

# Bacharelado em Sistemas de Informação

# BANCOS DE DADOS

## Aula 10

## Normalização



# Projetos de Bancos de Dados

✓ **Considere o Modelo Relacional abaixo para modelagem de um sistema de Controle dos sócios de um clube (os símbolos '{' e '}' indicam repetição)**

Sócio(CPF, nome-sócio, endereço-sócio, RG-sócio, {nome-dependente, data-nascimento, idade}, quantidade-dependentes, data-validade, código-categoria, descrição-categoria)

Mensalidade(CPF, ano, mês, nome-sócio, quantidade-dependentes, valor-mês, valor-total-ano)

Categoria(código-categoria, descrição-categoria, valor-mensalidade)

✓ **Existe algum problema nessas relações?**



# Projetos de Bancos de Dados

- ✓ Medidas informais para mensurar a qualidade de um projeto de BD:
  - semântica dos atributos
    - quanto mais fácil explicar significado da relação, melhor será o modelo de esquema da relação
  - redução de valores redundantes nas tuplas
  - redução de valores *null* nas tuplas
  - impedimento para a geração de valores ilegítimos nas tuplas



# Normalização

- ✓ **Codd 1972** – verificação de um conjunto de regras pra certificar se o esquema de uma relação satisfaz a uma *forma normal*.
  - *3 formas normais previstas: 1FN, 2FN, 3FN*
- ✓ **Boyce-Codd** – definição mais forte da *3FN*
- ✓ **4FN e 5FN** dependência multivalorada e junção



# Normalização

## ✓ Objetivo:

- gerar um conjunto de esquemas de relações:
  - sem **redundância** desnecessária
  - que permita **recuperar informações** de **forma fácil**
- projetar esquemas na **forma normal** apropriada
- cada relação deve ter atributos de um único assunto!

Para atingir esses objetivos



Conceitos de normalização

# Normalização

- ✓ Para atingir um bom projeto de Banco de Dados, às vezes é necessário **decompor uma relação em relações menores.**
- ✓ Propriedades a serem preservadas:
  - junção sem perda
  - preservação da dependência

# Projetos de Bancos de Dados

## ✓ Processo de normalização:

- Elaborado em torno do conceito de formas normais
- Formas normais: regras que devem ser obedecidas para que uma tabela seja considerada “bem projetada”
- Tem origem na definição de E.F. Codd\*

\*E.F.Codd. “Normalized Data Base Structure: A Brief Tutorial”, Proc 1971 ACM SIGFIDET Workshop on Data Description, Access, and Control, San Diego, Califórnia.



# Normalização

✓ O processo de normalização permite ao projetista **CONTROLAR** quanto da **consistência** é garantida pela maneira de construção do sistema, e quanto deve ser responsabilidade do SGBD.

✓ Normalizar  eficiência dos aplicativos 

✓ Normalizar  inconsistências 



# Normalização

- ✓ Processo para simplificar as relações através de regras
- ✓ São definidas basicamente três formas normais, em ordem crescente de simplicidade das relações:
  - ✱ Primeira Forma Normal (1FN)
  - ✱ Segunda Forma Normal (2FN)
  - ✱ Terceira Forma Normal (3FN)
- ✓ A aplicação das formas normais gera um BD com **mais relações**. Porém, essas relações são **mais simples**.
- ✓ Para a maioria dos projetos, a 3FN é suficiente. Porém, há outras formas que podem ser aplicadas:
  - ✓ FNBC: Forma Normal de Boyce-Codd
  - ✓ 4FN
  - ✓ 5FN



# Normalização

## ➤ PRIMEIRA FORMA NORMAL (1FN)

⇒ uma relação está na Primeira Forma Normal quando todos os seus atributos são **atômicos** e **monovalorados**:

⇒ **Atômicos**: simples, indivisíveis

- Exemplo: endereço não é atômico, porque é composto por rua, número e cidade.

⇒ **Monovalorados**: um único valor no domínio do atributo, isto é, não pode haver relações dentro de relações. Não pode haver repetições!

