados.

1. Cada músico que grava na *Musical* possui um nome, CPF, um endereço e um número de telefone. Músicos mal pagos podem compartilhar o mesmo endereço, e neste caso, não é desejável manter estes dados duplicados. Nenhum endereço possui mais de um telefone.
2. Cada instrumento usado nas canções gravadas na *Musical* tem um número identificador único e um nome (por exemplo, guitarra, sintetizador, flauta).
3. Cada álbum gravado na *Musical* tem um número único de identificação, um título, uma data de copyright e um formato (ex. CD, DVD, LP).
4. Cada música gravada na *Musical* tem um título e um autor.
5. Cada músico pode tocar vários instrumentos, e um dado instrumento pode ser tocado por vários músicos.
6. Cada álbum tem várias músicas, mas nenhuma música pode aparecer em mais de um álbum.
7. Cada música é tocada por um ou mais músicos, e um músico pode tocar em várias músicas.
8. Cada álbum tem exatamente um músico que atua como seu produtor. Um músico pode produzir vários álbuns.

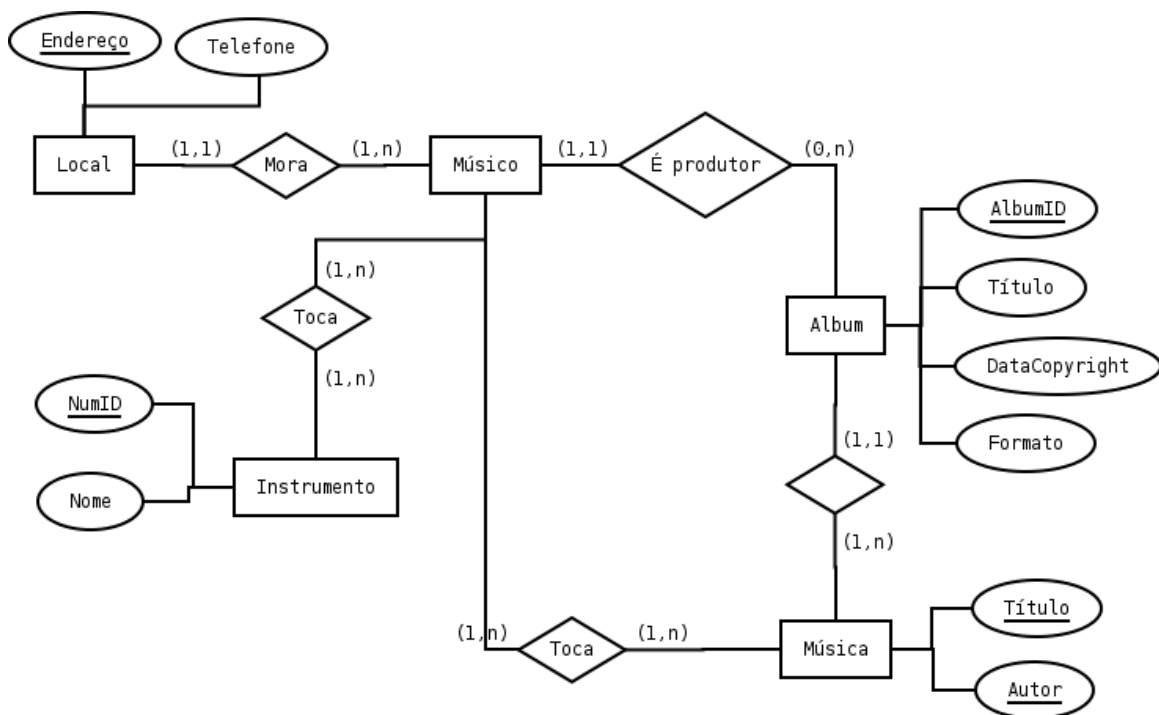
Desenha um diagrama ER que modele as restrições acima. Indique todos os identificadores e cardinalidades. Se houver alguma informação que não pode ser concluída a partir da lista acima, indique explicitamente o que vc assumiu e como isso foi refletido na modelagem (ex: a cardinalidade mínima da entidade X no relacionamento Y foi definida como 1, pois assumi que X era obrigatório no relacionamento Y).

A resposta abaixo foi modelada usando a ferramenta DIA.

Algumas alternativas para esta questão também serão aceitas:

- Colocar um identificador artificial para Música, e não usar Título + Autor como identificador.

