

Normalização de Relações

Normalização

- O que é ?
 - Processo para produzir um projeto lógico a partir do conhecimento de uma implementação existente em algum meio físico tal como um arquivo, formulário e relatório.
- Como é feito?
 - É um processo formal que avalia cada relação sob os critérios de cada forma normal e as decompõe, caso seja necessário.
- Objetivo?
 - O objetivo a ser alcançado é substituir um conjunto de relações por outro minimizando redundâncias e anomalias de atualização (inclusão, alteração e exclusão).

Normalização

- Os esquemas de relações que não atendam a certas condições (definidas pelas formas normais) são decompostos em esquemas de relações menores que possuem as propriedades desejadas.
- O procedimento de normalização proporciona ao projetista do banco de dados:
 - Uma estrutura formal para a análise de esquemas de relações com base em suas chaves e nas dependências funcionais de seus atributos.
 - Uma série de testes para garantir que o esquema do banco de dados atinja o grau de normalização desejado.

Normalização

- Embora existam varias formas normais de alto grau, como as 4FN e 5FN, a utilidade prática destas formas normais é questionável devido as restrições as quais são baseadas serem difíceis de entendidas e detectadas pelos projetistas e usuários.
- Em termos práticos, o usual é dar atenção a normalização até as formas normais 3FN ou BCNF.

Normalização

- **1ª FORMA NORMAL (1FN):** É considerada parte da definição formal de uma relação no modelo relacional. Uma relação está na Primeira Forma Normal (1FN) se todos os atributos contiverem valores atômicos e não existirem tuplas duplicadas.
- Ou seja, a 1FN impede que haja ‘relações dentro de relações’ e atributos multivalorados, como por exemplo, um vetor.

Primeira Forma Normal

- Considere o esquema de tabela **INSCRIÇÃO**, em que um sócio pode estar matriculado em várias atividades esportivas.
 - **INSCRIÇÃO** (matricula, nome, {atividade (modalidade, preço)}),
 - {} indica que o atributo atividade é multivalorado e composto.

INSCRIÇÃO

Matricula	Nome	Atividade	
		Modalidade	Preço
1267	João Ribeiro	natação	150,00
		musculação	120,00
		pilates	150,00
6590	Maria Ramos	natação	150,00
		pilates	150,00
8786	Antonio Vieira	Judô	80,00
0324	Ana Rocha	musculação	120,00
		pilates	150,00

Primeira Forma Normal

A tabela **INSCRIÇÃO** não está na **1FN**, pois as colunas **modalidade** e **preço** não possuem valores atômicos.

Processo para obter a 1FN:

- Eliminação dos atributos multivalorados.
- Determinação da nova chave primária: **Pk (matricula, modalidade)**

Anomalias Existentes:

1. Redundância.
2. Se Antonio Vieira deixar de ser sócio, perde-se a informação que judô custa 80,00.
3. Se quisermos incluir uma nova modalidade, por exemplo, hidroginástica, é preciso também incluir algum sócio para a modalidade.

Matricula	Nome	Modalidade	Preço
1267	João Ribeiro	Natação	150,00
1267	João Ribeiro	Musculação	120,00
1267	João Ribeiro	Pilates	150,00
6590	Maria Ramos	Natação	150,00
6590	Maria Ramos	Pilates	150,00
8786	Antonio Vieira	Judô	80,00
0324	Ana Rocha	Musculação	120,00
0324	Ana Rocha	Pilates	150,00

Segunda Forma Normal

Uma relação está na Segunda Forma Normal (2FN) se está na Primeira Forma Normal e não existir dependência funcional parcial entre a chave e os demais atributos.

Dependência Funcional:

É um relacionamento entre pelo menos dois atributos. Por exemplo, dado uma relação R (A, B, C, D), onde A, B, C, D são atributos, ocorre uma dependência funcional entre atributos de R se conhecendo o valor de um atributo, consegue-se determinar um único valor de outro atributo.

Notação: $A \rightarrow C$

Dependência Funcional

Exemplo:

Departamento

Código	Nome	Cidade
10	Jurídico	Rio de Janeiro
20	Pessoal	Rio de Janeiro
30	Financeiro	São Paulo

Dependências Funcionais:

Código -> Nome Nome -> Código
Código -> Cidade Nome -> Cidade

Dependência Funcional

OBS1:

Se $X \rightarrow (Y, Z)$, então $X \rightarrow Y$ e $X \rightarrow Z$

Se $(X, Y) \rightarrow Z$ não implica em $X \rightarrow Z$ ou $Y \rightarrow Z$.

