

SCC-240 Banco de Dados

Profa. Elaine Parros Machado de Sousa

Estagiários PAE:
Pedro Bugatti
Robson Cordeiro

Normalização - Parte 2



Normalização

- Formas Normais baseadas em dependências funcionais
 - baseadas em chave primária
 - 1ª FN
 - 2ª FN
 - 3ª FN**Parte 1**
 - 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 2**

ICMC - USP - GBDI

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

ICMC - USP - GBDI

3

Normalização

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

Exemplo

IdProp \rightarrow Municipio, NroLote, Área
Municipio, NroLote \rightarrow Área, IdProp
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?
 - 3ª FN?

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?
 - 3ª FN?

O que acontece com a DF Instrutor \rightarrow Curso?

Forma Normal de Boyce - Codd (BCNF)

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

ICMC - USP - GBDI

7

Exemplo

IdProp \rightarrow Município, NroLote, Área
Município, NroLote \rightarrow Área, IdProp
Área \rightarrow Preço

Lotes = { IdProp, Município, NroLote, Área, Preço }

- A relação **Lotes** está na BCNF?

Exemplo

IdProp \rightarrow Município, NroLote, Área
Município, NroLote \rightarrow Área, IdProp
Área \rightarrow Preço

Lotes = { IdProp, Município, NroLote, Área, Preço }

- **Normalizando para BCNF....**

Lotes = { IdProp, Município, NroLote, Área }

Área = { Área, Preço }

Exemplo

Aluno, Curso \rightarrow Instrutor
Instrutor \rightarrow Curso

Treinamento = { Aluno, Curso, Instrutor }

- A relação **Treinamento** está na BCNF?
- Alternativas 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?

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

ICMC - USP - GBDI

12

Discussão sobre DFs

- **Dependência Funcional** \Rightarrow mecanismo formal (fundamental) para definição de restrições e garantia de consistência em bases de dados relacionais

- $A \rightarrow B$

- E quanto às restrições semânticas que não podem ser especificadas com DFs?

ICMC - USP - GBDI

13

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???

ICMC - USP - GBDI

14

Exemplo (cont.)

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



Dependência Multivalorada

ICMC - USP - GBDI

15

Dependência Multivalorada

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

A \multimap 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

ICMC - USP - GBDI

16

Exemplo

Empregado = {Nome, Projeto, Dependente}

Nome \multimap Projeto

- 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**?

ICMC - USP - GBDI

17

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;

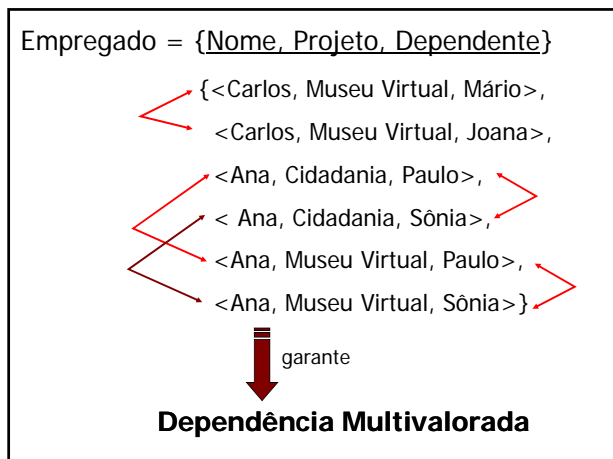
Empregado = {Nome, Projeto, Dependente}

PROBLEMA?

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

ICMC - USP - GBDI

18



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

ICMC - USP - GBDI 20

Dependência Multivalorada

- **Simetria** na definição de DM:
 - se $X \twoheadrightarrow Y$, então

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

Empregado = {Nome, Projeto, Dependente}

Nome \twoheadrightarrow Projeto

Nome \twoheadrightarrow Dependente

ICMC - USP - GBDI 21

Dependência Multivalorada

- **Dependências Multivaloradas**
 - ocorrem quando atributos multivalorados são desmembrados em múltiplas ocorrências de tuplas por causa da 1ª FN
 - identificadas pelo projetista da base de dados

ICMC - USP - GBDI 22

Nome \twoheadrightarrow Projeto Nome \twoheadrightarrow 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 $X \cup Y = R$ \Rightarrow **Dependência Multivalorada Trivial**
 - (b) $X \cup Y = R$ \Rightarrow **Dependência Multivalorada Não-Trivial**
 - caso contrário \Rightarrow **Dependência Multivalorada Não-Trivial**

ICMC - USP - GBDI 25

Exemplos

- DM Trivial (DMT)
 - Nome \twoheadrightarrow Projeto
 - Empregado = {Nome, Projeto}
- DM Não-Trivial (DMNT)
 - Nome \twoheadrightarrow Projeto
 - Nome \twoheadrightarrow Dependente
 - Empregado = {Nome, Projeto, Dependente}

ICMC - USP - GBDI

26

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!**

ICMC - USP - GBDI

27

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

ICMC - USP - GBDI

28

4ª Forma Normal

- Exemplos...
 - Empregado = {Nome, Dependente} **OK!!**
 - Nome \twoheadrightarrow Dependente
 - Empregado = {Nome, Projeto, Dependente} **Não!!**
 - Nome \twoheadrightarrow Projeto
 - Nome \twoheadrightarrow Dependente (por simetria)

ICMC - USP - GBDI

29

4ª Forma Normal

- Colocando a relação na 4ª FN....

Nome \twoheadrightarrow Projeto
Nome \twoheadrightarrow Dependente

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

Dependentes = {Nome, Dependente}

Projetos = {Nome, Projeto}

ICMC - USP - GBDI

30

4ª Forma Normal

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

ICMC - USP - GBDI

31

Exemplo

Nome -> Programa
Nome -> Orientado
Professor = {Nome, Programa, Orientado}

considerando: Nome -> Programa

Programa = {Nome, Programa}
Orientação = {Nome, Orientado}

ICMC - USP - GBDI

32

4ª Forma Normal

- evita redundância nas tuplas => evita inconsistências causadas por inclusão/remoção/alteração de novas tuplas
- lembrando: não violar uma DMNT significa replicar informação (*para evitar inferência incorreta de relacionamento entre atributos independentes*)... mas redundância pode gerar outras inconsistências...

ICMC - USP - GBDI

33

4ª Forma Normal

- normalização é importante 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**

ICMC - USP - GBDI

34

Observação

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



Orientação = {Nome, Orientado}
Programa = {Nome, Programa}

ICMC - USP - GBDI

35

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 => decomposição de relações
 - aumenta consistência
 - reduz desempenho => operações de junção

ICMC - USP - GBDI

36

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

USP - ICMC - GBDI

37