

---

# **Introdução às Bases de Dados**

Normalização de Esquemas Relacionais

---

FCUL, Departamento de Informática

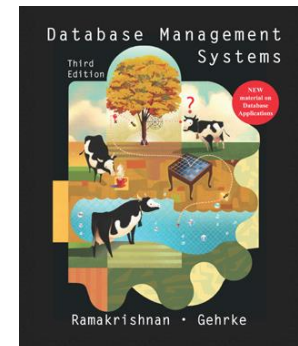
Ano Letivo 2021/2022

Ana Paula Afonso

# Sumário e Referências

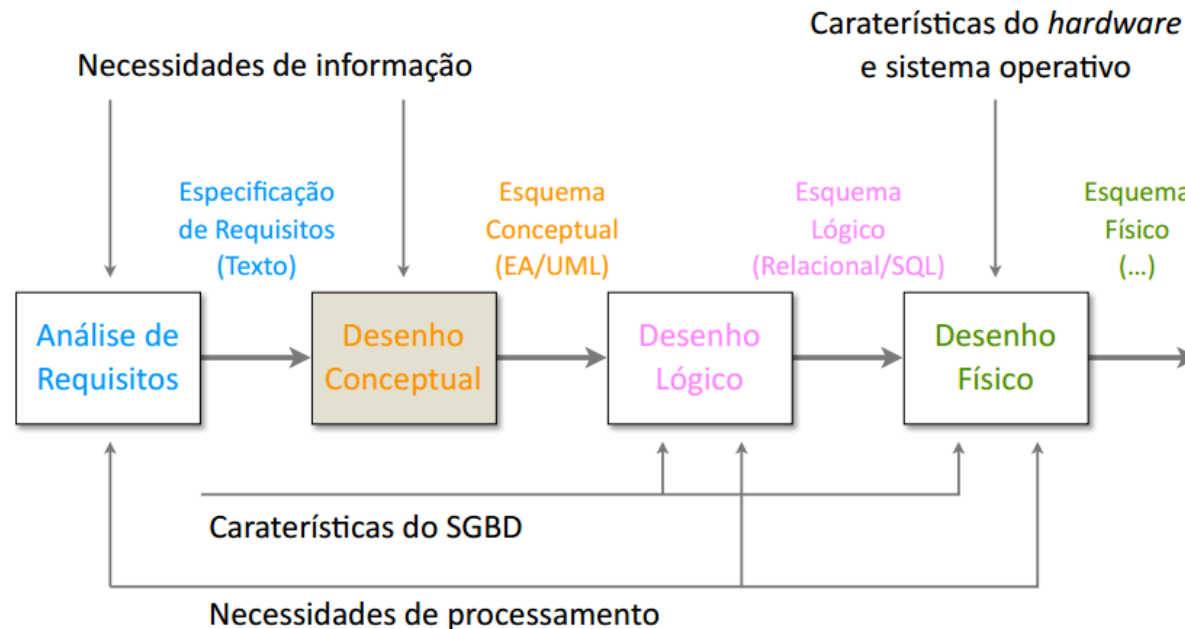
---

- Sumário
  - Motivação para a normalização de esquemas relacionais
  - Problemas da redundância de dados
    - Exemplo de relação com redundância
  - Dependências funcionais
  - Etapas da normalização
    - Principais formas normais
    - Exemplo de normalização
- Referências
  - R. Ramakrishnan (**capítulo 19, secção 19.1 - 19.4**)



# Motivação da Normalização

- Após a construção do modelo conceptual dos dados (Modelo E/A) é feita a transformação para um 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?



