

# *Bases de Dados*

## normalização

**Pedro Furtado**

Departamento de Engenharia Informática  
Faculdade de Ciências e Tecnologia da Universidade de Coimbra  
2018/2019

# Universal Relation and Anomalies

paciente

Sintomas-  
diagnostico

First Name	Middle Initial	Last Name	Address	City	Zip Code	Country	Phone Number	Date of Birth	PatientID	Symptoms	Diagnosis	Already Scheduled?	Schedule
Joaquim	N.	Gomes	Rua Castilho, 23	Lisboa	2720-234	Portugal	91-99893	2/10/00	19845	Accute headache	Concussion	yes	20/6/2015, 10:00
Manuel	M.	Silva	Rua Augusta, 18	Coimbra	3000-542	Portugal	96-39494	2/10/90	21537	Sharp pain in lower abdomen	Food poisoning	no	-

# U. Relation and Anomalies

paciente

Sintomas-  
diagnostico

First Name	Middle Initial	Last Name	Address	City	Zip Code	Country	Phone Number	Date of Birth	PatientID	Symptoms	Diagnosis	Already Scheduled?	Schedule
Joaquim	N.	Gomes	Rua Castilho, 23	Lisboa	2720-234	Portugal	91-99893	2/10/00	19845	Accute headache	Concussion	yes	20/6/2015, 10:00
Manuel	M.	Silva	Rua Augusta, 18	Coimbra	3000-542	Portugal	96-39494	2/10/90	21537	Sharp pain in lower abdomen	Food poisoning	no	-



Deletion anomaly

se apago diagnostico, tb apago doente!!!  
Se apago doente, tb apago diagnostico

First Name	Middle Initial	Last Name	Address	City	Zip Code	Country	Phone Number	Date of Birth	PatientID	Symptoms	Diagnosis	Already Scheduled?	Schedule
Joaquim	N.	Gomes	Rua Castilho, 23	Lisboa	2720-234	Portugal	91-99893	2/10/00	19845	Accute headache	Concussion	yes	20/6/2015, 10:00
Manuel	M.	Silva	Rua Augusta, 18	Coimbra	3000-542	Portugal	96-39494	2/10/90	21537	Sharp pain in lower abdomen	Food poisoning	no	-

# U. Relation and Anomalies

First Name	Middle Initial	Last Name	Address	City	Zip Code	Country	Phone Number	Date of Birth	PatientID	Symptoms	Diagnosis	Already Scheduled?	Schedule
Joaquim	N.	Gomes	Rua Castilho, 23	Lisboa	2720-234	Portugal	91-99893	2/10/00	19845	Accute headache	Concussion	yes	20/6/2015, 10:00
Manuel	M.	Silva	Rua Augusta, 18	Coimbra	3000-542	Portugal	96-39494	2/10/90	21537	Sharp pain in lower abdomen	Food poisoning	no	-



Insertion Anomaly

Insertion anomaly  
insiro diagnostico, insiro de novo o doente!!!

First Name	Middle Initial	Last Name	Address	City	Zip Code	Country	Phone Number	Date of Birth	PatientID	Symptoms	Diagnosis	Already Scheduled?	Schedule
Joaquim	N.	Gomes	Rua Castilho, 23	Lisboa	2720-234	Portugal	91-99893	2/10/00	19845	Accute headache	Concussion	yes	20/6/2015, 10:00
Manuel	M.	Silva	Rua Augusta, 18	Coimbra	3000-542	Portugal	96-39494	2/10/90	21537	Sharp pain in lower abdomen	Food poisoning	no	-
Joaquim	N.	Gomes	Rua Castilho, 23	Lisboa	2720-234	Portugal	91-99893	2/10/00	19845	Pain in the mouth	Dental abscess	no	-

# U. Relation and Anomalies

First Name	Middle Initial	Last Name	Address	City	Zip Code	Country	Phone Number	Date of Birth	PatientID	Symptoms	Diagnosis	Already Scheduled?	Schedule
Joaquim	N.	Gomes	Rua Castilho, 23	Lisboa	2720-234	Portugal	91-99893	2/10/00	19845	Accute headache	Concussion	yes	20/6/2015, 10:00
Manuel	M.	Silva	Rua Augusta, 18	Coimbra	3000-542	Portugal	96-39494	2/10/90	21537	Sharp pain in lower abdomen	Food poisoning	no	-



Update anomaly

Joaquim mora onde??

