

# SCC-240 Banco de Dados

Profa. Elaine Parros Machado de Sousa

## Normalização - 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

**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

IdProp  $\rightarrow$  Municipio, NroLote, Área, Preço

Municipio, NroLote  $\rightarrow$  Área, IdProp, Preço

Área  $\rightarrow$  Preço

**Lotes = {IdProp, Municipio, NroLote, Área, Preço }**

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

- 2ª FN genérica?
- 3ª FN genérica?

# Exemplo

Aluno, Curso  $\rightarrow$  Instrutor

Instrutor  $\rightarrow$  Curso

**Treinamento = {Aluno, Curso, Instrutor}**

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

- 2ª FN genérica?
- 3ª FN genérica?

O que acontece com 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 **superchave** em R

# Exemplo

$\text{IdProp} \rightarrow \text{Municipio}, \text{NroLote}, \text{Área}, \text{Preço}$

$\text{Municipio}, \text{NroLote} \rightarrow \text{Área}, \text{IdProp}, \text{Preço}$

$\text{Área} \rightarrow \text{Preço}$

**Lotes** = { **IdProp**, **Municipio, NroLote**, **Área**, **Preço** }

➤ A relação **Lotes** está na BCNF?



# Exemplo

$\text{IdProp} \rightarrow \text{Municipio, NroLote, \u00c1rea, Pre\u00e7o}$

$\text{Municipio, NroLote} \rightarrow \text{\u00c1rea, IdProp, Pre\u00e7o}$

$\text{\u00c1rea} \rightarrow \text{Pre\u00e7o}$

$\text{Lotes} = \{ \underline{\text{IdProp}}, \underline{\underline{\text{Municipio, NroLote}}}, \text{\u00c1rea, Pre\u00e7o} \}$

➤ **Normalizando para 3<sup>a</sup> FN Gen\u00e9rica e BCNF....**

$\text{Lotes} = \{ \underline{\text{IdProp}}, \underline{\underline{\text{Municipio, NroLote}}}, \text{\u00c1rea} \}$

$\text{\u00c1rea} = \{ \underline{\text{\u00c1rea}}, \text{Pre\u00e7o} \}$



# 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?**

# Observação

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

# 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

# Discussão sobre DFs

- Relembrando...
  - **Dependência Funcional**  $\Rightarrow$  mecanismo formal (fundamental) para definição de restrições e garantia de consistência em bases de dados relacionais
    - $A \rightarrow B$
  - E as restrições semânticas que não podem ser especificadas com DFs?

EX.?????

# Outro Exemplo

- Informação sobre empregados a ser armazenada na base de dados de uma empresa:
  - nome do empregado
  - nomes dos projetos em que trabalha
  - nomes de seus dependentes
- atributos: **nome, projeto, dependente**
- dependências funcionais???

# Exemplo (cont.)

- atributos: **nome, projeto, dependente**
  - semanticamente:
    - um conjunto de valores de **projeto** é determinado por um valor de **nome**, e somente por **nome**
    - idem para **dependente**
    - **projeto** e **dependente** não têm relação alguma...



## 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

# Exemplo

**Empregado = {Nome, Projeto, Dependente}**

**Nome ->> Projeto**

**Nome ->> Dependente**

➤ **Dados:**

- Carlos trabalha no projeto Museu Virtual e tem 2 dependentes: Mário e Joana;
- Ana trabalha nos projetos Museu Virtual e Cidadania, e tem 2 dependentes: Paulo e Sônia;

➤ Como **armazenar** os dados na relação Empregado de maneira a manter a **semântica**?

## Exemplo (cont.)

- Carlos trabalha no projeto Museu Virtual e tem 2 dependentes: Mário e Joana;
- Ana trabalha nos projetos Museu Virtual e Cidadania, e tem 2 dependentes: Paulo e Sônia;

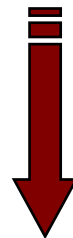
Empregado = {Nome, Projeto, Dependente}

**Anomalia?**

{<Carlos, Museu Virtual, Mário>,  
<Carlos, Museu Virtual, Joana>,  
<Ana, Cidadania, Paulo>,  
<Ana, Museu Virtual, Sônia>}

Empregado = {Nome, Projeto, Dependente}

{<Carlos, Museu Virtual, Mário>,  
<Carlos, Museu Virtual, Joana>,  
<Ana, Cidadania, Paulo>,  
< Ana, Cidadania, Sônia>,  
<Ana, Museu Virtual, Paulo>,  
<Ana, Museu Virtual, Sônia>}



atende

**Dependências Multivaloradas  
na relação Empregado**

# Dependência Multivalorada

- **Dependência Multivalorada  $X \twoheadrightarrow Y$**   
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  $t_4[Z] = t_1[Z]$ ,
      - $Z = R - (X \cup Y)$

