



Cadeira de Tecnologias de Informação

Ano lectivo 2008/09

Normalização de Tabelas

Prof^a Ana Lucas (Responsável)

Mestre Cristiane Pedron

Mestre Fernando Naves

Eng^a Filipa Pires da Silva

Dr. José Camacho

Dr. Luís Vaz Henriques

Normalizar Porquê ?

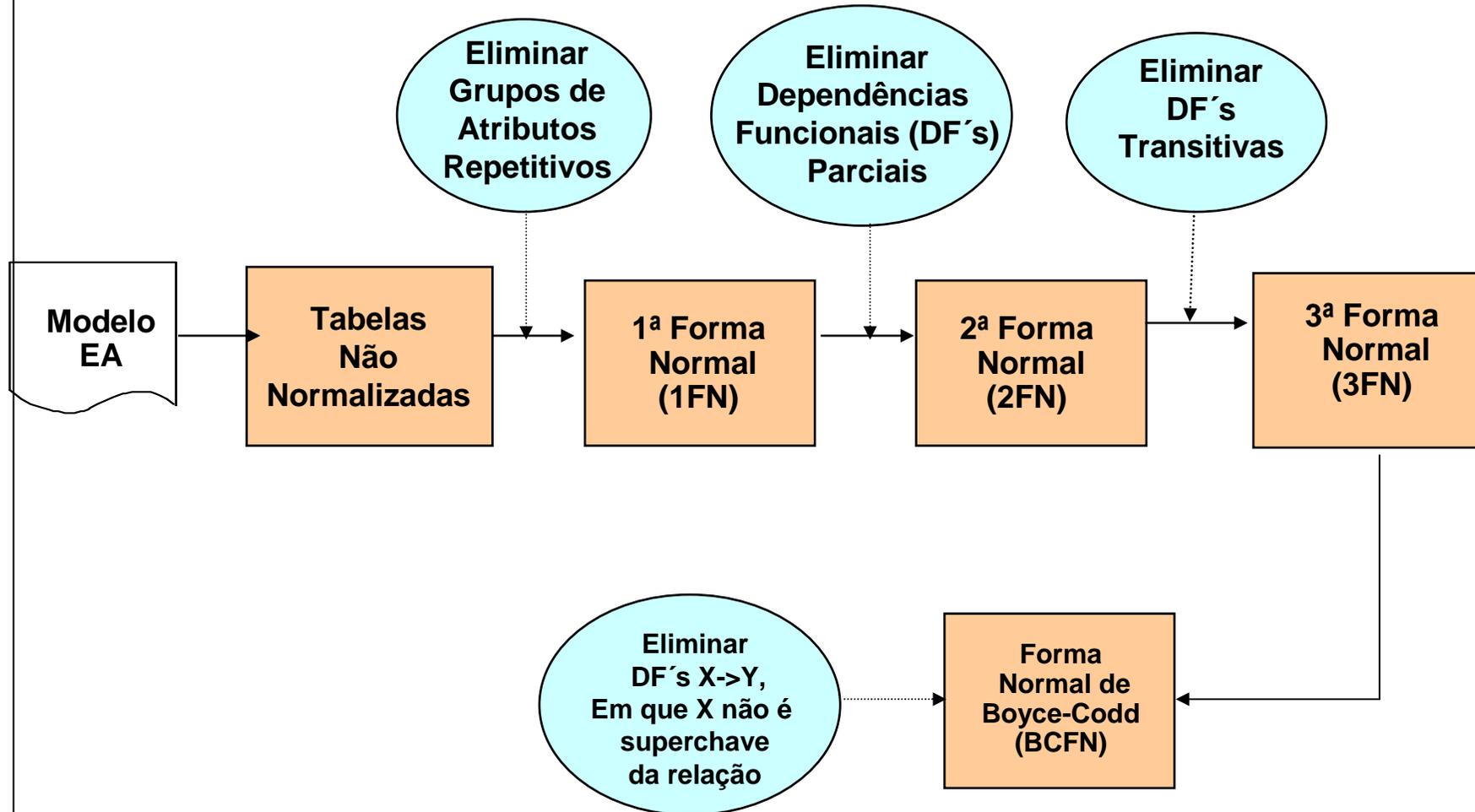
Após a construção do modelo conceptual dos dados (Modelo Entidade/Associação) é feita a transformação para o modelo lógico (Esquema de Tabelas).