Modification  
Anomaly

First Name	Middle Initial	Last Name	Address	City	Zip Code	Country	Phone Number	Date of Birth	PatientID	Symptoms	Diagnosis	Already Scheduled?	Schedule
Joaquim	N.	Gomes	Rua Murillo, 2	Coimbra	3000-210	Portugal	91-99893	2/10/00	19845	Accute headache	Concussion	yes	20/6/2015, 10:00
Manuel	M.	Silva	Rua Augusta, 18	Coimbra	3000-542	Portugal	96-39494	2/10/90	21537	Sharp pain in lower abdomen	Food poisoning	no	-
Joaquim	N.	Gomes	Rua Castilho, 23	Lisboa	2720-234	Portugal	91-99893	2/10/00	19845	Pain in the mouth	Dental abscess	no	-

# NORMALIZAÇÃO

# Normalização

- Normalização: processo que identifica a localização correcta de cada atributo e a estrutura correcta das tabelas existentes numa base de dados.
- Se uma tabela está mais normalizada isso significa que tem menos dados redundantes.
- A normalização de uma tabela é obtida pela divisão da tabela em duas ou mais tabelas, de acordo com um procedimento bem específico.

# Dependência funcional

- B é **funcionalmente dependente** de A se:
  - >a cada valor de A estiver associado um e um só valor de B

Representa-se por:  $A \rightarrow B$

- Diz-se que A é o **determinante** de B se  $A \rightarrow B$  e se B não for funcionalmente dependente de nenhum subconjunto de A.
- Exemplo: numCC -> nome da pessoa



# Diagramas de dependência funcional

- **Notação simbólica**

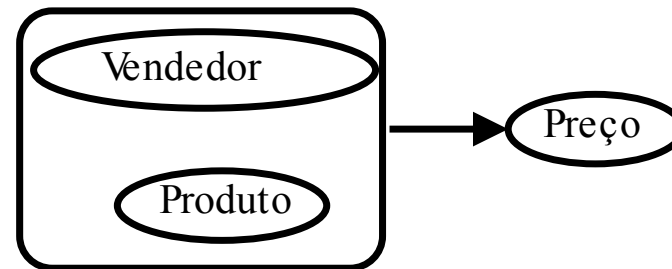
**$A \rightarrow B$**

- **Exemplo:**

$\{\text{Vendedor}, \text{Produto}\} \rightarrow \text{Preço}$

Porque é que pode acontecer que vendedor e produto determinem o preço, e não apenas produto?

- **Diagrama de dependência funcional**



# Identificação de dependências funcionais: exemplo

- A identificação das dependências funcionais **não se pode obter apenas a partir da observação de alguns registos (linhas específicas de tabelas)**, mas sim através das próprias propriedades dos atributos.

- Exemplo:

**Funcionarios**

N_funcionario	Nome	Apelido	Departamento
1021	António	Silva	900
1022	João	Reis	900
1023	Rui	Paiva	800
1024	Ana	Madeira	700
1025	Luís a	Saraiva	800

- O atributo Apelido é funcionalmente dependente de N\_funcionario.
  - Se se conhecer o número do funcionário (N\_funcionario, que é unívoco), então pode-se determinar o seu apelido (um funcionário só pode ter um apelido).
- **E o atributo N\_funcionario, será funcionalmente dependente de Apelido?**

# Dependências funcionais: outro exemplo

**Classificacao**

Nome_aluno	Numero	Disciplina	Data_matricula	Nota
António Silva	1234343	Português	12-03-2004	12
António Silva	1234343	Análise	12-03-2004	14
Joana Antunes	1275432	Álgebra	13-03-2004	10
Joana Antunes	1275432	Português	13-03-2004	11

**Nota é funcionalmente dependente de quê?**

- Nota é funcionalmente dependente de Numero?
- Nota é funcionalmente dependente de Disciplina?
- Nota depende funcionalmente de Nome\_aluno e Disciplina?
- Nota depende funcionalmente de Numero e Disciplina?

