

Banco de Dados I

Normalização
Steven Louback

Normalização

- O processo de normalização pode ser visto como um processo pelo qual um esquema de tabelas (relações) insatisfatório é quebrado de forma que seus atributos formem relações menores que sejam mais adequadas.
- O objetivo é evitar que as anomalias de atualização ocorram.

Normalização

- A normalização serve para analisar tabelas e organizá-las de forma que a sua estrutura seja simples, relacional e estável, para que o gerenciamento possa ser também simples, eficiente e seguro.
- Os objetivos são evitar a perda e a repetição da informação e atingir uma forma de representação adequada para o que se deseja armazenar.
- Oferece mecanismos para analisar o projeto do BD e a identificação de erros.

Normalização

- Oferece métodos para corrigir problemas.
- A normalização é a explicação formal para coisas que são óbvias ao se projetar um Banco de Dados.
 - A intuição do projetista o leva a dividir um esquema de Banco de Dados em várias tabelas, mas a normalização explica porque as coisas são assim.

Normalização

- Um bom projeto de um esquema de entidades, e sua consequente conversão para um esquema relacional, segundo as regras vistas, praticamente deixa o esquema relacional normalizado.
- Assim, utiliza-se a normalização somente para *validar um projeto relacional*.
- A normalização é necessária (embora não suficiente) a um bom projeto relacional.

Normalização

- Principais formas normais existentes
 - 1FN
 - 2FN
 - 3FN
 - Forma normal de Boyce/Codd (FNBC) ou Nova 3FN.
 - 4FN
 - 5FN

Normalização

- Erros encontrados em projetos de Banco de Dados:
 - Repetição de informação;
 - Perda de informações;
 - Inabilidade de representar certas informações;

Normalização

Nro_ped	Data	Nro_Peca	Descrição	Qt_comprada	Preço
1000	10/02/2010	AX12	Bicicleta	2	1000
1010	10/02/2010	BT04	TV	3	5000
1020	12/02/2010	CD09	DVD	4	600
1030	12/02/2010	EF12	Aparelho de som	3	800
1040	18/02/2010	BT04	TV	4	7000
1050	18/02/2010	CD09	Home Theater	5	8000
1060	20/02/2010	BT04	TV	2	4000

Normalização

- Problemas apresentados (anomalias)
 - Inconsistência: não há nada no projeto impedindo que a peça CD09 tenha duas ou mais descrições diferentes no Banco de Dados.
 - Redundância de dados: A descrição de um mesmo produto encontra-se em diversas linhas.
 - Anomalia de modificação: uma mudança na descrição da peça BT04 requer várias mudanças
 - Anomalia de inclusão : não se pode inserir um novo produto sem que esteja associado a um pedido.
 - Anomalia de exclusão: Se for necessário excluir um pedido as informações do produto que estiver associado somente a esse pedido serão perdidas.

Normalização

Solicitação de Material			
Número da Solicitação:			Data da Solicitação
Código Funcionário		Nome	
Cod.Prod.	Descrição	Quantidade	

Normalização

- Tabela não normalizada

Número da Solicitação, Data da Solicitação,
Código do Funcionário, Nome do Funcionário,
(Cód. produto, Descrição, Quantidade).

Normalização

Primeira Forma Normal (1 FN)

- Uma relação está na 1 FN se, e apenas se, todos os domínios contiverem apenas valores atômicos.
- Uma tabela na 1FN não permite atributos multivalorados ou atributos compostos.

Normalização

Para tornar uma relação não-normalizada em uma relação normalizada existem duas alternativas:

- 1. Construir uma única tabela com redundância de dados**
- 2. Para cada tabela aninhada criar uma nova tabela. Esta nova tabela irá conter os atributos da tabela aninhada e o(s) atributo(s) que compõe a chave primária da(s) tabela(s) na(s) qual(is) a tabela está aninhada.**

- Tabela com redundância de dados

Número da Solicitação, Data da Solicitação, Código do Funcionário,
Nome do Funcionário, Cód. produto, Descrição, Quantidade

Normalização

- Com tabelas aninhadas

Tabela Solicitação

Número da Solicitação, Data da Solicitação, Código do Funcionário, Nome do Funcionário

Tabela Item

Número da Solicitação, Cód. produto, Descrição, Quantidade

Normalização

- **Dependências Funcionais**
- Existe uma dependência funcional entre atributos de uma relação, quando a um valor de um atributo X está sempre associado um mesmo valor para um atributo Y . Assim, se um valor V_x ocorrer em linhas diferentes da relação R , o valor V_y será o mesmo em todas as linhas onde V_x ocorrer.

