



Bases de Dados

Normalização - Parte 2

Profa. Elaine Parros Machado de Sousa

Normalização

- Formas Normais baseadas em dependências funcionais
 - baseadas em chave primária
 - 1ª FN
 - 2ª FN
 - 3ª FN
 - baseadas em chaves candidatas
 - definições genéricas de 2ª FN e 3ª FN
 - FN de Boyce-Codd (BCNF)
- Forma Normal baseada em dependências multivaloradas
 - 4ª FN

Parte 1

Parte 2

Normalização

- Formas Normais baseadas em dependências funcionais
 - baseadas em chave primária
 - 1ª FN
 - 2ª FN
 - 3ª FN
 - baseadas em chaves candidatas
 - definições genéricas de 2ª FN e 3ª FN
 - FN de Boyce-Codd (BCNF)
- Forma Normal baseada em dependências multivaloradas
 - 4ª FN

Normalização

- **Definições genéricas** \Rightarrow levam todas as chaves candidatas em consideração
 - **2ª FN genérica** \Rightarrow todos os atributos **não primários** possuem **dependência total**, transitiva ou não, de **todas as chaves** (primária, secundária, ...)
 - **3ª FN genérica** \Rightarrow todos os atributos **não primários** possuem **dependência total, não transitiva**, de **todas as chaves** (primária, secundária, ...)

Exemplo

Matrícula → Municipio, NroLote, Área, Preço

Municipio, NroLote → Área, Matrícula, Preço

Área → Preço

Lotes = {Matricula, Municipio, NroLote, Área, Preço }

➤ Analise a relação **Lotes** considerando as definições genéricas de 2ª FN e 3ª FN

- 2ª FN genérica? **OK!!**
- 3ª FN genérica? **NÃO!!**

Exemplo

Matrícula \rightarrow Municipio, NroLote, Área, Preço

Municipio, NroLote \rightarrow Área, Matrícula, Preço

Área \rightarrow Preço

Lotes = {Matricula, Municipio, NroLote, Área, Preço }

➤ **Normalizando para 3ª FN Genérica**

Lotes = {Matricula, Municipio, NroLote, Área}

Área = {Área, Preço}



Exemplo

Aluno, Curso \rightarrow Instrutor

Instrutor \rightarrow Curso

Treinamento = {Aluno, Curso, Instrutor}

➤ Analisando a relação **Treinamento** considerando as definições genéricas de 2ª FN e 3ª FN...

- 2ª FN genérica? **OK!!**
- 3ª FN genérica? **OK!!**

Mas... e a DF Instrutor \rightarrow Curso?

Forma Normal de Boyce - Codd (BCNF)

- **BCNF** \Rightarrow extensão da 3ª FN genérica
- uma relação R **está** na **BCNF** se:
 - estiver na 3ª FN genérica
 - **para toda DF não-trivial $X \rightarrow A$** válida para a relação R, **X** é uma **chave** em R

Forma Normal de Boyce - Codd (BCNF)

- Na prática:
 - maioria das relações em 3FN genérica também está na BCNF
 - Exceção:
 - quando $X \rightarrow A$ e:
 - X não é chave
 - A é atributo primário

Exemplo

Aluno, Curso \rightarrow Instrutor

Instrutor \rightarrow Curso

Treinamento = {Aluno, Curso, Instrutor}

- A relação **Treinamento** está na BCNF?
- Alternativas de normalização por decomposição?

Exemplo

Aluno, Curso \rightarrow Instrutor

Instrutor \rightarrow Curso

Treinamento = {Aluno, Curso, Instrutor}

➤ Alternativas de decomposição:

- 1) {Aluno, Instrutor} e {Aluno, Curso}
- 2) {Curso, Instrutor} e {Curso, Aluno}
- 3) {Instrutor, Aluno} e {Instrutor, Curso}



**Quais os problemas de cada alternativa?
Qual a melhor opção? Por que?**

Exemplo

Aluno, Curso \rightarrow Instrutor


Instrutor \rightarrow Curso

Treinamento = {Instrutor, Aluno}



Instrutor = {Instrutor, Curso}

**Mas essa solução está realmente boa?
Garantindo consistência dos dados
(considerando as DFs)?**

- 
- OBS: Na melhor solução anterior ainda pode acontecer aluno (Zé) fazer mesmo curso com instrutores diferentes:
 - > Ana, Zé - Ana, C++
 - > João, Zé - João, C++