**{Numero, Disciplina} → Nota ?**

# NORMALIZAÇÃO evita dados redundantes

Nome	BI	Morada	N_Telefone	Cidade
António Silva	1234343	'Rua xpto'	9002233	Coimbra
Joaquim Alves Dias	1256673	'Rua Ypto'	5563452	Guarda
Maria Teresa Horta	1275432	'Rua Zpto'	9005335	Lisboa
Joana Antunes	1022634	'Rua Tpto'	8013454	Lisboa
Luísa Saraiva	1342664	'Rua Upto'	6474354	Lisboa

UC

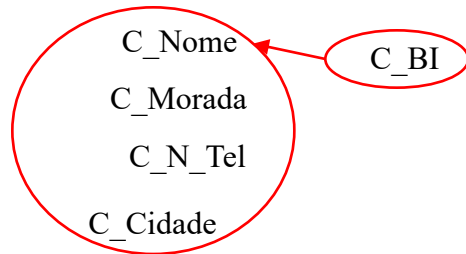
## Hierarquia de funções:

Nome	...	ChefeNome	ChefeBI	ChefeMorada	ChefeTel	ChefeCidade
António Silva	...	Joana Antunes	1022634	'Rua Tpto'	8013454	Lisboa
Joaquim Alves Dias	...	Joana Antunes	1022634	'Rua Tpto'	8013454	Lisboa
Maria Teresa Horta	...	Joana Antunes	1022634	'Rua Tpto'	8013454	Lisboa
Joana Antunes	...	?				
Luísa Saraiva	...					

bd

# Eliminar Dados redundantes

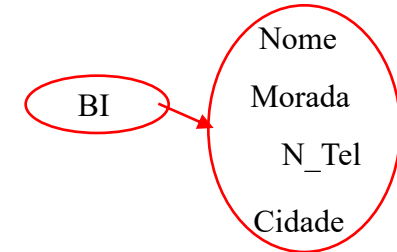
## Chefe



## Chefe\_Emp



## Empregado



## Empregado e Chefe

Nome	BI	Morada	N_Telefone	Cidade
António Silva	1234343	'Rua xpto'	9002233	Coimbra
Joaquim Alves Dias	1256673	'Rua Ypto'	5563452	Guarda
Maria Teresa Horta	1275432	'Rua Zpto'	9005335	Lisboa
Joana Antunes	1022634	'Rua Tpto'	8013454	Lisboa
Luísa Saraiva	1342664	'Rua Upto'	6474354	Lisboa

BI	C_BI
1234343	1022634
1256673	1022634
1275432	1022634
1022634	
1342664	

# Normalização

- Normalização: processo que identifica a localização correcta de cada atributo e a estrutura correcta das tabelas existentes numa base de dados.
- Se uma tabela está mais normalizada, isso significa que tem menos dados redundantes.
- A normalização de uma tabela é obtida pela divisão da tabela em duas ou mais tabelas, de acordo com um procedimento bem específico.

# 1a Forma Normal (1FN)

- Uma relação está na 1a Forma Normal (1FN) quando:
  - os domínios de todos os seus atributos contêm apenas valores atômicos.
  - não há conjuntos de atributos repetidos descrevendo a mesma característica.

**ERROS, NAO 1FN:**

Nome	BI	Morada	Telefones	Cidade
António Silva	1234343	'Rua xpto'	9002233, 4333232, 566564333	Coimbra
Joaquim Alves Dias	1256673	'Rua Ypto'	5563452, 34343434	Guarda

Nome	BI	Morada	Telefone1	Telefone2	Telefone3	Cidade
António Silva	1234343	'Rua xpto'	9002233	4333232	566564333	Coimbra
Joaquim Alves Dias	1256673	'Rua Ypto'	5563452	34343434		Guarda

# 1FN não é suficiente porque ...

- Pode conter informação redundante, pelo que as operações (inserção, actualização e apagamento) sobre a base de dados podem levar à perda de integridade. Há três tipos de problemas:
  - Anomalias de inserção;
  - Anomalias de actualização;
  - Anomalias de apagamento.