O Desenho de Tabelas obtido representa a estrutura da informação de um modo natural e completo.

Mas terá o mínimo de redundância possível ?

A Normalização tem como objectivo avaliar a qualidade do Desenho de Tabelas e transformá-lo (em caso de necessidade) num Desenho (Conjunto de Tabelas) equivalente, menos redundante e mais estável.

Etapas



Exemplo

Numa Escola, pretende-se manter informação sobre:

- os Estudantes da escola (nº interno, nome e curso a que pertencem);**
- as Disciplinas que são ministradas na escola (nº disciplina e nome);**
- os Professores contratados pela escola (código, nome e grau académico);**
- Inscrições dos Alunos em Disciplinas;**
- Notas obtidas pelos alunos nas disciplinas em que estão inscritos.**

Definição de 1ª FORMA NORMAL

Uma relação está na 1ª Forma Normal (1NF) se e só se cada linha contém exactamente um valor para cada atributo.

Dado que as Relações(Tabelas) são estruturas bidimensionais, então no cruzamento de uma linha com uma coluna (atributo) só é possível armazenar valores atómicos.

Relação Não Normalizada

Relação NOTAS

<u>Nº Estudante</u>	<u>Nome Estudante</u>	<u>Curso</u>	<u>Nº Disciplina</u>	<u>Nome Disciplina</u>	<u>Cod Professor</u>	<u>Nome Professor</u>	<u>Categoria Professor</u>	<u>Nota</u>
21934	Antunes	Informática	04	Álgebra	21	Gil Alves	PA	15
			14	Análise Sist.	87	Ana Lopes	PC	-
			23	Progr.Linear	43	Paulo Pinto	AS	16
42346	Bernardo	Matemática	08	Topologia	32	Nuno Neves	AE	10
			04	Álgebra	21	Gil Alves	PA	-
			12	Geometria	21	Gil Alves	PA	18
			16	Lógica	32	Nuno Neves	AE	13
54323	Correia	Estatística	04	Álgebra	21	Gil Alves	PA	11
			08	Topologia	32	Nuno Neves	AE	10
...

Dados em forma tabular, conforme são visualizados pelo utilizador.

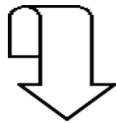
Esta Tabela foi desenhada para representar Inscrições e Aprovações de Alunos em Disciplinas. Cada estudante está inscrito e/ou já foi aprovado em várias disciplinas, com a informação correspondente.

Assim, temos um grupo de atributos repetitivo :

{NºDisciplina, NomeDisciplina, Cod.Professor, Nome Professor, CategoriaProfessor, Nota}

1ª Forma Normal

NºEstudante	Nome Estudante	Curso	Nº Disciplina	Nome Disciplina	Cod Professor	Nome Professor	Nota
-------------	----------------	-------	---------------	-----------------	---------------	----------------	------



Relação ESTUDANTE

Chave

<u>NºEstud.</u>	Nome Estudante	Curso
21934	Antunes	Informática
42346	Bernardo	Matemática
54323	Correia	Estatística
...



Relação NOTAS

Chave

<u>NºEstudante</u>	<u>Nº Disciplina</u>	Nome Disciplina	Cod Professor	Nome Professor	Nota
21934	04	Álgebra	21	Gil Alves	15
21934	14	Análise Sist.	87	Ana Lopes	12
21934	23	Progr.Linear	43	Paulo Pinto	16
42346	08	Topologia	32	Nuno Neves	10
42346	04	Álgebra	21	Gil Alves	12
42346	12	Geometria	21	Gil Alves	18
42346	16	Lógica	32	Nuno Neves	13
54323	04	Álgebra	21	Gil Alves	11
54323	08	Topologia	32	Nuno Neves	10
...