Solução alternativa: João Marcos Della Torre Divino (Eng Comp 2017)

Normalização

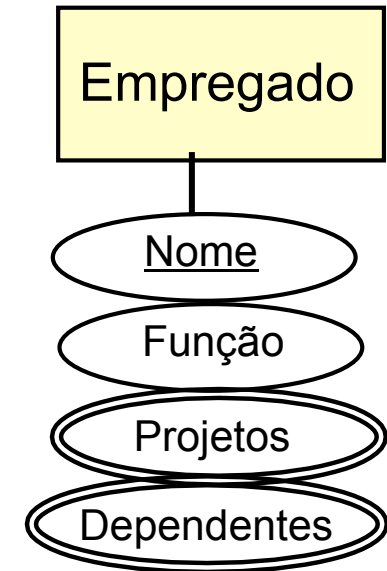
- Formas Normais baseadas em dependências funcionais
 - baseadas em chave primária
 - 1ª FN
 - 2ª FN
 - 3ª FN
 - baseadas em chaves candidatas
 - definições genéricas de 2ª FN e 3ª FN
 - FN de Boyce-Codd (BCNF)
- Forma Normal baseada em **dependências multivaloradas**
 - 4ª FN

Exemplo (simplificado...)

- Dados sobre empregados na base de dados de uma empresa:
 - nome e função do empregado
 - nomes dos projetos em que trabalha
 - nomes de seus dependentes
- dependências funcionais???
- Nome \rightarrow Função
- Projeto e Dependente?

Empregado = {Nome, Função}

Empregado_PD = {Nome, Projeto, Dependente}



Exemplo (cont.)

Empregado = {Nome, Função}

Empregado_PD = {Nome, Projeto, Dependente}

- semanticamente:
 - um **conjunto** de valores de **Projeto** é determinado por **um** valor de **Nome** de empregado, e somente por **Nome**
 - idem para **Dependente**
 - **Projeto** e **Dependente** não têm relação de dependência...



Exemplo de Cenário de Dependência Multivalorada

Dependência Multivalorada

- **Dependência Multivalorada (DM):** restrição entre dois conjuntos de atributos

$$\mathbf{A -\>> B}$$

- **A multidetermina B** (ou B é multidependente de A) \Rightarrow o **conjunto de valores** de B é determinado pelo valor de A, e **somente** pelo valor de A

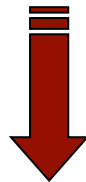
Dependência Multivalorada

- DM são semânticas (assim como as DF)
- **Simetria** na definição de DM:
 - se $X \twoheadrightarrow Y$, então

$$X \twoheadrightarrow Z, Z = R - (X \cup Y)$$

Empregado_PD = {Nome, Projeto, Dependente}

Nome \twoheadrightarrow Projeto



Nome \twoheadrightarrow Dependente

Dependência Multivalorada

- Dada uma DM $X \twoheadrightarrow Y$ em R

- se:

- (a) $Y \subseteq X$ ou

- (b) $X \cup Y = R$



**Dependência
Multivalorada Trivial**

- caso contrário



**Dependência
Multivalorada Não-Trivial**

Exemplos

- DM Trivial (DMT)

Nome -» Projeto

Empregado={Nome, Projeto}

- DM Não-Trivial (DMNT)

Nome -» Projeto

Nome -» Dependente

Empregado={Nome, Projeto, Dependente}

Exemplo

Empregado_PD = {Nome, Projeto, Dependente}

Nome -» Projeto

Nome -» Dependente

➤ **Dados:**

- Ana trabalha nos projetos Museu Virtual e Cidadania, e tem 2 dependentes: Paulo e Sônia;

➤ Como **armazenar** os dados na tabela Empregado_PD?

Exemplo (cont.)