## 2a Forma Normal (2NF)

- Uma tabela está na 2FN quando estiver na 1FN e os atributos que não são chave **dependerem inteiramente da chave**.
- Exemplo: consideremos a tabela Mestrando (que já está na 1FN)**

Mestrando

Nº_aluno	Orientador	Gabinete	Disciplina	Nota
1022	Dr. Ferreira	A23	Física	15
1022	Dr. Ferreira	A23	Química	12
1022	Dr. Ferreira	A23	História	16
4123	Dr. Altino	B12	Português	12
4123	Dr. Altino	B12	Inglês	14
4123	Dr. Altino	B12	Filosofia	11
7239	Dr. Ferreira	A23	Química	16

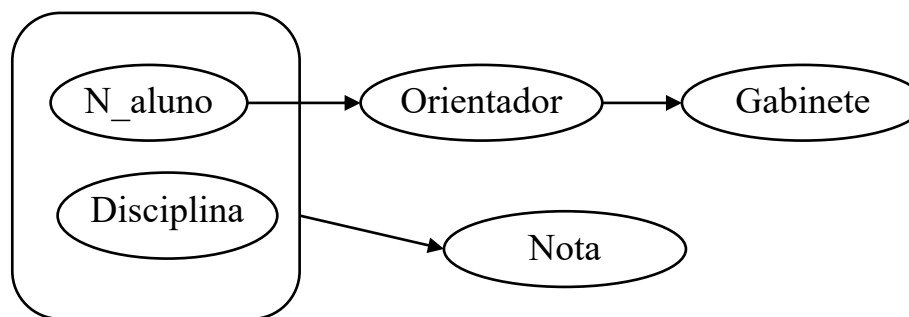
- Para verificarmos se a tabela Mestrando está na 2FN, perguntamos:
- 1) **Qual a chave desta tabela?** Se a chave for apenas um atributo então está na 2FN. Se a chave for concatenada (mais do que um atributo), perguntamos:
- 2) Há atributos que não são chave e que dependem apenas de parte da chave? Se não, está na 2FN.

# Não 2a Forma Normal (2NF)

Mestrando

Nº_aluno	Orientador	Gabinete	Disciplina	Nota
1022	Dr. Ferreira	A23	Física	15
1022	Dr. Ferreira	A23	Química	12
1022	Dr. Ferreira	A23	História	16
4123	Dr. Altino	B12	Português	12
4123	Dr. Altino	B12	Inglês	14
4123	Dr. Altino	B12	Filosofia	11
7239	Dr. Ferreira	A23	Química	16

- Chave **{No\_aluno, Disciplina}**
  - Porque não pode ser só **No\_aluno**?