# Normalização

## ➤ Como deixar as relações na 1FN:

**Se existirem atributos compostos:**

- Substituí-los por atributos atômicos

**Exemplo:**

Aluno (RA, nome, endereço)

Aluno (RA, nome, rua , número , bairro , cidade , estado)

# Normalização

## ➤ PRIMEIRA FORMA NORMAL (1FN)

### – Atributo monovalorado

- aquele que tem apenas um valor (não uma lista).

### – Exemplo:

Aluno (RA, nome, {código-disciplina, nome-disciplina})

Consiste em uma lista de todas as disciplinas nas quais o aluno se matriculou. Pode ser uma ou várias disciplinas (usaremos o símbolo { } para indicar repetições)

# Normalização

## ➤ Como deixar as relações na 1FN:

**Se existirem atributos multivalorados:**

**Duas  
situações**

**Quantidade de valores é  
pequena e conhecida  
previamente.**

**Quantidade de valores é  
desconhecida, grande ou  
variável.**

# Normalização

## ➤ Como deixar as relações na 1FN:

- Quantidade de valores é **pequena e conhecida previamente**
  - ➡ Substitui-se o atributo multivalorado por um **conjunto de atributos de mesmo domínio**, cada um monovalorado representando uma ocorrência do valor.

### Exemplo:

Aluno (RA, nome, {notas-bimestrais})

Aluno (RA, nome, nota1, nota2, nota3, nota4)

# Normalização

## ➤ Como deixar as relações na 1FN:

- Quantidade de valores é **desconhecida, grande ou variável**.
  - ➡ retira-se da relação o atributo multivalorado
  - ➡ cria-se uma nova relação que tem o mesmo conjunto de atributos chave, mais o atributo multivalorado. Verificar a formação da chave primária.

### Exemplo:

Aluno (RA, nome, {código-disciplina, nome-disciplina})



# Normalização

## ➤ Como deixar as relações na 1FN:

- Quantidade de valores é desconhecida, grande ou variável.

### Exemplo:

Aluno (RA, nome, {código-disciplina, nome-disciplina})



Aluno (RA, nome)

Matrícula (RA, código-disciplina, nome-disciplina)



# Exercícios

**Deixe os BDs abaixo na Primeira Forma Normal (atributos entre { } indicam repetição)**

## 1. Biblioteca

Livro(cod-livro, título, cod-editora, nome-editora, endereço-editora, {cod-autor, nome-autor})

## 2. Controle de Projetos

ProjetoEmpr(cod-proj, tipo, descrição, {cod-empregado, nome-empregado, categoria-empregado, salario-categoria, data-início-projeto, data-fim-projeto})

## 3. Fábrica de Móveis

Movel(cod-movel, descrição, preço-unitario, {cod-mão-obra, descrição-mão-obra, preço-mão-obra, quantidade-usada-mo}, {cod-matéria-prima, descrição-mat-prima, preço-mat-prima, quantidade-usada-mp})

Mao-obra(cod-mão-obra, descrição-mão-obra, preço-mão-obra)

Materia-prima(cod-matéria-prima, descrição-mat-prima, preço-mat-prima)

## 4. Clínica Médica

Médico(CRM, med-nome, med-telefone, {med-data-consulta, med-hora-consulta, codigo-paciente})

Paciente(pac-código, pac-nome, pac-ddd, pac-telefone, {cod-convênio, descrição-convênio})

Consulta(CRM-Medico, codigo-paciente, data-consulta, hora-consulta, valor-consulta, cod-convênio, descrição-convênio)

## 5. Folha de Pagamento

Funcionario(func-código, func-nome, func-endereço, func-data-nascimento, numero-banco, agencia-banco, cidade-banco, {func-data-falta, func-motivo-falta})

Pagamento(func-código, func-mês, func-ano, valor-depositado)



# Dependência Funcional

- ✓ Propriedade semântica ou do significados dos atributos
  - ✓ Projetistas de um BD usam o significado dos atributos de uma relação para especificar dependências funcionais
  - ✓ Generaliza a noção de superchave (conjunto de atributos de uma relação que identifica unicamente cada tupla)
- ⇒ Seja  $\alpha \subseteq R$  e  $\beta \subseteq R$ .
- ⇒ A dependência funcional  $\alpha \rightarrow \beta$  realiza-se em  $R$  se:
- ⇒ em qualquer relação válida  $r(R)$ , para todos os pares de tuplas  $t_1$  e  $t_2$  em  $r$  tal que  $t_1[\alpha] = t_2[\alpha]$ , então  $t_1[\beta] = t_2[\beta]$  será também verdade.