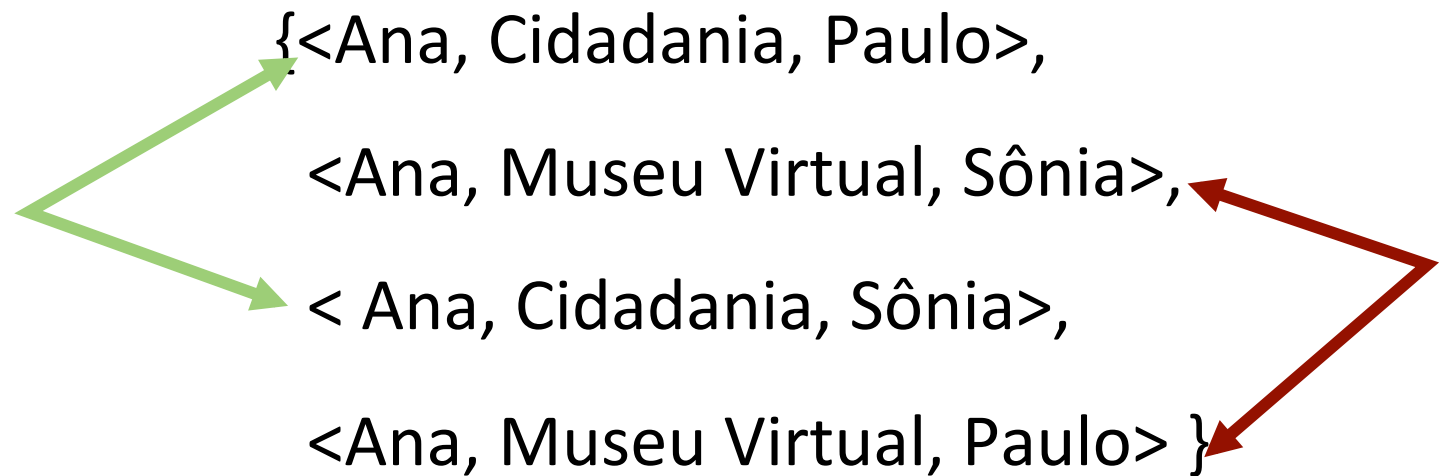
- Ana trabalha nos projetos Museu Virtual e Cidadania, e tem 2 dependentes: Paulo e Sônia;

Empregado_PD = {Nome, Projeto, Dependente}
{<Ana, Cidadania, Paulo>,
<Ana, Museu Virtual, Sônia>}

Anomalias?

Para evitar anomalias de remoção e perda de dados...combinação????

Empregado_PD = {Nome, Projeto, Dependente}




Ruim!!!! Redundância!!!

Dependência Multivalorada

- **Dependência Multivalorada Não Trivial**
 $X \twoheadrightarrow Y$ definida para uma relação **R**:

- se existe $t_1[X] = t_2[X]$, então também existem:

- $t_3[X] = t_4[X] = t_1[X] = t_2[X]$
- $t_3[Y] = t_1[Y]$ e $t_4[Y] = t_2[Y]$
- $t_3[Z] = t_2[Z]$ e $t_4[Z] = t_1[Z]$,
 - $Z = R - (X \cup Y)$



	X	Y	Z
	{ <u>Nome</u> , Projeto, Dependente}		
t_1	<Ana, Cidadania, Paulo>		
t_2	<Ana, Museu Virtual, Sônia>		
t_3	<Ana, Cidadania, Sônia>		
t_4	<Ana, Museu Virtual, Paulo>		

Nome -» Projeto

Nome -» Dependente

Empregado = {Nome, Projeto, Dependente}

{<Ana, Cidadania, Paulo>,

<Ana, Museu Virtual, Sônia>,


< Ana, Cidadania, Sônia>,

<Ana, Museu Virtual, Paulo> }

⇒ Dados ok em relação às DMs....

⇒ Mas... muita redundância!!!

Dependência Multivalorada

- Problema da **DM Não-Trivial**:
 - requer **redundância** nas tuplas
 - Exemplo:
 - Empregado={Nome, Projeto, Dependente}
 - está na BCNF, mas ainda há redundância de dados
- Solução  **4ª FN!**

4ª Forma Normal

- Relação R está na 4ª Forma Normal se:
 - todas as dependências multivaloradas são **triviais** ou
 - para cada dependência multivalorada não-trivial $A \twoheadrightarrow B$, **A é uma chave** (completa) em R

4ª Forma Normal

■ Exemplos...

- Empregado_D={Nome, Dependente}

OK!!

Nome -» Dependente

- Empregado_PD={Nome, Projeto, Dependente}

Nome -» Projeto

Nome -» Dependente

Não!!

4ª Forma Normal

- Intuição:

Nome -» Projeto

Nome -» Dependente

~~Empregado_PD = {Nome, Projeto, Dependente}~~



Dependentes = {Nome, Dependente}

Projetos = {Nome, Projeto}

4ª Forma Normal

- Normalizando a relação para a 4ª FN....