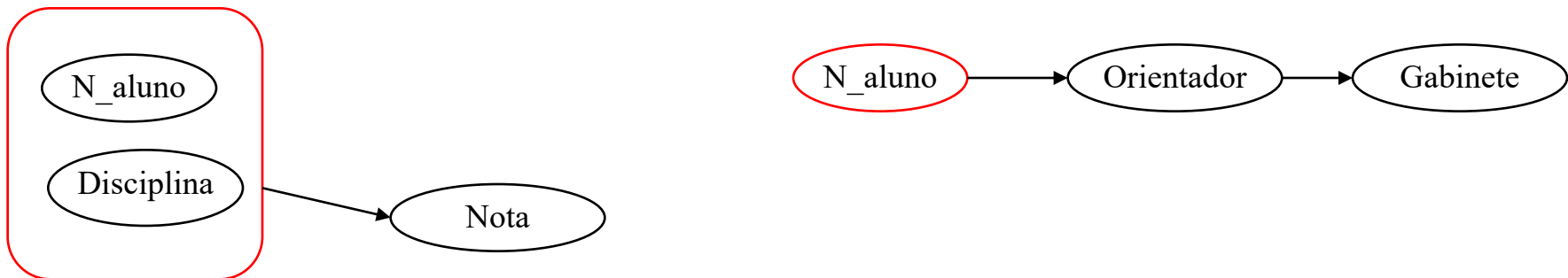
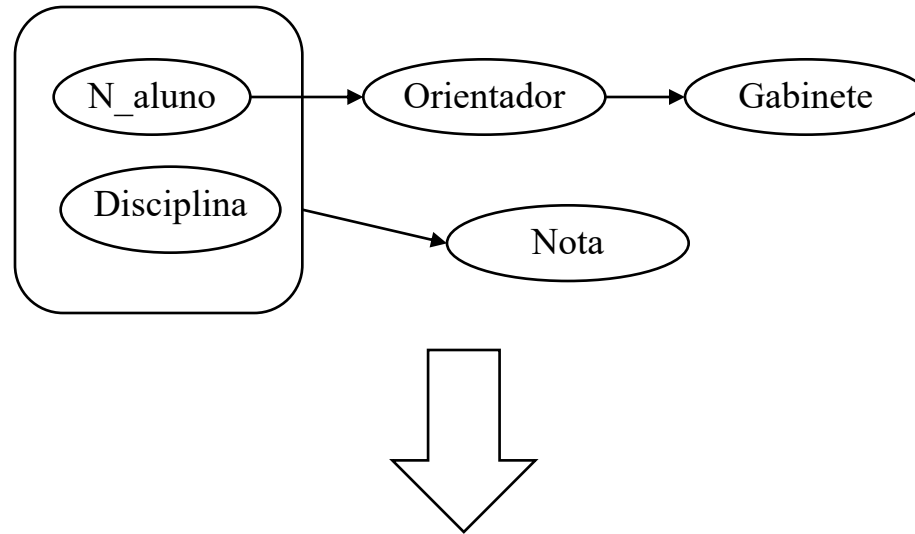
# Não 2a Forma Normal (2NF)

**Mestrando**

Nº_aluno	Orientador	Gabinete	Disciplina	Nota
1022	Dr. Ferreira	A23	Física	15
1022	Dr. Ferreira	A23	Química	12
1022	Dr. Ferreira	A23	História	16
4123	Dr. Altino	B12	Português	12
4123	Dr. Altino	B12	Inglês	14
4123	Dr. Altino	B12	Filosofia	11
7239	Dr. Ferreira	A23	Química	16

- Há atributos que só dependam de parte da chave?
  - Orientador depende do No\_aluno (embora Nota dependa de toda a chave)
- **Conclusão: Não está na 2FN**

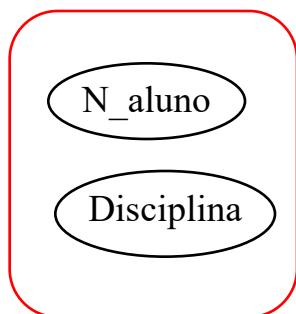
# 2a Forma Normal (2NF)



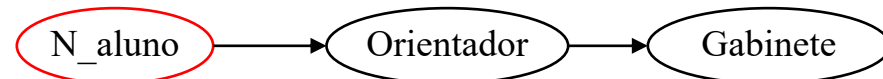
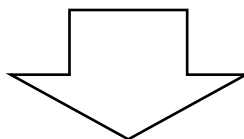
# 2a Forma Normal (2NF)

**Mestrando**

Nº_aluno	Orientador	Gabinete	Disciplina	Nota
1022	Dr. Ferreira	A23	Física	15
1022	Dr. Ferreira	A23	Química	12
1022	Dr. Ferreira	A23	História	16
4123	Dr. Altino	B12	Português	12
4123	Dr. Altino	B12	Inglês	14
4123	Dr. Altino	B12	Filosofia	11
7239	Dr. Ferreira	A23	Química	16



Nota



**Aluno\_disciplina**

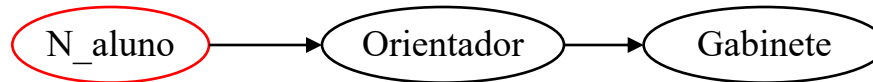
Nº_aluno	Disciplina	Nota
1022	Física	15
1022	Química	12
1022	História	16
4123	Português	12
4123	Inglês	14
4123	Filosofia	11
7239	Química	16