Assumimos que uma Disciplina tem um só Professor (Responsável) (Nº_Disciplina → Cod_Professor), mas um Professor pode ser responsável por várias Disciplinas

1ª FORMA NORMAL

Problemas na relação NOTAS

– REMOÇÃO

Se quisermos apagar a informação sobre todos os alunos aprovados ou inscritos numa determinada disciplina, então perdemos toda a informação dessa disciplina e do respectivo professor.

– ACTUALIZAÇÃO

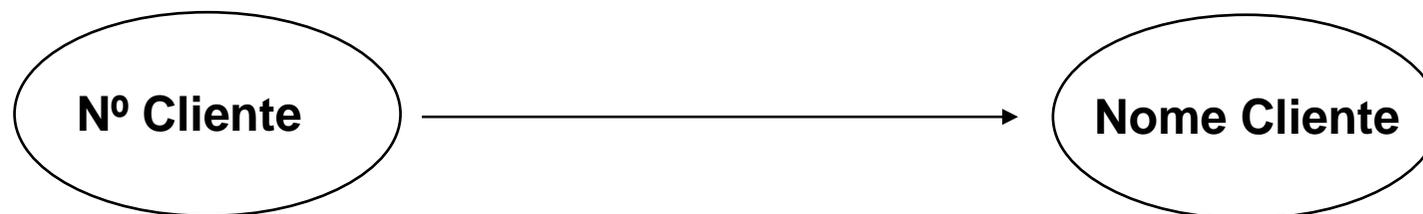
Se pretendermos modificar o nome de uma disciplina (por exemplo Álgebra passa a Álgebra e Aplicações) é necessário percorrer toda a tabela e fazer essa modificação para todos os alunos inscritos ou aprovados nessa disciplina;

No caso de falhar a aplicação de modificação em alguma ocorrência, então teríamos dados inconsistentes.

Dependências Funcionais

Por vezes, dois atributos (ou dois grupos de atributos) estão intrinsecamente ligados entre si.

Ex:



 Num determinado instante, em qualquer ponto da B.D. onde figurem estes dois atributos, a um mesmo nº de cliente corresponderá necessariamente o mesmo nome (*).

Escrevemos N^a Cliente → Nome Cliente

(*) Note que o inverso poderá não ser verdade!

Chaves e Descritores

- **Chaves Candidatas**
 - são os atributos de uma entidade que identificam, de forma inequívoca, uma ocorrência específica dessa entidade, distinguindo-a das restantes. A chave tem de conter todos os atributos necessários à identificação de todos os atributos da Tabela, ou seja Chave → Atributo, para todos os atributos da Tabela
- **Chave Primária**
 - é a chave candidata escolhida
- **Descritores**
 - são os atributos que apenas descrevem ou caracterizam as ocorrências de uma entidade, não pertencendo a nenhuma chave candidata

Definição de 2ª FORMA NORMAL

Uma relação está na 2ª Forma Normal se está na primeira e se todos os atributos descritores (não pertencentes a nenhuma Chave Candidata) dependerem da totalidade da chave (e não apenas de parte dela – Dependências Funcionais Parciais).

Dependências Parciais

Chave

Relação NOTAS

<u>Nº Estudante</u>	<u>Nº Disciplina</u>	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor	Nota
21934	04	Álgebra	21	Gil Alves	PA	15
21934	14	Análise Sist.	87	Ana Lopes	PC	12
21934	23	Progr.Linear	43	Paulo Pinto	AS	16
42346	08	Topologia	32	Nuno Neves	AE	10
42346	04	Álgebra	21	Gil Alves	PA	12
42346	12	Geometria	21	Gil Alves	PA	18
42346	16	Lógica	32	Nuno Neves	AE	13
54323	04	Álgebra	21	Gil Alves	PA	11
54323	08	Topologia	32	Nuno Neves	AE	10
...