➤ dada uma DMNT **A -» B** na relação R,
substitui-se R por:

- **A ∪ B** e

- **R - B**

A **B**
Nome -» Projeto

Empregado_PD = {Nome, Projeto, Dependente}



Projetos = {Nome, Projeto}

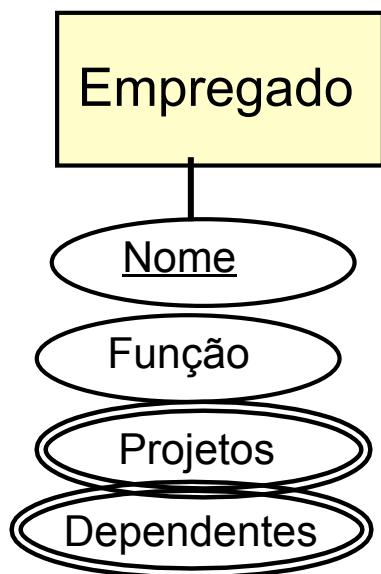
Dependentes = {Nome, Dependente}

4ª Forma Normal

- Evita redundância de tuplas \Rightarrow evita anomalias de inclusão/remoção/alteração
- Cenários típicos de normalização para 4ªFN (ocorrências de DMNT):
 - atributos **multivalorados independentes** misturados em uma única tabela
 - Ex: Empregado_PD={Nome, Projeto, Dependente}
 - atributos **multivalorados** ou **CR N:N** armazenados **de maneira incorreta** em múltiplas linhas
 - Ex: Pessoa={CPF, telefone, nome, data-nasc, ...}

Observação

- Lembrando o mapeamento ME-R → Modelo Relacional
 - atributos multivalorados definem novas relações
 - sem redundância e sem anomalias



Empregado = {Nome, Função}

Projetos = {Nome, Projeto}

Dependentes = {Nome, Dependente}

Considerações Finais (cont.)...

■ Normalização:

- ✓ uma relação por vez –
 - análise de 1ªFN, 2ª FN, 3ª FN, BCNF e 4ª FN
- ✓ **forma normal de uma relação** \Rightarrow forma normal mais restrita atendida
- ✓ **forma normal da base de dados** \Rightarrow forma normal mais restrita atendida por todas as relações

Considerações Finais (cont.)...

■ Lembrando...

■ Propriedades desejáveis:

- 1) decomposição **sem perda de junção**
(sem geração de tuplas ilegítimas)
- 2) decomposição com **preservação de dependências** (possibilidade de avaliar a DF \Rightarrow atributos na mesma tabela)

Sugestão de Leitura

- **ELMASRI, R; NAVATHE, S.B. – *Sistemas de Banco de Dados*, Addison Wesley**
 - 4ª Edição
 - **Capítulos 10 e 11**
 - 6ª Edição
 - **Capítulo 15**

EXERCÍCIO

Considere as seguintes relações e suas dependências funcionais.

1) Lotes = {ID-propriedade, municipio, nro_lote, regioao, area, preço, imposto}

ID-propriedade \rightarrow municipio, nro_lote, regioao, area, preço, imposto

municipio, nro_lote \rightarrow ID-propriedade, regioao, area, preço, imposto

municipio \rightarrow imposto

area \rightarrow preço

regiao \rightarrow municipio

2) Carros_Vendidos = {carro, data-venda, vendedor, %comissão, desconto-do-dia}

carro \rightarrow data-venda;

data-venda \rightarrow desconto-do-dia;

vendedor \rightarrow %comissão

carro \twoheadrightarrow vendedor

EXERCÍCIO

2) Carros_Vendidos = {carro, data-venda, vendedor, %comissão,
desconto-do-dia}

carro → data-venda;

data-venda → desconto-do-dia;

vendedor → %comissão

carro -» vendedor

EXERCÍCIO

- a) Indique se as formas normais 1FN, 2FN (baseada em PK e genérica), 3NF (baseada em PK e genérica) , BCNF e 4FN são atendidas.
- b) Normalize as relações para atender a cada uma dessas formas (faça todos os passos). Explique quais propriedades de decomposição são atendidas ou não, e o porquê.
- c) Discuta a influência dessa normalização no desempenho e dê exemplos.