



Cadeira de Tecnologias de Informação

Ano lectivo 2009/2010

Normalização Relacional

Tópicos

- Normalizar porquê?
- Etapas da modelização da informação
- Objectivos da normalização
- Etapas da normalização
- Como decompor uma relação
- Exemplo de normalização

Normalizar Porquê?

Após a construção do modelo conceptual dos dados (Diagrama de Classes) é feita a transformação para o modelo lógico (Esquema Relacional).

O Esquema Relacional 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 Esquema Relacional e **decompo-lo, em caso de necessidade, noutro Esquema Relacional equivalente, menos redundante e mais estável.**

Etapas da Modelização da Informação

Etapa 1

Análise das necessidades informacionais e elaboração do Diagrama de Classes

Análise e Concepção

Diagrama de Classes

Etapa 2

Transformação do Diagrama de Classes num Esquema Relacional

Desenho

Relações não normalizadas

Etapa 3

Normalização do Esquema Relacional

Relações normalizadas

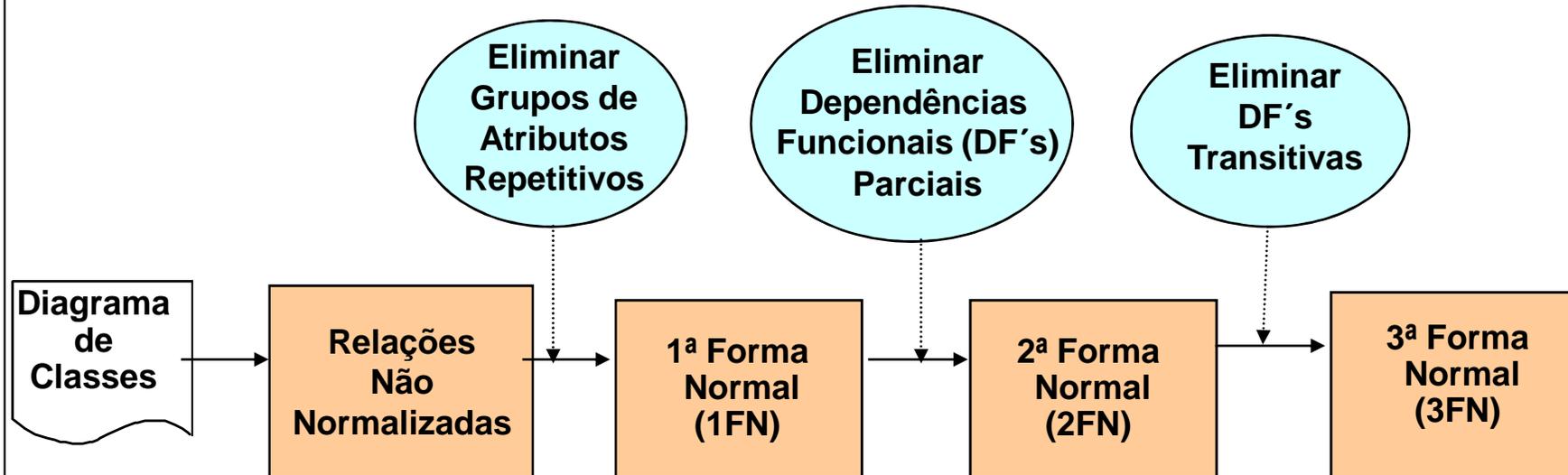
Objectivos da Normalização

- ✓ **Minimizar a redundância de dados**
- ✓ **Conseguir que cada relação possa ser consistente e correctamente actualizada**

N_Func	Nome	Categoria	Vencimento	Disciplina
10234	Tiago Silva	Prof. Associado	4.000 €	GF1
10345	Tomás Ribeiro	Prof. Auxiliar	3.000 €	TI
11908	Maria Marques	Assistente	2.000 €	SIG
16000	Carlos Antunes	Prof. Auxiliar	3.000 €	TI
17241	Raquel Silva	Assistente	2.000 €	SIG