Fonte: António Ferreira, Guião SIBD, 2016

# Normalização

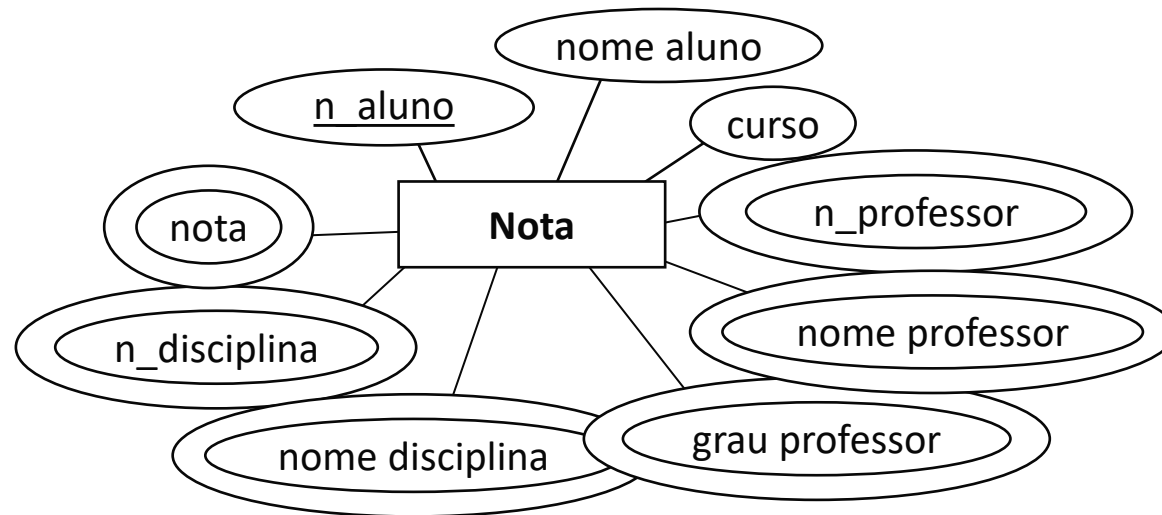
---

- Dado um esquema relacional por vezes
  - Os mesmos dados existem armazenados em múltiplos locais (**redundância**)
  - Existência de **anomalias** aquando da escrita de dados (incoerências dos dados)
- **Normalização** permite melhorar a qualidade do esquema através da
  - Eliminação da redundância dos dados
  - Prevenção de anomalias
- **Normalização** é uma abordagem que envolve a
  - Decomposição sucessiva de relações até se obter um conjunto de relações
  - ... sem redundâncias e que permitam inserções, atualizações e remoções sem incoerências

# Exemplo de Relação não Normalizada

---

Possível EA para a tabela Nota (**INCORRETO**)



Notação: atributo com vários valores



# Anomalias de Relação não Normalizada

<u>n_aluno</u>	nome aluno	curso	n_disciplina	nome disciplina	n_professor	nome professor	grau professor	nota
21934	Antunes	Informática	04	Álgebra	21	Gil Algébrico	PA	15
			14	Análise Sist.	87	Ana Listada	PC	12
			23	P.Linear	43	Plínio	AS	16
42346	Bernardo	Matemática	08	Topologia	32	Topo Lógico	AE	10
			04	Álgebra	21	Gil Algébrico	PA	12
			12	Geometria	21	Gil Algébrico	PA	18
			16	Lógica	32	Topo Lógico	AE	13
54323	Correia	Estatística	04	Álgebra	21	Gil Algébrico	PA	11
			08	Topologia	32	Topo Lógico	AE	10
...	...	...	...	...	...	...	...	...

- **Inserção** de um novo professor requer indicação de outros dados
- **Atualização** do grau do professor tem de afetar várias linhas
- **Remoção** de um professor elimina os dados da disciplina a que está afeto

# Problemas da Redundância dos Dados

---

- A redundância introduz problemas (**anomalias**) de coerência e manutenção
  - **Anomalia de inserção** - informação que é independente não pode ser inserida de forma separada na Base de Dados
  - **Anomalia de atualização** - a modificação de informação num conjunto de ocorrências implica a criação de inconsistências ou a necessidade de alterar informação noutras instâncias da Base de Dados que são independentes das primeiras
  - **Anomalia de remoção** - a remoção de informação acarreta a perda de outra informação independente contida na Base de Dados

# Principais Formas Normais

---

- **Primeira** forma normal: 1FN
  - Colunas da relação guardam apenas **um** valor por linha
  - Tipicamente verificado em BD relacionais
- **Segunda** forma normal: 2FN = 1FN +
  - Colunas não pertencentes às chaves candidatas da relação...
  - ...dependem da **totalidade** das colunas de cada chave
  - Trivial, se chaves da relação tiverem apenas uma coluna
- **Terceira** forma normal: 3FN = 2FN +
  - Colunas não pertencentes às chaves candidatas...
  - ...dependem **apenas** das chaves candidatas