Normalização

- Pode-se então afirmar que:

$$X \rightarrow Y$$

X determina Y ou Y é dependente funcional de X ou ainda X é o determinante de Y

Exemplo: Cód. Produto \rightarrow Descrição

- Alguns atributos podem depender funcionalmente de um conjunto de atributos e não apenas de um único.

Exemplo: (Número da Solicitação , Cód. produto) \rightarrow Quantidade

Normalização

Segunda Forma Normal (2 FN)

- Uma relação está na segunda forma normal se e apenas se, estiver na 1FN, e cada atributo não-chave for totalmente dependente da chave primária.

Normalização

Para chegar a 2FN deve-se:

1. Verificar se existem colunas não-chave parcialmente dependentes de algum dos atributos da chave. Esta verificação é feita apenas em tabelas que tem a sua chave primária composta por mais de um atributo
2. Se existir um atributo que dependa apenas parcialmente da chave primária:
 - i. Criar tabela (se não existe ainda) cuja chave primária será a parte da chave primária que é determinante do atributo analisado
 - ii. Mover da tabela original para a tabela criada o atributo dependente parcialmente da chave (Mover é *diferente* de copiar...)

Normalização

Assim, considerando cada uma das duas formas utilizadas para gerar as tabelas na 1
FN:

1 - Construir uma única tabela com redundância de dados

1 FN

Número da Solicitação, Data da Solicitação, Código do Funcionário, Nome do Funcionário, Cód. produto, Descrição, Quantidade

2FN

Número da Solicitação, Data da Solicitação, Código do Funcionário, Nome do Funcionário

Número da Solicitação, Cód. produto, Quantidade

Cód. produto, Descrição

Normalização

2 – Com tabelas aninhadas

1FN

Número da Solicitação, Data da Solicitação, Código do Funcionário, Nome do Funcionário

Número da Solicitação, Cód. produto, Descrição, Quantidade

2FN

Número da Solicitação, Data da Solicitação, Código do Funcionário, Nome do Funcionário

Número da Solicitação, Cód. produto, Quantidade

Cód. produto, Descrição

Normalização

Terceira Forma Normal (3FN)

- Uma relação está na terceira forma normal se e apenas se, estiver na 2FN, e não tiver dependências transitivas.

Normalização

- *Dependência transitiva*: ocorre quando um atributo não-chave, além de depender da chave primária da tabela, depende funcionalmente de outro atributo ou combinação de atributos não-chave. Assim, em uma tabela na 3FN não existem atributos não-chave que tenham dependência de outros atributos não chave.

Normalização

Para chegar a 3FN deve-se:

1. Verificar se existem colunas não-chave dependentes de algum outro atributo ou conjunto de atributos não-chave. Esta verificação é feita apenas em tabelas que possuírem pelo menos 2 atributos não-chave.
2. Se existir um atributo que dependa de outro atributo (ou conjunto de atributos) não-chave:
 - i. Criar tabela (se não existe ainda) cuja chave primária será o atributo ou conjunto de atributos que determinam o valor do atributo não-chave
 - ii. Mover da tabela original para a tabela criada o atributo dependente

Normalização

2FN

Número da Solicitação, Data da Solicitação, Código do Funcionário,
Nome do Funcionário

Número da Solicitação, Cód. produto, Quantidade

Cód. produto, Descrição

3FN

Número da Solicitação, Data da Solicitação, Código do Funcionário

Código do Funcionário, Nome do Funcionário

Número da Solicitação, Cód. produto, Quantidade

Cód. produto, Descrição

Exercícios

Aplicar as Formas Normais cabíveis, nas questões abaixo.

Você deve transformar os esquemas abaixo em conjuntos de esquemas e justificar sua normalização de acordo com as regras de normalização.

- 1) Empregado (Número Empregado, Nome do Empregado, Número do Departamento, Nome do Departamento, Número do Gerente, Nome do Gerente, Número do Projeto, Nome do Projeto, Dia de Início do Projeto, Número de horas trabalhadas no projeto).

Exercícios

- 2) Ordem_Compra (cd_ordem_compra, dt_emissão, cd_fornecedor, nm_fornecedor, endereço_fornecedor, cd_material (n vezes), descrição_material (n vezes), qt_comprada (nvezes), vl_unitário (n vezes), vl_total_item (n vezes), vl_total_ordem).
- 3) Tabela de Notas Fiscais (Num_NF, Série, Data emissão, Cod. Cliente, Nome cliente, Endereço cliente, CGC cliente, Código Mercadoria, Descrição Mercadoria, Quantidade vendida, Preço de venda, Total da venda da Mercadoria e Total Geral da Nota). Cada nota pode ter mais do que uma mercadoria.