Os atributos **Nome Disciplina**, **Cod-Professor**, **Nome-Professor** e **Categoria-Professor** dependem apenas do atributo **Nº Disciplina** (que está estritamente contido na chave da tabela)

2ª Forma Normal

<u>Nº</u> <u>Estudante</u>	<u>Nº</u> <u>Disciplina</u>	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor	Nota
-------------------------------	--------------------------------	--------------------	------------------	-------------------	------------------------	------

Chave e atributos que dependem da totalidade desta

Relação NOTA

<u>Nº</u> <u>Estudante</u>	<u>Nº</u> <u>Disciplina</u>	Nota
21934	04	15
21934	14	12
21934	23	16
42346	08	10
42346	04	12
42346	12	18
42346	16	13
54323	04	11
54323	08	10
...

Atributos que dependem de parte da chave mais a referida parte da chave

Relação DISCIPLINA

<u>Nº</u> <u>Disciplina</u>	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor
04	Álgebra	21	Gil Alves	PA
14	Análise Sist.	87	Ana Lopes	PC
23	Progr.Linear	43	Paulo Pinto	AS
08	Topologia	32	Nuno Neves	AE
12	Geometria	21	Gil Alves	PA
16	Lógica	32	Nuno Neves	AE
...

Definição de 3ª FORMA NORMAL

Uma relação está na 3ª Forma Normal se está na 2ª Forma Normal e se não existirem atributos descritores (não pertencentes a nenhuma Chave Candidata) a dependerem funcionalmente de outros atributos descritores (não chaves) - Dependências Transitivas.

Assim sendo, cada atributo deve depender apenas das Chaves Candidatas da relação

Dependências Funcionais Directas

Chave

Relação DISCIPLINAS

<u>Nº Disciplina</u>	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor
04	Álgebra	21	Gil Alves	PA
14	Análise Sist.	87	Ana Lopes	PC
23	Progr.Linear	43	Paulo Pinto	AS
08	Topologia	32	Nuno Neves	AE
12	Geometria	21	Gil Alves	PA
16	Lógica	32	Nuno Neves	AE
...

Os atributos **Nome Professor** e **Categoria Professor** dependem do atributo **Cod.Professor** (que não é chave da tabela) e portanto as DFs $N^{\circ}Disciplina \rightarrow Nome_Professor$ e $N^{\circ}Disciplina \rightarrow Categoria_Professor$, não são DFs directas (diz-se então que são Transitivas, porque, p.ex., $N^{\circ}Disciplina \rightarrow CodProfessor$ e $CodProfessor \rightarrow Nome_Professor$).

3ª Forma Normal

<u>Nº Disciplina</u>	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor
--------------------------	--------------------	------------------	-------------------	------------------------

Atributos que dependem
do atributo não chave +
esse atributo

Atributos que apenas
dependem da chave

Relação PROFESSOR

Relação DISCIPLINA

<u>Cod Professor</u>	Nome Professor	Categoria Professor
21	Gil Alves	PA
87	Ana Lopes	PC
43	Paulo Pinto	AS
32	Nuno Neves	AE
...

<u>Nº Disciplina</u>	Nome Disciplina	Cod Professor
04	Álgebra	21
14	Análise Sist.	87
23	Progr.Linear	43
08	Topologia	32
12	Geometria	21
16	Lógica	32
...

Forma Normal de Boyce-Codd (FNBC)

Superchave – é qualquer conjunto de atributos que contém uma chave

Ex: se BI é chave da tabela funcionário, nome+BI é uma superchave da mesma tabela

Uma relação está na Forma Normal de Boyce-Codd se e só se, para todas as dependências funcionais $X \rightarrow Y$ existentes na relação, X é chave ou superchave da relação

Forma Normal de Boyce-Codd (FNBC)

Só em casos muito raros uma relação está na 3FN e não está na FNBC.