Dizemos que:

$\alpha$  determina funcionalmente  $\beta$   
ou  
 $\beta$  depende funcionalmente de  $\alpha$ .

# Dependência Funcional

- ✓ Baseia-se no reconhecimento que os valores de alguns atributos podem ser determinados a partir de outros.
- ✓ Esse conhecimento não pode ser inferido pelo SGBD → deve ser **identificado** durante a fase de PROJETO do BD.



# Dependência Funcional

## ✓ Exemplo:

▪ **RA** → nome, idade

**Aluno (RA, nome, disciplina, idade)**

{2,Rodolfo, CC-302, 20  
2, Rodolfo, CC-304, 20  
10,Eduardo, CC-304, 21  
10,Eduardo, CC-308, 21  
25,Vanessa, CC-304, 20  
32,César, CC-304, 22  
38,Mariana, CC-303, 21  
54,Érica, CC-302,20 }

Sempre que o **RA** se repete,  
**nome** e **idade** se repetem?

**SIM !!!**

**nome** e **idade** SÃO  
FUNCIONALMENTE DEPENDENTES  
de **RA**

# Dependência Funcional

## ✓ Exemplo:

▪ **idade** → **disciplina**

**Aluno (RA, nome, disciplina, idade)**

{2,Rodolfo, CC-302, 20  
2, Rodolfo, CC-304, 20  
10,Eduardo, CC-304, 21  
10,Eduardo, CC-308, 21  
25,Vanessa, CC-304, 20  
32,César, CC-304, 22  
38,Mariana, CC-303, 21  
54,Érica, CC-302,20 }

Sempre que o **idade** se repete,  
**disciplina** se repete?

**NÃO !!!**

**disciplina NÃO É**  
FUNCIONALMENTE DEPENDENTE  
de **idade**



# Dependência Funcional

## ➡ Alguns exemplos de exercícios anteriores:

- ✓ livro (livro-código, livro-título, livro-editora, livro-ano)  
**livro-código → livro-título, livro-editora, livro-ano**
- ✓ usuário(usu-código, usu-nome, usu-endereço, usu-cidade, usu-curso)  
**usu-código → usu-nome, usu-endereço, usu-cidade, usu-curso**
- ✓ venda (cod-cliente, cod-produto, data, quantidade, forma-pagamento)  
**cod-cliente, cod-produto, data → quantidade, forma-pagamento**
- ✓ autor(aut-código, aut-nome)  
**aut-código → aut-nome**
- ✓ pagamento (cod-obra-arte, num-prestação, valor, data-vencimento, data-pagamento)  
**cod-obra-arte, num-prestação → valor, data-vencimento, data-pagamento**

# Dependência Funcional

- ✓ As dependências são informações **semânticas** fornecidas pelo projetista.
- ✓ As dependências funcionais fazem parte da base intencional. **Refletem a intenção do projetista!**
- ✓ São usadas para evitar redundância e, assim, **construir um BD mais eficiente!**



# Normalização

## ➤ SEGUNDA FORMA NORMAL (2FN)

### ➤ Dependência funcional total:

- dependência funcional  $\alpha \rightarrow \beta$  é chamada de *dependência funcional total* se a remoção de qualquer atributo de  $\alpha$  implicar que a dependência não mais será assegurada.

➡ uma relação está na 2FN se:

- ✖ está na 1ª Forma Normal
- ✖ todos os atributos que não participam da chave primária são **funcionalmente dependentes** de toda a chave primária.
- ✖ caso haja atributos dependentes de somente parte da chave primária, a relação deve ser normalizada.