# Exemplo de Relação não Normalizada

---

## Nota

<u>n_aluno</u>	nome aluno	curso	n_disciplina	nome disciplina	n_professor	nome professor	grau professor	nota
21934	Antunes	Informática	04	Álgebra	21	Gil Algébrico	PA	15
			14	Análise Sist.	87	Ana Listada	PC	12
			23	P.Linear	43	Plínio	AS	16
42346	Bernardo	Matemática	08	Topologia	32	Topo Lógico	AE	10
			04	Álgebra	21	Gil Algébrico	PA	12
			12	Geometria	21	Gil Algébrico	PA	18
			16	Lógica	32	Topo Lógico	AE	13
54323	Correia	Estatística	04	Álgebra	21	Gil Algébrico	PA	11
			08	Topologia	32	Topo Lógico	AE	10
...	...	...	...	...	...	...	...	...

# Passagem/Decomposição para a 1ª Forma Normal (1FN)

<u>n_aluno</u>	nome aluno	curso						
			n disciplina	nome disciplina	n_professor	nome professor	grau professor	nota

**Aluno**

*Chave*

<u>n_aluno</u>	nome	curso
21934	Antunes	Informática
42346	Bernardo	Matemática
54323	Correia	Estatística
...	...	...

**Nota**

*Chave*

<u>n_aluno</u>	<u>n_disciplina</u>	nome disciplina	n_professor	nome professor	grau professor	nota
21934	04	Álgebra	21	Gil Algébrico	PA	15
21934	14	Análise Sist.	87	Ana Listada	PC	12
21934	23	P.Linear	43	Plínio	AS	16
42346	08	Topologia	32	Topo Lógico	AE	10
42346	04	Álgebra	21	Gil Algébrico	PA	12
42346	12	Geometria	21	Gil Algébrico	PA	18
42346	16	Lógica	32	Topo Lógico	AE	13
54323	04	Álgebra	21	Gil Algébrico	PA	11
54323	08	Topologia	32	Topo Lógico	AE	10
...	...	...	...	...	...	...

# Anomalias da Relação Nota

---

Nota

<i>Chave</i>						
<u>n_aluno</u>	<u>n_disciplina</u>	nome disciplina	n_professor	nome professor	grau professor	nota
21934	04	Álgebra	21	Gil Algébrico	PA	15
21934	14	Análise Sist.	87	Ana Listada	PC	12
21934	23	P.Linear	43	Plínio	AS	16
42346	08	Topologia	32	Topo Lógico	AE	10
42346	04	Álgebra	21	Gil Algébrico	PA	12
42346	12	Geometria	21	Gil Algébrico	PA	18
42346	16	Lógica	32	Topo Lógico	AE	13
54323	04	Álgebra	21	Gil Algébrico	PA	11
54323	08	Topologia	32	Topo Lógico	AE	10
...	...	...	...	...	...	...

# Anomalias da Relação Nota

---

- **Inserção**

- Não se pode inserir informação sobre uma nova disciplina ...
- ... enquanto não existirem alunos inscritos para a nova disciplina (o atributo `n_aluno` faz parte da chave da relação)

- **Remoção**

- Quando se apaga a informação sobre todos os alunos que têm uma determinada disciplina ....
- ... perde-se toda a informação dessa disciplina e do respetivo professor

- **Atualização**

- Quando se modifica o nome de uma disciplina ...
- ... é necessário percorrer toda a relação e fazer essa modificação para todos os alunos que tivessem essa disciplina
- No caso de falhar a aplicação de modificação em alguma ocorrência, então ter-se-iam dados inconsistentes

# Principais Formas Normais

---

- **Primeira** forma normal: 1FN
  - Colunas da relação guardam apenas **um** valor por linha
  - Tipicamente verificado em bases de dados relacionais
- **Segunda** forma normal: 2FN = 1FN +
  - Se **não existirem dependências** funcionais entre **subconjuntos próprios da chave e atributos não chave**
  - Colunas não pertencentes às chaves candidatas da relação...
  - ...dependem da **totalidade** das colunas de cada chave
  - As relações com chaves simples (1 único atributo) que estejam na 1FN estão sempre na 2FN

