

Banco de Dados I

Normalização de dados

Márcio Daniel Puntel
marcio.puntel@ulbra.edu.br

Normalização de dados

- Normalização é um processo onde uma tabela relacional não normalizada é transformada em um conjunto de tabelas normalizadas para deixar a modelagem em conformidade com a realidade.

Processo de Normalização

- Etapas do processo de normalização :
 - O documento é representado na forma de uma tabela não normalizada.
 - O documento vai sendo decomposto em tabelas normalizadas (corretas). A normalização se dá, normalmente, em três passos.
 - Uma forma normal é um conjunto de regras que uma tabela deve obedecer. Estas regras destinam-se a eliminar as redundâncias de dados.

Documento NN

Exemplo de documento a normalizar: Relatório de Alocação a Projeto

CÓDIGO DO PROJETO: LSC001 TIPO: Novo Desenv. DESCRIÇÃO: Sist. Estoque.

NOEMP	NOME	CATEG	SALÁRIO	INÍCIO	TEMPO
				PROJETO	ALOCÇÃO
2146	João	A1	40	01/11/91	24
3145	Sílvia	A2	40	02/10/91	24
6126	José	B1	90	03/10/92	18
1181	Carlos	A2	40	01/11/92	12

CÓDIGO DO PROJETO: PAG02 TIPO: Manutenção DESCRIÇÃO: Sistema de RH.

NOEMP	NOME	CATEG	SALÁRIO	INÍCIO	TEMPO
				PROJETO	ALOCÇÃO
1181	Carlos	A2	40	01/11/93	14
5672	Luís	A1	40	12/10/91	24
6126	José	B1	90	31/11/92	11

Forma Não Normalizada

- Uma tabela não normalizada (NN) contém valores de atributos não atômicos, isto é, contém tabelas embutidas (grupos repetidos, arrays).
- Representação não normalizada do documento exemplo:
PROJ (CODPROJ, TIPOPROJ, DESCR,
 (NOEMP, NOME, CAT, SAL,
DATAINÍCIO,TEMPOALOC))
- Deve-se observar a representação do embutimento de tabelas através de parênteses e a indicação das chaves primárias em cada nível de embutimento.

Primeira Forma Normal

- Uma tabela na primeira forma normal (PFN) não contém tabelas embutidas.
- A passagem à primeira forma normal é feita nos seguintes passos:
- 1. Para cada tabela embutida (cada uma abre parênteses), inclusive a mais externa, é criada uma tabela na PFN que contém:
 - as chaves primárias de cada tabela externa à tabela embutida;
 - os atributos da própria tabela embutida.
- 2. São definidas as chaves primárias das tabelas na PFN.

Primeira Forma Normal

Passagem à PFN - decomposição de tabelas

- Tabela1: corresponde ao nível externo
PROJ (CODPROJ, TIPOPROJ, DESCR)
- Tabela2: corresponde à tabela embutida
PROJEMP (CODPROJ, NOEMP, NOME, CAT, SAL,
DATAINÍCIO, TEMPOALOC)

Primeira Forma Normal

Passagem à PFN - identificação de chaves

- Tabela1: corresponde ao nível externo

PROJ (CODPROJ, TIPOPROJ, DESCR)

- Tabela2: corresponde à tabela embutida

PROJEMP (CODPROJ, NOEMP, NOME, CAT, SAL,
DATAINÍCIO, TEMPOALOC)

Segunda Forma Normal

- Uma tabela está na segunda forma normal (SFN) quando, além de estar na PFN, cada atributo não chave primária depende funcionalmente de toda a chave primária e não de apenas parte dela.
- Ao passar uma tabela para a SFN é necessário considerar apenas tabelas que tenham:
 - chave primária composta;
 - pelo menos um atributo não chave.

Segunda Forma Normal

Para o caso do exemplo:

PROJ (CODPROJ, TIPOPROJ, DESCR)

PROJEMP (CODPROJ, NOEMP,
DATAINÍCIO,TEMPOALOC)

EMP (NOEMP, NOME, CAT, SAL)

Terceira Forma Normal

- Uma tabela está na terceira forma normal (TFN) quando, além de estar na SFN, cada atributo não chave primária depende diretamente da chave primária, isto é, não há dependências entre atributos não chave;
- Na passagem à TFN, basta considerar tabelas com mais de um atributo não chave.