# Normalização

## ➤ SEGUNDA FORMA NORMAL (2FN)

➤ para verificar se uma relação está na 2FN, pergunta-se:

**1. Qual é a chave primária da relação?**

✧ Se a chave primária for **composta** por dois ou mais atributos, pergunta-se:

**2. Há algum atributo que não é chave que depende funcionalmente de somente parte da chave?**

# Normalização

## ➤ SEGUNDA FORMA NORMAL (2FN)

### ➤ Exemplo:

Matrícula (RA, código-disciplina, nome-aluno, nome-disciplina, quant-aula-disciplina, nota, frequência)

**1. Qual é a chave primária da relação?**

***RA + código-disciplina***

**2. Há algum atributo que não é chave que depende funcionalmente de somente parte da chave?**

***RA → nome-aluno***

***código-disciplina → nome-disciplina, quant-aula-disciplina***



# Normalização

## ➤ Como deixar as relações na 2FN (considerando que já estão na 1FN):

1. Verificam-se os grupos de atributos que dependem da mesma parte da chave ⇒ retiram-se da relação todos os atributos de um desses grupos
2. Cria-se uma nova relação, que tem esse grupo como atributos **não chaves** e os **atributos que determinam funcionalmente esse grupo como chave**
3. Repetem-se os procedimentos 1 e 2 para cada grupo, até que todas as relações contenham somente atributos que dependem da chave toda.



# Normalização

➤ **Como deixar as relações na 2FN (considerando que já estão na 1FN):**

➤ Exemplo:

Matrícula (RA, código-disciplina, nome-aluno, nome-disciplina, quant-aula-disciplina, nota, frequência)

Primeiro grupo: *RA → nome-aluno*  
*Aluno(RA, nome-aluno)*

Segundo grupo: *código-disciplina → nome-disciplina, quant-aula-disciplina*  
*Disciplina(código-disciplina, nome-disciplina, quant-aula-disciplina)*

# Normalização

➤ Como deixar as relações na 2FN (considerando que já estão na 1FN):

➤ Exemplo:

Matrícula (RA, código-disciplina, nome-aluno, nome-disciplina, quant-aula-disciplina, nota, frequência)

## BD final:

Matrícula (RA, código-disciplina, nota, frequência)

Aluno(RA, nome-aluno)

Disciplina(código-disciplina, nome-disciplina, quant-aula-disciplina)

# Exercícios

**Deixe os BDs abaixo na Segunda Forma Normal (atributos entre { } indicam repetição)**

## 1. Controle de Projetos

Projeto(cod-proj, tipo-proj, descrição-proj, {cod-empregado, nome-empr, cod-categoria-empr, descricao-categoria, salario-categoria, data-início-func-proj, data-fim-func-proj, quantidade-dias-no-projeto})

## 2. Sistema Acadêmico

Aluno(RA, nome-aluno, cod-curso, nome-curso, {cod-disc, nome-disc, ano-cursada, cod-prof-disc, nome-prof-disc, nota-disc, freq-disc })

## 3. Folha de Pagamento

Funcionario(func-código, func-nome, func-endereço, func-data-nascimento, {código-cargo, nome-cargo, data-início, data-fim, func-tempo-no-cargo}, {data-pagto, codigo-banco, codigo-agência, nome-banco, valor-pagto})

## 4. Clínica Médica

Médico(CRM, med-nome, med-telefone, {num\_consulta, data-consulta, hora-consulta, cod-paciente, nome-paciente, nome-convenio, quantidade-exames-solicitados {cod-exame-solicitado, descrição-exame-solicitado, resultado-exame-solicitado} })  
Exame(cod-exame, descrição-exame)



# Normalização

## ➤ TERCEIRA FORMA NORMAL (3FN)

### ➤ Dependência funcional transitiva:

- dependência funcional é chamada de *dependência funcional transitiva* se existir um conjunto de atributos **Z** que não é chave e ambas  $\alpha \rightarrow Z$  e  $Z \rightarrow X$  estão asseguradas.

⇒ uma relação está na 3FN se:

- ✖ está na 2ª Forma Normal
- ✖ todos os atributos que não participam da chave primária **não são funcionalmente dependentes** de outros atributos que não são chaves.