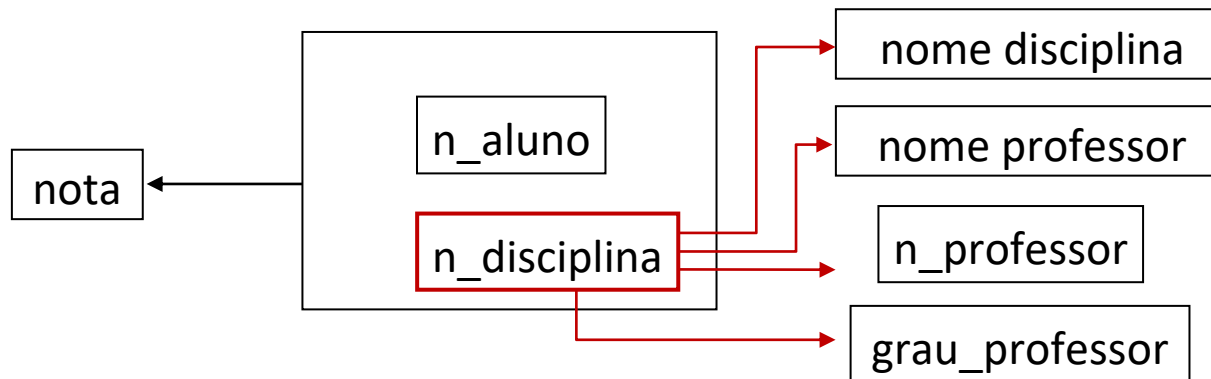


# Decomposição na 2FN

Disciplina (n\_disciplina, n\_aluno, nota,  
nome\_disciplina, n\_professor, nome\_professor, grau\_professor)

Existem **DFs** entre atributo subconjunto da chave (n\_disciplina) e atributos não chave

n\_disciplina → nome\_disciplina, n\_professor, nome\_professor, grau\_professor



# Decomposição na 2FN

---

<u>n_aluno</u>	<u>n_disciplina</u>	nome disciplina	cod professor	nome professor	grau professor	nota
----------------	---------------------	-----------------	---------------	----------------	----------------	------

**Nota**

<u>n_aluno</u>	<u>n_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
...	...	...

**Disciplina**

<u>n_disciplina</u>	nome disciplina	cod professor	nome professor	grau professor
04	Álgebra	21	Gil Algébrico	PA
14	Análise Sist.	87	Ana Listada	PC
23	P.Linear	43	Plínio	AS
08	Topologia	32	Topo Lógico	AE
12	Geometria	21	Gil Algébrico	PA
16	Lógica	32	Topo Lógico	AE
...	...	...	...	...

# Decomposição na 2FN

---

- O esquema atual tem atualmente 3 relações
  - Aluno, Disciplina e Nota
  - Inicialmente era apenas 1 relação (Nota)
  - As 3 relações estão na 2FN ...
  - ... colunas não chave dependem da totalidade da chave
- Contudo, ainda existem anomalias na relação **Disciplina**



# Anomalias na Relação Disciplina

---

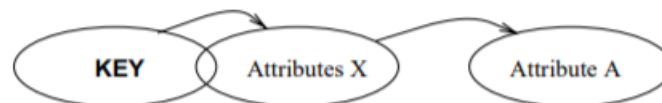
- **Inserção** de um novo professor exige que lhe seja atribuída pelo menos uma disciplina (a chave é n\_disciplina)
- **Remoção** de uma disciplina elimina a informação do professor se esse professor não dá aulas a outras disciplinas, então perde-se a sua informação
- **Atualização** do grau de um professor exige efetuar a alteração em todas as disciplinas atribuídas a esse professor

<u>n_disciplina</u>	nome disciplina	n professor	nome professor	grau professor
04	Álgebra	21	Gil Algébrico	PA
14	Análise Sist.	87	Ana Listada	PC
23	P.Linear	43	Plínio	AS
08	Topologia	32	Topo Lógico	AE
12	Geometria	21	Gil Algébrico	PA
16	Lógica	32	Topo Lógico	AE
...	...	...	...	...