OBS2:

Uma coluna depende parcialmente da chave se para que seu valor seja determinado não é necessário conhecer a chave como um todo, mas sim somente uma parte da chave.

OBS3:

Se a chave da relação é definida por somente uma coluna, então a relação está na 2FN.

Segunda Forma Normal

Processo para obtenção da 2FN:

1 – Para as colunas que não fazem parte da chave analisar se o seu valor é determinado por parte ou totalidade da chave.

INSCRIÇÃO

Matricula	Nome	Modalidade	Preço
1267	João Ribeiro	Natação	150,00
1267	João Ribeiro	Musculação	120,00
1267	João Ribeiro	Pilates	150,00
6590	Maria Ramos	Natação	150,00
6590	Maria Ramos	Pilates	150,00
8786	Antonio Vieira	Judô	80,00
0324	Ana Rocha	Musculação	120,00
0324	Ana Rocha	Pilates	150,00

Análise das dependências funcionais:

Matricula -> nome

modalidade-> preço

Segunda Forma Normal

2 – As colunas que são dependentes funcionais parciais da **Pk** formarão novas tabelas cuja chave primária será à parte da chave de qual dependem.

SÓCIO

matricula	nome
1267	João Ribeiro
6590	Maria Ramos
8786	Antônio Vieira
0324	Ana Rocha

Pk (matricula)

ATIVIDADE

modalidade	preço
Natação	150,00
Musculação	120,00
Pilates	150,00
Judô	80,00

Pk (modalidade)

Segunda Forma Normal

3 – As colunas com dependência funcional parcial são removidas da relação original

INSCRIÇÃO

Matricula	Modalidade
1267	Natação
1267	Musculação
1267	Pilates
6590	Natação
6590	Pilates
8786	Judô
0324	Musculação
0324	Pilates

Esquema final na 2FN:

Inscrição (matricula, modalidade)

Sócio (matricula, nome)

Atividade (modalidade, preço)

Obs: Anomalias Anteriormente

Citadas Resolvidas

Pk (matricula, modalidade)

Terceira Forma Normal

Uma relação está na **Terceira Forma Normal** se está na **Segunda Forma Normal**, e não existir dependência funcional entre atributos que não pertencem chave.

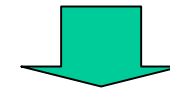
Exemplo: Um cliente mora em um único edifício e um edifício tem uma única taxa.

ADMINISTRAÇÃO

Cliente	Edifício	Taxa
100	Barão de Mauá	500,00
200	Cidade	800,00
300	Metropolitan	1.000,00
400	Cidade	800,00

Pk (Cliente)

Análise das Dependências funcionais:
Cliente -> Edifício, Cliente -> Taxa
Edifício -> Taxa



Logo a relação está na **2FN** mas **não** está na **3FN**

Terceira Forma Normal

Uma relação está na **Terceira Forma Normal** se está na **Segunda Forma Normal**, e não existir dependência funcional entre atributos que não pertencem chave.

Exemplo: Um cliente mora em um único edifício e um edifício tem uma única taxa.

ADMINISTRAÇÃO

Cliente	Edifício	Taxa
100	Barão de Mauá	500,00
200	Cidade	800,00
300	Metropolitan	1.000,00
400	Cidade	800,00

Pk (Cliente)

Anomalias:

Se o cliente 300 for excluído, o valor da taxa para o edifício “Metropolitan” é perdido

Se a Administradora quiser incluir o valor da taxa para um novo edifício, ela não poderá fazer até que surja um cliente para o edifício.

Terceira Forma Normal

Processo para obtenção da 3FN

- 1 - Identificar as colunas que não fazem parte da chave primária.
- 2 - Analisar se existe dependência funcional entre estes atributos.
- 3- As colunas que têm seu valor determinado por uma coluna não chave, formarão novas tabelas cujas chaves primárias serão as colunas que as determinam.

Administração

Cliente	Edifício
100	Barão de Mauá
200	Cidade
300	Metropolitan
400	Cidade

PK (Cliente)

Condomínio

Edifício	Taxa
Barão de Mauá	500,00
Metropolitan	1.000,00
Cidade	800,00

PK (Edifício)

Esquema final na 3FN:

Administração (cliente, edifício)

Condomínio (edifício, taxa)

OBS: ANOMALIAS ANTERIORMENTE CITADAS RESOLVIDAS