# Normalização

## ➤ TERCEIRA FORMA NORMAL (3FN)

### ➤ Caso 1: atributos calculados

Aluno (RA, nome-aluno, data-nascimento, idade)

*Idade é dependente de data-nascimento porque pode ser calculada a partir deste atributo (data-nascimento).*



# Normalização

## ➤ TERCEIRA FORMA NORMAL (3FN)

### ➤ Caso 2: dependência funcional transitiva

Aluno (RA, nome-aluno, cod-curso, nome-curso, título do curso)

*Há algum atributo que não é chave que depende funcionalmente de outro atributo não chave?*

Aluno (RA, nome-aluno, **cod-curso**, **nome-curso**, **título-curso**)

***cod-curso → nome-curso, título-curso***

# Normalização

## ➤ Como deixar as relações na 3FN (considerando que já estão na 2FN):

1. Eliminam-se todos os atributos que são calculados a partir de outros atributos presentes no BD.
2. Para cada grupo de atributos não-chaves dependentes funcionalmente de outros atributos não-chaves, cria-se uma nova relação, que tem os atributos dependentes como **não-chaves** e os **atributos que causam a dependência como chave primária**.
3. Repete-se o procedimento 2 para cada grupo, até que todas as relações não contenham atributos dependentes de atributos não-chaves.



# Normalização

➤ Como deixar as relações na 3FN (considerando que já estão na 2FN):

➤ Exemplo: atributos calculados

Aluno (RA, nome-aluno, data-nascimento, **idade**)

**BD final:**

Aluno(RA, nome-aluno, data-nascimento)

# Normalização

➤ Como deixar as relações na 3FN (considerando que já estão na 2FN):

➤ Exemplo: dependência funcional

Aluno (RA, nome-aluno, **cod-curso**, **nome-curso**, **título do curso**)



## BD final:

Aluno(RA, nome-aluno, cod-curso)

Curso (cod-curso, nome-curso, título-curso)



# Normalização

## ➤ Forma Normal de Boyce-Codd (FNBC)

- também utiliza o conceito de dependência funcional
- uma relação R está na FNBC se para todas as dependências funcionais na forma  $\alpha \rightarrow \beta$  ao menos uma das seguintes se realiza:
  - $\alpha \rightarrow \beta$  é uma dependência funcional trivial ( $\beta$  está contido em  $\alpha$ )
  - $\alpha \rightarrow \beta$  é uma superchave para a relação R

# Normalização

## ➤ Forma Normal de Boyce-Codd (FNBC)

- Em outras palavras:
  - uma relação está na FNBC se e somente se, para todas as dependências funcionais  $\alpha \rightarrow \beta$  existentes na relação,  $\alpha$  é chave candidata da relação, ou seja, os determinantes (lado esquerdo da dependência funcional) são chaves candidatas;

- Exemplo:

Emprestimo\_livro(ISBN, cod\_emprestimo, data\_devolução,  
cod\_cliente)



# Normalização

## ➤ Forma Normal de Boyce-Codd (FNBC)

### ■ Exemplo:

Emprestimo\_livro(ISBN, cod\_emprestimo, data\_devolução,  
cod\_cliente)

Considere que:

- ISBN, cod\_emprestimo → data\_devolução, cod\_cliente
- cod\_emprestimo é chave candidata (ocorrência única)

Então, o BD ficaria:

Emprestimo\_livro(ISBN, cod\_emprestimo)

Emprestimo(cod\_emprestimo, data\_devolução, cod\_cliente)

# Normalização

## ➤ Forma Normal de Boyce-Codd (FNBC)

- Em outras palavras:
  - a FNBC é mais rígida que a 3FN, mas, na prática quase todas relações que estão na 3FN também estão na FNBC;
  - as relações que estão na 3FN e que têm uma única chave candidata, estão automaticamente na FNBC;
  - uma relação que está na 3FN e que não tem chaves candidatas sobrepostas está na FNBC;
  - a FNBC não faz referência explícita à 1FN e à 2FN e nem ao conceito de dependência transitiva.