**Aluno\_orientador**

Nº_aluno	Orientador	Gabinete
1022	Dr. Ferreira	A23
4123	Dr. Altino	B12
7239	Dr. Ferreira	A23

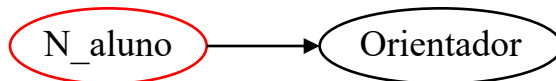
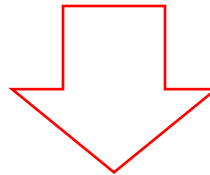
# 3a Forma Normal (3FN)

- Uma tabela está na 3FN quando estiver na 2FN e nenhum atributo que não seja chave depender de outro que também não seja chave.



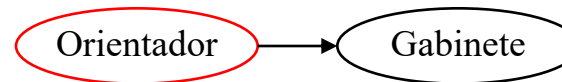
**Aluno\_orientador**

Nº_aluno	Orientador	Gabinete
1022	Dr. Ferreira	A23
4123	Dr. Altino	B12
7239	Dr. Ferreira	A23



**Aluno\_orientador**

Nº_aluno	Orientador
1022	Dr. Ferreira
4123	Dr. Altino
7239	Dr. Ferreira



**Orientador\_gabinete**

Orientador	Gabinete
Dr. Ferreira	A23
Dr. Altino	B12

# Processo de Normalização

**Mestrando**

Nº_aluno	Orientador	Gabinete	Disciplina	Nota
1022	Dr. Ferreira	A23	Física	15
1022	Dr. Ferreira	A23	Química	12
1022	Dr. Ferreira	A23	História	16
4123	Dr. Altino	B12	Português	12
4123	Dr. Altino	B12	Inglês	14
4123	Dr. Altino	B12	Filosofia	11
7239	Dr. Ferreira	A23	Química	16

**Aluno\_disciplina**

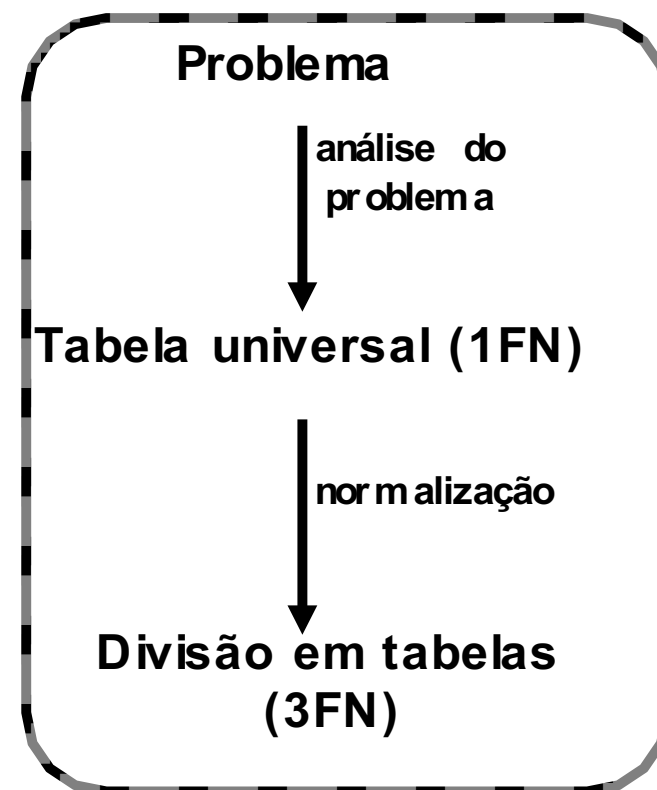
Nº_aluno	Disciplina	Nota
1022	Física	15
1022	Química	12
1022	História	16
4123	Português	12
4123	Inglês	14
4123	Filosofia	11
7239	Química	16