- Criar uma entidade para telefona, e relacioná-la com a entidade Local



Questão 6

Os usuários do aeroporto DXX estão reclamando da falta de organização do aeroporto. Por causa disso, os responsáveis pelo aeroporto decidiram que toda a informação relacionada ao aeroporto deve ser armazenada em um banco de dados, e você foi contratado para projetá-lo. Sua primeira tarefa é organizar a informação sobre todos os aviões estacionados e mantidos no aeroporto. A informação relevante para este caso é a seguinte:

- Cada avião tem um número de registro, e um modelo.
- O aeroporto acomoda vários modelos de aeronaves, e cada modelo é identificado pelo número do modelo (ex. DC-10) e possui capacidade e peso.
- Vários técnicos trabalham no aeroporto. Você precisa armazenar nome, CPF, endereço, telefone e salário de cada técnico.
- Cada técnico é especialista em um ou mais modelos de avião, e sua especialidade pode se sobrepor à de outros técnicos.
- Controladores de voo devem fazer um exame médico anual. Para cada controlador de voo, você deve armazenar a data do exame mais recente.
- Todos os empregados do aeroporto (incluindo os técnicos) pertencem a um sindicato. Você deve armazenar o número de filiação de cada empregado a seu sindicato.
- Você pode assumir que cada empregado é identificado unicamente pelo seu CPF.
- O aeroporto possui vários testes que são usados periodicamente para assegurar que os aviões estão em condições seguras de operação. Cada teste tem um número designado pela Administração Federal de Aviação (FAA), um nome, e uma pontuação máxima.

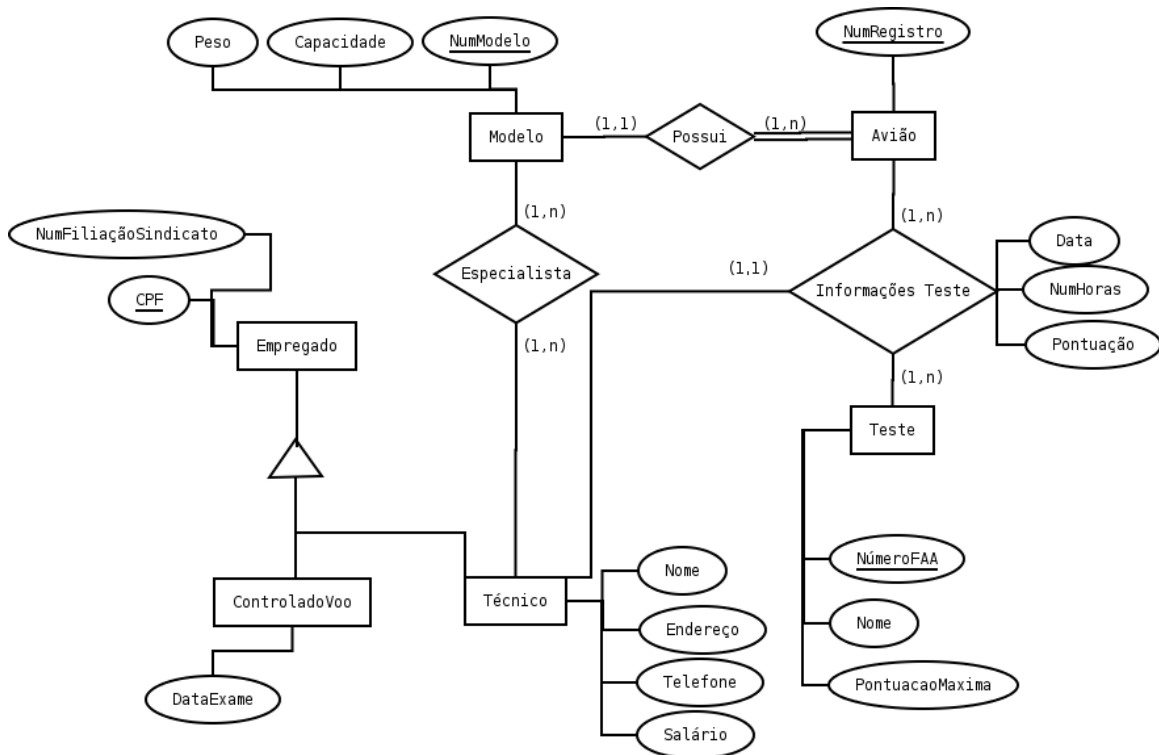
- A FAA exige que o aeroporto registre os testes realizados em um dado avião. As informações necessárias são: qual o teste realizado, qual o técnico que o realizou, a data, o número de horas que o técnico gastou realizando o teste, e a pontuação obtida pelo avião no teste.

Desenhe o diagrama ER que modela as restrições acima. Não esqueça de explicitar os identificadores e cardinalidades dos relacionamentos. Se houver alguma informação que não pode ser concluída a partir da lista acima, indique explicitamente o que vc assumiu e como isso foi refletido na modelagem.

A resposta abaixo foi modelada usando a ferramenta DIA.

Algumas alternativas para esta questão também serão aceitas:

- Colocar nome, endereço, telefone e salário como atributos de empregado ao invés de Técnico



Questão 7

Discuta duas estratégias para verificação de modelos ER quanto à sua corretude.