# Normalização

- Às vezes a aplicação simples da FNBC (sem aplicar 1FN, 2FN e 3FN) pode gerar perda de dependência.
- Por isso, as 3 metas de projeto para um BD relacional são:
  - FNBC
  - junção sem perda
  - preservação da dependência
- Se não for possível atingir essas metas, é preferível:
  - 3FN
  - junção sem perda
  - preservação da dependência

# Normalização

## ➤ QUARTA FORMA NORMAL (4FN)

- considera dependências multivaloradas: repetição de atributos não-chaves, gerando redundância desnecessária;
- Exemplo:

Música(nome, intérprete, gravadora)

nome	intérprete	gravadora
Aquarela do Brasil	Toquinho	Polygram
Aquarela do Brasil	Tim Maia	Som Livre
Aquarela do Brasil	Toquinho	Som Livre
Andança	Beth Carvalho	Som Livre
Andança	Roupa Nova	Polygram
Andança	Beth Carvalho	Polygram

# Normalização

## ➤ QUARTA FORMA NORMAL (4FN) MusicalInterprete

- Exemplo:  
Música(nome, intérprete, gravadora)

nome	intérprete	gravadora
Aquarela do Brasil	Toquinho	Polygram
Aquarela do Brasil	Tim Maia	Som Livre
Aquarela do Brasil	Toquinho	Som Livre
Andança	Beth Carvalho	Som Livre
Andança	Roupa Nova	Polygram
Andança	Beth Carvalho	Polygram

### MusicalInterprete

nome	intérprete
Aquarela do Brasil	Toquinho
Aquarela do Brasil	Tim Maia
Andança	Beth Carvalho
Andança	Roupa Nova

### MusicaGravadora

nome	gravadora
Aquarela do Brasil	Polygram
Aquarela do Brasil	Som Livre
Andança	Som Livre
Andança	Polygram



# Normalização

## ➤ QUINTA FORMA NORMAL (5FN)

- verifica se é possível dividir as relações na 4FN em relações mais simples:

- Exemplo:

MusicalInterprete(nome, intérprete)

### MusicalInterprete

nome	intérprete
Aquarela do Brasil	Toquinho
Aquarela do Brasil	Tim Maia
Andança	Beth Carvalho
Andança	Roupa Nova

Como deixar mais simples ???

# Normalização

## ➤ QUINTA FORMA NORMAL (5FN)

- verifica se é possível dividir as relações na 4FN em relações mais simples:

- Exemplo:

MusicalInterprete(nome, intérprete)

Pode ser dividida em:

Musica(codm, nome)

Interprete(codi, nome)

MusicalInterprete(codm, codi)

### Musica

codm	nome
1	Aquarela do Brasil
2	Andança

### Interprete

codm	nome
1	Toquinho
2	Roupa Nova

# Normalização

- ✓ A normalização de relações é feita na grande maioria das vezes decompondo-se uma relação em duas ou mais.
- ✓ No entanto, é frequente que nas consultas a BDs, seja necessário **recompor** a relação original.
- ✓ Isso não causa nenhum transtorno, uma vez que a não normalização de relações causa **anomalias de atualização**, não de consulta.
- ✓ A decomposição deve ser feita com cuidado, para **evitar perda de informação na recomposição**.



# Exercícios

**Deixe os BDs abaixo na Terceira Forma Normal (atributos entre { } indicam repetição)**

## 1. Controle de Projetos

ProjetoEmpr(cod-proj, tipo-proj, descrição-proj, {cod-empregado, nome-empr, cod-categoria-empr, descricao-categoria, salario-categoria, data-início-func-proj, data-fim-func-proj, quantidade-dias-no-projeto})

## 2. Sistema Acadêmico

Aluno(RA, nome-aluno, cod-curso, nome-curso, {cod-disc, nome-disc, ano-cursada, cod-prof-disc, nome-prof-disc, nota-disc, freq-disc })

## 3. Folha de Pagamento

Funcionario(func-código, func-nome, func-endereço, func-data-nascimento, {código-cargo, nome-cargo, data-início, data-fim, func-tempo-no-cargo}, {data-pagto, codigo-banco, codigo-agência, nome-banco, valor-pagto})

## 4. Clínica Médica

Médico(CRM, med-nome, med-telefone, {num\_consulta, data-consulta, hora-consulta, cod-paciente, nome-paciente, nome-convenio, quantidade-exames-solicitados {cod-exame-solicitado, descrição-exame-solicitado, resultado-exame-solicitado} })

Exame(cod-exame, descrição-exame)



# Bacharelado em Sistemas de Informação

# BANCOS DE DADOS

## Aula 10

## Normalização