As relações que estão na 3ª Forma Normal e que têm **uma única chave candidata, estão automaticamente na FNBC.** (adaptado de Vincent e Srinivasan, 1993)

Mais concretamente, uma relação na 3FN **que não tem várias chaves candidatas overlapping está na FNBC.**

(Vincent e Srinivasan, 1993)

Forma Normal de Boyce-Codd (FNBC)

Ex: seja a relação Encomenda (n_fornecedor, nome_fornecedor, n_produto, quantidade)

Vamos assumir que quer o n_fornecedor, quer o nome_fornecedor são únicos.

Esta relação tem duas chaves candidatas: (n_fornecedor, n_produto) e (nome_fornecedor, n_produto)

A relação está na 3ª FN, qualquer que seja a chave escolhida

Dependências Funcionais presentes na relação:

- n_fornecedor, n_produto-> quantidade
- nome_fornecedor, n_produto-> quantidade
- **n_fornecedor-> nome_fornecedor**
- **nome_fornecedor-> n_fornecedor**

A relação não está na FNBC porque nem n_fornecedor, nem nome_fornecedor são chaves ou superchaves da relação

Forma Normal de Boyce-Codd (FNBC)

Anomalias da Relação Encomenda

n_fornecedor	nome_fornecedor	n_produto	quantidade
1	Silva&Silva	1000	300
1	Silva&Silva	1010	250
1	Silva&Silva	1050	700

- **O nome do fornecedor é repetido para cada produto encomendado**
- **Se o fornecedor mudar de nome, essa mudança tem de ser efectuada em todas as linhas da relação referentes a esse fornecedor**

A relação encomenda vai ser decomposta nas seguintes duas relações:

Fornecedor (n_fornecedor, nome_fornecedor)

Encomenda (n_fornecedor, n_produto, quantidade)

Estas relações estão na FNBC

Normalização - Exemplo

1ª Forma Normal

Uma relação está na 1ª Forma Normal se e só se cada linha contém exactamente um valor para cada atributo

Exemplo (dados referentes a uma encomenda de livros - estrutura não normalizada):

Pedido_Livro (Nome-cliente, ISBN, Data-pedido, Título, Autor(es), Quantidade, Preço, Valor-total)

Na 1ª Forma Normal:

Pedido_Livro (Nome-cliente, ISBN, Data-pedido, Título, Quantidade, Preço, Valor-total)

Autoria (ISBN, Autor)

Normalização - Exemplo

2ª Forma Normal

Uma relação está na 2ª Forma Normal se está na primeira e se todos os atributos descritores dependerem da totalidade das chaves candidatas.

Exemplo (na 1ª FN):

Pedido_Livro (Nome-cliente, ISBN, Data-pedido, Título, Quantidade, Preço, Valor-total)

Autoria (ISBN, Autor)

Aplicando a 2ª Forma Normal:

Pedido_Livro (Nome-cliente, ISBN, Data-pedido, Quantidade, Valor-total)

Autoria (ISBN, Autor)

Livro (ISBN, Título, Preço)

Normalização - Exemplo

3ª Forma Normal

Uma relação está na 3ª Forma Normal se está na 2ª Forma Normal e se não existirem atributos descritores a dependerem funcionalmente de outros atributos descritores (que não pertencem a nenhuma chave candidata).

Exemplo (na 2ª FN):

Pedido_Livro (Nome-cliente, ISBN, Data-pedido, Quantidade, Valor-total)

Autoria (ISBN, Autor)

Livro (ISBN, Título, Preço)

Este esquema relacional está na 3ª Forma Normal porque, em nenhuma das relações existem dependências funcionais transitivas e está na FNBC porque cada relação tem uma única chave candidata