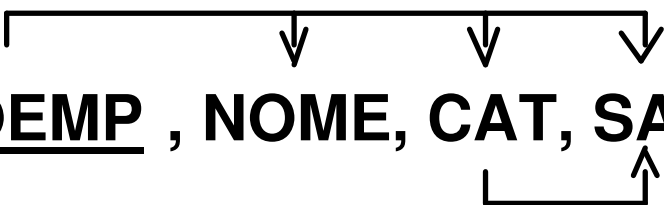
Terceira Forma Normal

Para o caso do exemplo:

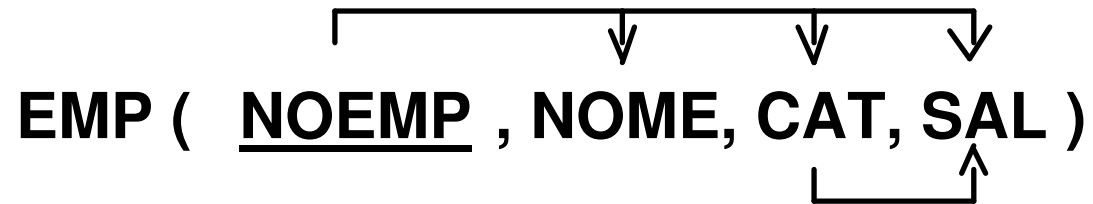
PROJ (CODPROJ, TIPOPROJ, DESCR)

PROJEMP (CODPROJ, NOEMP,
DATAINÍCIO,TEMPOALOC)

EMP (NOEMP , NOME, CAT, SAL)



Terceira Forma Normal



EMP é subdividida em duas:

EMP (NOEMP, NOME, CAT)

CATEGORIA (CAT, SAL)

Forma Normal

Documento normalizado:

PROJ (CODPROJ, TIPOPROJ, DESCR)

PROJEMP (CODPROJ, NOEMP,
DATAINÍCIO,TEMPOALOC)

EMP (NOEMP, NOME, CAT)

CATEGORIA (CAT, SAL)

Forma Normal

- Exemplo genérico de Passagem à PFN

Tabela NN:

(A1, A2, A3, A4, A5
 (B1, B2, B3, B4
 (C1, C2, C3)
 (D1, D2))
 (E1, E2, E3))

Forma Normal

PFN - Subdivisão de Tabelas

1. (A1, A2, A3, A4, A5)
2. (A1, A2, B1, B2, B3, B4)
3. (A1, A2, B1, C1, C2, C3)
4. (A1, A2, B1, D1, D2)
5. (A1, A2, E1, E2, E3)

Normalizar

"Uma fábrica de móveis vai informatizar os pedidos de produtos especificados por funcionários para revendedores. Cada pedido contém um número, data, prazo de entrega, nome do cliente (revendedor), endereço, cidade, U.F., CGC e inscrição estadual do revendedor. Além disso, um pedido contém vários itens de pedido. Cada item de pedido registra o código do produto, descrição, setor, quantidade, valor unitário, desconto e subtotal."

Normalizar

- “Uma empresa de ônibus possui um número, nome (empresa), endereço e nome do gerente. Cada empresa de ônibus possui vários ônibus e motoristas. Um ônibus pertence a somente uma empresa e um motorista pode trabalhar em mais de uma empresa de ônibus. Um ônibus possui um número, placa, ano de fabricação, tipo e capacidade (número de passageiros de acordo com o tipo do ônibus). Um motorista possui um número (especificado pela Secretaria Municipal de Transportes), nome e tempo de experiência. Uma linha de ônibus contém um número, itinerário, é executada por vários ônibus, pode ser percorrida várias vezes (cada uma delas com um horário de partida e de chegada) ao dia em sentidos opostos ou no mesmo sentido. Por exemplo, a linha 31 sai às 8 horas no sentido centro/bairro e também às 8 horas no sentido bairro/centro.”

Referências

- <http://www.slideshare.net/adorepump/normalizao-de-banco-de-dados-presentation>
- KORTH, H. F.; SILBERSCHATZ, A.. **Sistema de Banco de Dados**. Rio de Janeiro: Elsevier, 2006.
- DATE, C. J.. **Introdução a Sistemas de Banco de Dados**. Rio de Janeiro: Elsevier, 2004.