Um diagrama ER pode conter dois tipos de erros: sintáticos e semânticos. Erros sintáticos são aqueles que acontecem quando as regras de formação do modelo ER não são seguidas. Por exemplo, quando incluímos um relacionamento de uma entidade com um atributo, quando especializamos um relacionamento, ou quando associamos atributos a atributos. Para verificar este tipo de erro, basta checar se existem construções inválidas

no modelo. Erros semânticos acontecem quando um diagrama não expressa a realidade modelada de forma correta. Este tipo de erro é mais difícil de ser verificado. O ideal, nestes casos, é envolver o usuário para que ele ajude na validação.

Questão 8

Explique a diferença entre a independência lógica e física de dados.

A independência lógica de dados significa que os usuários ficam imunes às mudanças na estrutura lógica dos dados, enquanto que a independência física de dados isola os usuários de mudanças no armazenamento físico dos dados.

Foi visto um exemplo de independência lógica de dados na resposta do exercício 7. Considere a relação Aluno daquele exemplo (porém, agora assuma que ela não foi substituída pelas outras duas menores). Poderia ter sido definido que a relação Aluno seria armazenada como um arquivo seqüencial com índice sobre o atributo aid. Outra alternativa seria armazenar Aluno com o agrupamento físico das tuplas baseado num índice agrupado sobre anome ou ainda criar índices sobre aid e cra ou ainda armazená-la de modo ordenado com base no atributo cra. Essas alternativas de armazenamento não são visíveis aos usuários, a menos no que diz respeito ao impacto que causam no desempenho de consultas sobre essa relação. Os usuários apenas vêem a relação como um conjunto de tuplas. Isso é o que significa a independência física dos dados.

Questão 9

O que é independência lógica de dados e por que é importante? Mostre um exemplo.

Independência lógica de dados significa que os usuários ficam imunes às mudanças na estrutura lógica dos dados, por ex., inclusão de novos atributos numa relação.

Exemplo:

Supondo que, por necessidade de alguma aplicação, a relação ALUNO (aid, anome, cra) seja substituída por duas relações, a saber: ALUNO_NOME (aid, anome) e ALUNO_CRA (aid, cra). Os programas de aplicações que operam sobre a relação ALUNO, ficam imunes a essa mudança, por conta de uma definição de visão denominada ALUNO (aid, anome, cra), como sendo a junção natural das relações ALUNO_NOME e ALUNO_CRA.

Assim, os programas de aplicação que se referem à ALUNO não precisam ser modificados quando a relação ALUNO é substituída pelas outras duas relações. A única mudança que ocorre está no escopo do SGBD que ao invés de armazenar e recuperar diretamente as tuplas de ALUNO, fará a obtenção dessas tuplas, quando necessário, usando a definição da visão. Isso fica transparente para o programa de aplicação.

Questão 10

Quais são as funções do Administrador do Banco de Dados?

- Autorizar o acesso à base de dados (segurança e autorização)
- Coordenar e monitorar o uso da base de dados (auxiliar nos procedimentos de reconstrução da base em caso de falhas e manter a base operacional)
- Definir a aquisição de software e recursos de hardware
- Controlar o uso dos recursos e
- Monitorar a eficiência das operações .

Sintonia fina (*database tuning*): O ABD é responsável pela evolução da base de dados, em especial o esquema físico, para adequar o desempenho do SGBD às mudanças de requisitos dos usuários.

Questão 11

Explique a diferença entre os esquemas externo, interno, e lógico.

Como essas camadas de esquemas distintos se relacionam com os conceitos de independência lógica e física de dados?

Esquemas externos permitem que o acesso aos dados seja particularizado (e autorizado) ao contexto individual dos usuários ou grupos de usuários. Esquemas lógicos descrevem todos os dados que estão armazenados numa determinada base de dados. Enquanto existem inúmeras visões para uma mesma base de dados, existe apenas um único esquema lógico para todos os usuários de uma base de dados. Esquemas internos (físicos) definem como são armazenadas (no disco ou em outro meio físico) as relações descritas no esquema lógico.

Esquemas externos viabilizam a independência lógica dos dados, enquanto que o os esquemas lógicos provêem a independência física dos dados.

Questão 12

Quais são as principais vantagens de usar um SGBD ?

Independência de dados e acesso eficiente. Programas de aplicação independem dos detalhes de representação e armazenamento dos dados. Os esquemas lógico e externo provêem independência das decisões quanto ao armazenamento físico e projeto lógico respectivamente. Além disso, o SGBD possui mecanismos eficientes de armazenamento e acesso aos dados, contando com a gerência de arquivos muito grandes, estruturas de índices e otimização de consultas.