- ✓ **Problemas de Actualização** – Para alterar o salário de uma categoria tem de se alterar várias linhas
- ✓ **Problemas de Anulação** - Se se apagar o registo do Tiago Silva deixa de se ter informação relativa ao salário de um Prof. Associado

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 categoria);
- **Inscrições dos Alunos em Disciplinas;**
- **Notas obtidas pelos alunos nas disciplinas em que estão inscritos;**
- **Considera-se que a uma Disciplina está afecto um único Professor (responsável), mas que um Professor pode ser responsável por várias Disciplinas.**

Definição de 1ª FORMA NORMAL

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

Dado que as 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

Nº Estudante	Nome Estudante	Curso	Nº Disciplina	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor	Nota
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 relação 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

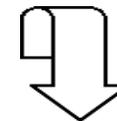
NºEstudante	Nome Estudante	Curso	Nº Disciplina	Nome Disciplina	Cod Professor	Nome Professor	Nota
-------------	----------------	-------	---------------	-----------------	---------------	----------------	------



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

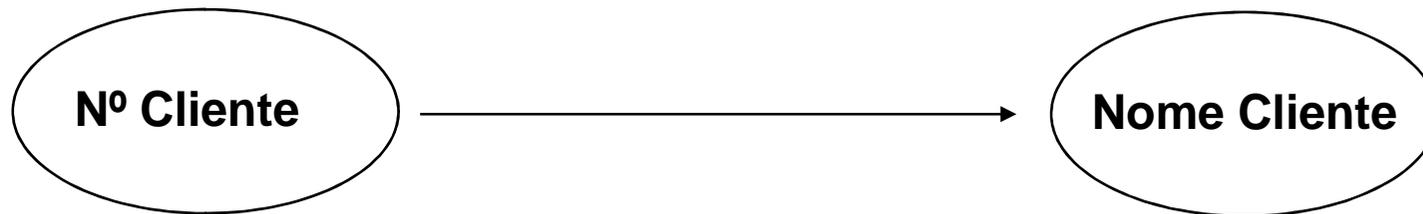
Se pretendermos modificar o nome de uma disciplina (por exemplo Álgebra passa a Álgebra e Aplicações) é necessário percorrer toda a relação e fazer essa modificação para todos os alunos inscritos ou aprovados nessa disciplina;

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 base de dados onde figurem estes dois atributos, a um mesmo nº de cliente corresponderá necessariamente o mesmo nome (*).

Isto significa que é válida a Dependência Funcional (DF)

Nº Cliente → Nome Cliente

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

Definições (1/2)

- **Chaves Candidatas**

São os atributos de uma relação que identificam, de forma inequívoca, uma ocorrência específica dessa relação, distinguindo-a das restantes. A chave tem de conter todos os atributos necessários à identificação de todos os restantes atributos da relação, ou seja
Chave → Atributo, para todos os atributos da relação

- **Chave Primária**

É a chave candidata escolhida

- **Descritores**

São os atributos que apenas descrevem ou caracterizam as ocorrências de uma relação.

Definições (2/2)

- **Atributos Primos**

São os atributos que pertencem a pelo menos uma chave candidata

Ex: se as chaves candidatas de Factura são (n_cliente, n_factura) e (nome_cliente, n_factura) os atributos n_cliente, nome_cliente e n_factura são primos.

- **Atributos Não Primos**

São os atributos que não pertencem a nenhuma chave candidata Ex: data_factura, montante_global,...

Definição de 2ª FORMA NORMAL

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

Dependências Funcionais 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 relação)

2ª Forma Normal

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

Chave e atributos que dependem da totalidade desta

Relação NOTA

<u>Nº</u> Estudante	<u>Nº</u> Disciplina	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> Disciplina	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 (3FN) se está na 2ª Forma Normal e se não existirem atributos não primos a dependerem funcionalmente de outros atributos não primos - Dependências Funcionais Transitivas.