# Dependência Multivalorada

- **Dependências Multivaloradas**

- cenário típico: quando atributos multivalorados são desmembrados em múltiplas ocorrências de tuplas por causa da 1ª FN
- identificadas pelo projetista da base de dados

Nome ->> Projeto

Nome ->> Dependente

Empregado = {Nome, Projeto, Dependente}

{<Carlos, Museu Virtual, Mário>,  
<Carlos, Museu Virtual, Joana>,  
<Ana, Cidadania, Paulo>,  
< Ana, Cidadania, Sônia>,  
<Ana, Museu Virtual, Paulo>,  
<Ana, Museu Virtual, Sônia>}

**=> Empregado atende às DMs....**

**=> Mas... qual é o PROBLEMA???**

# Dependência Multivalorada

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

- se:

- (a)  $Y \subseteq X$  ou



**Dependência  
Multivalorada Trivial**

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

- 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}

# Dependência Multivalorada

- Problema da DM Não-Trivial:
  - requer **redundância** nas tuplas
  - como garantir **consistência**?
  - Exemplo:
    - Empregado={Nome, Projeto, Dependente}
    - está na BCNF, mas ainda vulnerável a inconsistências....

 **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 superchave** em R

# 4ª Forma Normal

## ■ Exemplos...

- Empregado={Nome, Dependente}

Nome -> Dependente

**OK!!**

- Empregado={Nome, Projeto, Dependente}

Nome -> Projeto

Nome -> Dependente

**Não!!**

# 4ª Forma Normal

- Intuição:

Nome  $\rightarrow$  Projeto

Nome  $\rightarrow$  Dependente

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



Dependentes = {Nome, Dependente}

Projetos = {Nome, Projeto}

# 4ª Forma Normal

- Normalizando a relação para a 4ª FN....
  - dada uma DMNT  $A \twoheadrightarrow B$  na relação  $R$ ,  
substitui-se  $R$  por:
    - $A \cup B$  e
    - $R - B$

# Exemplo

Nome -> Programa

Nome -> Orientado

Professor = {Nome, Programa, Orientado}

considerando: Nome -> Programa

Programa = {Nome, Programa}

Orientação = {Nome, Orientado}

# 4ª Forma Normal

- Evita redundância nas tuplas  $\Rightarrow$  evita inconsistências causadas por inclusão/remoção/alteração de novas tuplas
  - lembrando: não violar uma DMNT significa replicar informação, mas redundância pode gerar outras inconsistências...

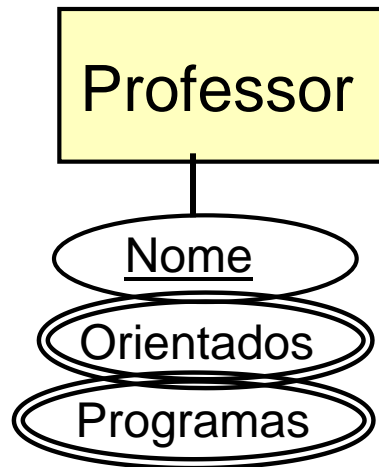


# 4ª Forma Normal

- Normalização para 4ªFN é aplicada quando atributos multivalorados **independentes** são misturados na mesma relação (DMNT)
- Reduz espaço de armazenamento
- Mais restrita que BCNF
- Propriedade desejada: **decomposição sem perda de junção**

# Observação

- Mapeamento ME-R → Modelo Relacional
  - atributos multivalorados definem novas relações
    - sem redundância e sem anomalias



Orientação = {Nome, Orientado}

Programa = {Nome, Programa}

# Considerações Finais quanto à Normalização

- 1ªFN, 2ª FN, 3ª FN, BCNF e 4ª FN são consideradas para cada relação
  - base de dados é considerada normalizada para uma determinada FN quando todas as suas relações estiverem nessa FN
- Normalização  $\Rightarrow$  decomposição de relações
  - aumenta consistência
  - reduz desempenho  $\Rightarrow$  operações de junção

# Sugestão de Leitura

- **ELMASRI, R; NAVATHE, S.B.** – *Sistemas de Banco de Dados*, Addison Wesley, 4ª Edição.
  - **Capítulo 10** – Dependência Funcional e normalização em um banco de dados relacional
  - **Capítulo 11** – Algoritmos para Projeto de Banco de Dados Relacional e Demais Dependências