Tempo reduzido de desenvolvimento da aplicação. É decorrente dos recursos do SGBD disponíveis para realizar de modo eficiente, diversas funções que teriam que ser codificadas nos programas de aplicação sempre que fossem necessários. Dentre essas funções destacam-se: controle de concorrência, reconstrução em caso de falha, especificação de consultas em linguagens de alto nível, etc. Apenas o código específico da lógica da aplicação deve ser programado. Mesmo assim, existem diversas ferramentas de apoio ao desenvolvimento de aplicações disponíveis na maioria dos fabricantes de SGBDs.

Integridade de dados e segurança. O mecanismo de visão e recursos de autorização do SGBD permitem um controle de acesso aos dados muito poderoso. Além disso, modificações que violam a semântica dos dados podem ser detectadas e descartadas pelo SGBD a partir das especificações das restrições de integridade.

Administração de dados. Como o SGBD oferece um guarda-chuva de gerência de grandes coleções de dados e operações que podem ser compartilhadas por diversos usuários, as tarefas de controle, manutenção e administração desses dados é facilitada. Um bom ABD pode dispensar os usuários das tarefas de sintonia fina da representação dos dados, cópias de segurança periódicas, etc.

Questão 13

Qual a diferença entre redundância de dados controlada e não controlada ?

Num cenário de programas de aplicações que utilizam sistemas de arquivos para gerenciar os dados, inevitavelmente ocorre redundância entre os dados armazenados. Essa redundância é dita não controlada, uma vez que ela ocorre de modo independente e sem um planejamento ou centralização, o que caracteriza a falta de controle da redundância. Já num SGBD, existe uma redundância mínima no armazenamento dos dados usados por programas de aplicações. Em geral, os atributos chave primária e estrangeira aparecem tanto na tabela que referencia quanto na tabela referenciada. Entretanto essa redundância é controlada pelo SGBD, na medida em que esse relacionamento é definido e acompanhado ao longo das operações realizadas no âmbito do SGBD.

Questão 14

Considere a figura 1.2 do livro texto (Elmasri & Navathe), reproduzida a seguir:

- (a) Apresente, de modo informal, em português, cinco consultas (três do tipo leitura e duas do tipo de modificação) envolvendo essas tabelas.
- (b) Identifique três relacionamentos entre essas tabelas.
- (c) Pense em diferentes usuários para essa base de dados. Que tipo de aplicação cada usuário precisaria ? A que categoria cada um desses usuários pertence, e que tipo de interface eles precisariam ?

ALUNO	Nome	Numero	Turma	Curso_Hab
	Smith	17	1	CC
	Brown	8	2	CC

CURSO	NomedoCurso	NumerodoCurso	Creditos	Departamento
	Introdução à Ciência da Computação	CC1310	4	CC
	Estruturas de dados	CC3320	4	CC
	Matemática Discreta	MAT2410	3	MATH
	Banco de dados	CC3380	3	CC

DISCIPLINA	IdentificadordeDisciplina	NumerodoCurso	Semestre	Ano	Instrutor
	85	MAT2410	Segundo Semestre	98	King
	92	CC1310	Segundo Semestre	98	Anderson
	102	CC3320	Primeiro Semestre	99	Knuth
	112	MAT2410	Segundo Semestre	99	Chang
	119	CC1310	Segundo Semestre	99	Anderson
	135	CC3380	Segundo Semestre	99	Stone

HISTORICO_ESCOLAR	NumerodoAluno	Identificador_Disciplinas	Nota
	17	112	B
	17	119	C
	8	85	A
	8	92	A
	8	102	B
	8	135	A

PRE_REQUISITO	NumerodoCurso	NumerodoPre_requisito
	CC3380	CC3320
	CC3380	MAT2410
	CC3320	CC1310

(a) Algumas consultas:

Quais são os cursos do departamento “CC” ?

Quais são os instrutores do curso de “Matemática Discreta” ?

Qual o nome dos alunos que já tiraram A em alguma disciplina ?

Modifique a Nota do aluno “Smith” na disciplina 112 para “C”

Remova o pré-requisito do curso de “Estrutura de Dados”

(b) Alguns relacionamentos entre tabelas:

ALUNO (Numero) e HISTORICO_ESCOLAR (NumerodoAluno)

DISCIPLINA (IdentificadordeDisciplina) e HISTORICO_ESCOLAR

CURSO (NumerodoCurso) e DISCIPLINA

CURSO (NumerodoCurso) e PRE_REQUISITO

(c) Classes de usuários (categoria, aplicação, interface):

- a. Projetista da Base de Dados – definir a estrutura da tabela ALUNO, interface de gerência do esquema lógico.

- b. ABD – otimização de desempenho das consultas que buscam determinados conceitos, interface de monitoramento de execução.
- c. Usuário Final – modificar as notas dos alunos, interface de consulta à base de dados.
- d. ABD – autorização de usuários que podem modificar as notas dos alunos, interface de gerência de usuários