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

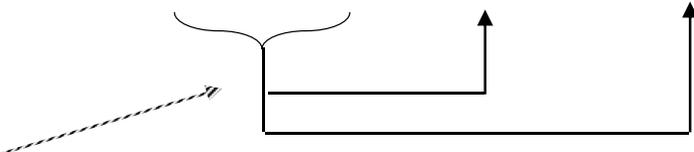
Se a relação contiver **atributos calculados** (p.ex. total parcial numa linha de factura), devem retirar-se estes atributos (total parcial=preço×quantidade, portanto preço, quantidade → total parcial, o que viola a 3FN)

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

<u>Nº</u> <u>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
...

Outra Definição de 3ª FORMA NORMAL

Uma relação está na 3ª Forma Normal se cada atributo não primo depende da chave primária (1FN), de toda a chave (2FN) e de nada mais do que da chave (3FN)

Formas Normais - Resumo



Como Decompôr uma Relação

Teorema da Decomposição Binária

Dada uma relação $R(A, B, C, D, E, F)$ e dada a dependência funcional $A \rightarrow B$, a relação é sempre decomponível em $R_1(A, B)$ e $R_2(A, C, D, E, F)$

Normalização – Exemplo (1/3)

Normalizar até à 3FN a relação resultante da factura apresentada

Bill #: 90141						
Patient #: 12345			Date: 7/20/08			
Patient Name: Mary Baker			Date admitted: 7/14/08			
Patient Address: 300 Oak Street			Discharge date: 7/17/08			
City-State-Zip: Boulder, CO 80638						
Cost Center	Cost Name	Date Charged	Item Code	Desc	Charge	Bal Due
100	Room & Board	7/14/08	2000	Semi-prv room	200	
		7/14/08	2005	Television	5	
		7/15/08	2000	Semi-prv room	200	
				Subtotal		405
110	Laboratory	7/14/08	1580	Glucose	25	
		7/15/08	1585	Culture	20	
				Subtotal		45
				Balance due		450

Normalização – Exemplo (2/3)

Começam por se retirar os atributos calculados Subtotal e Balance due

Patient Bill (Bill#, Patient #, Date, Patient Name, Date admitted, Discharge date, Patient Address, City-State-Zip, Cost Center, Cost Name, Date Charged, Item Code, Desc, Charge)

Bill# → Patient #, Date, Patient Name, Date admitted, Discharge date, Patient Address, City-State-Zip

Patient_Bill_v1 (Bill#, Patient #, Date, Patient Name, Date admitted, Discharge date, Patient Address, City-State-Zip, Discharge date) 2FN

Patient # → Patient Name, Patient Address, City-State-Zip

Patient_Bill_Line (Bill#, Cost Center, Cost Name, Date Charged, Item Code, Desc, Charge) 1FN

Cost Center → Cost Name

Item Code → Desc, Charge, Cost Center

Normalização – Exemplo (3/3)

Patient_Bill_v1 (Bill#, Patient #, Date, Patient Name, Date admitted, Discharge date, Patient Address, City-State-Zip, Discharge date) – 2FN

Patient # → Patient Name, Patient Address, City-State-Zip

Patient_Ident (Patient #, Patient Name, Patient Address, City-State-Zip) 3FN

Patient_Bill_v2 (Bill#, Patient #, Date, Date admitted, Discharge date) 3FN

Patient_Bill_Line (Bill#, Cost Center, Cost Name, Date Charged, Item Code, Desc, Charge) 1FN

Cost Center → Cost Name

Cost_Ident (Cost Center, Cost Name) 3FN

Patient_Bill_Line_v1 (Bill#, Cost Center, Date Charged, Item Code, Desc, Charge) 1FN

Item Code → Desc, Charge, Cost Center

Item_Ident (Item Code, Desc, Charge, Cost Center) 3FN

Patient_Bill_Line_v2 (Bill#, Item Code, Date Charged) 3FN

Pretendemos com esta aula sobre Normalização que os alunos compreendessem:

- **Normalizar porquê?**
- **Etapas da modelização da informação**
- **Objectivos da normalização**
- **Etapas da normalização**