# Principais Formas Normais

---

- **Primeira** forma normal: 1FN
  - Colunas da relação guardam apenas **um** valor por linha
  - Tipicamente verificado em bases de dados relacionais
- **Segunda** forma normal: 2FN = 1FN +
  - Colunas não pertencentes às chaves candidatas da relação...
  - ...dependem da **totalidade** das colunas de cada chave
  - Trivial, se chaves da relação tiverem apenas uma coluna
- **Terceira** forma normal: 3FN = 2FN +
  - Colunas não pertencentes às chaves candidatas...
  - ...dependem **apenas** das chaves candidatas
  - As relações com um único atributo não chave que estejam na 2FN estão sempre na 3FN

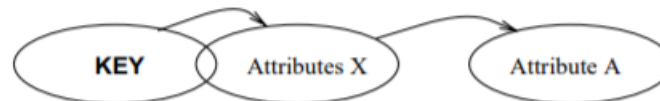
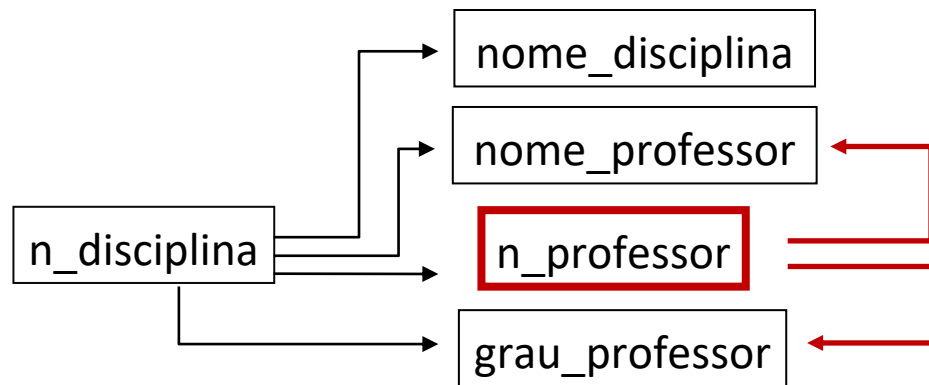


Case 1: A not in KEY

# Decomposição na 3FN

Disciplina (n\_disciplina, nome\_disciplina,  
n\_professor, nome\_professor, grau\_professor)

Existem as **DFs** entre um atributo não pertencente à chave (n\_professor) e atributos não chave  
cod\_professor → nome\_professor, grau\_professor

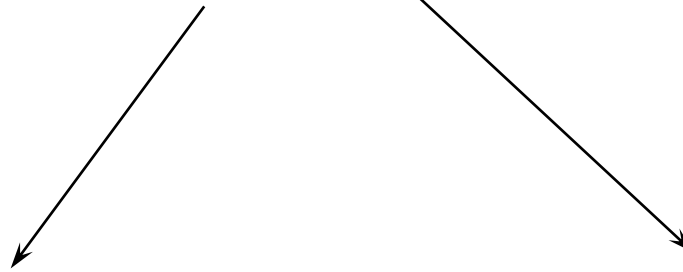


Case 1: A not in KEY

# Decomposição na 3FN

---

<u>n_ disciplina</u>	nome disciplina	cod professor	nome professor	grau professor
--------------------------	--------------------	------------------	-------------------	-------------------



## Disciplina

<u>n_ disciplina</u>	nome disciplina	cod professor
04	Álgebra	21
14	Análise Sist.	87
23	P.Linear	43
08	Topologia	32
12	Geometria	21
16	Lógica	32
...	...	...

## Professor

cod professor	nome professor	grau professor
21	Gil Algébrico	PA
87	Ana Listada	PC
43	Plínio	AS
32	Topo Lógico	AE
...	...	...

# Decomposição na 3FN

---

- Esquema atual tem 4 relações

Aluno (n\_aluno, nome, curso)

Nota (n\_aluno, n\_disciplina, nota)

Disciplina (n\_disciplina, nome\_disciplina, n\_professor)

Professor (n\_professor, nome\_professor, grau\_professor)

- Esquema relacional está na 3FN
  - Para cada relação, não existem dependências entre colunas não chave
  - Todas as dependências são das chaves candidatas para as outras colunas
- **Esquema normalizado é que seria traduzido para comandos SQL**