**Aluno\_orientador**

Nº_aluno	Orientador
1022	Dr. Ferreira
4123	Dr. Altino
7239	Dr. Ferreira

**Orientador\_gabinete**

Orientador	Gabinete
Dr. Ferreira	A23
Dr. Altino	B12



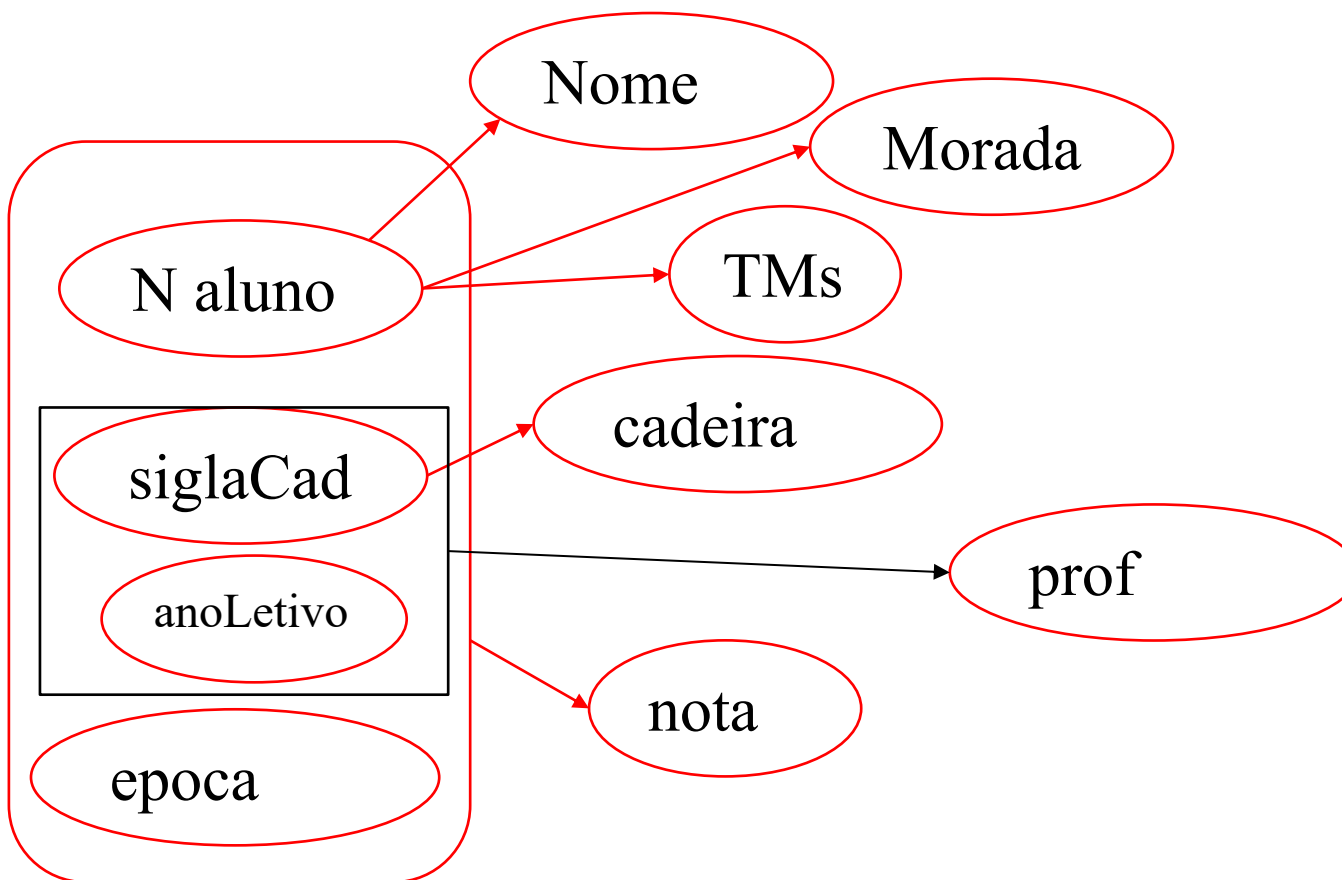
# Exemplo

Notas: naluno, nome, morada, TM, siglaCad, cadeira, prof, ano, epoca, nota



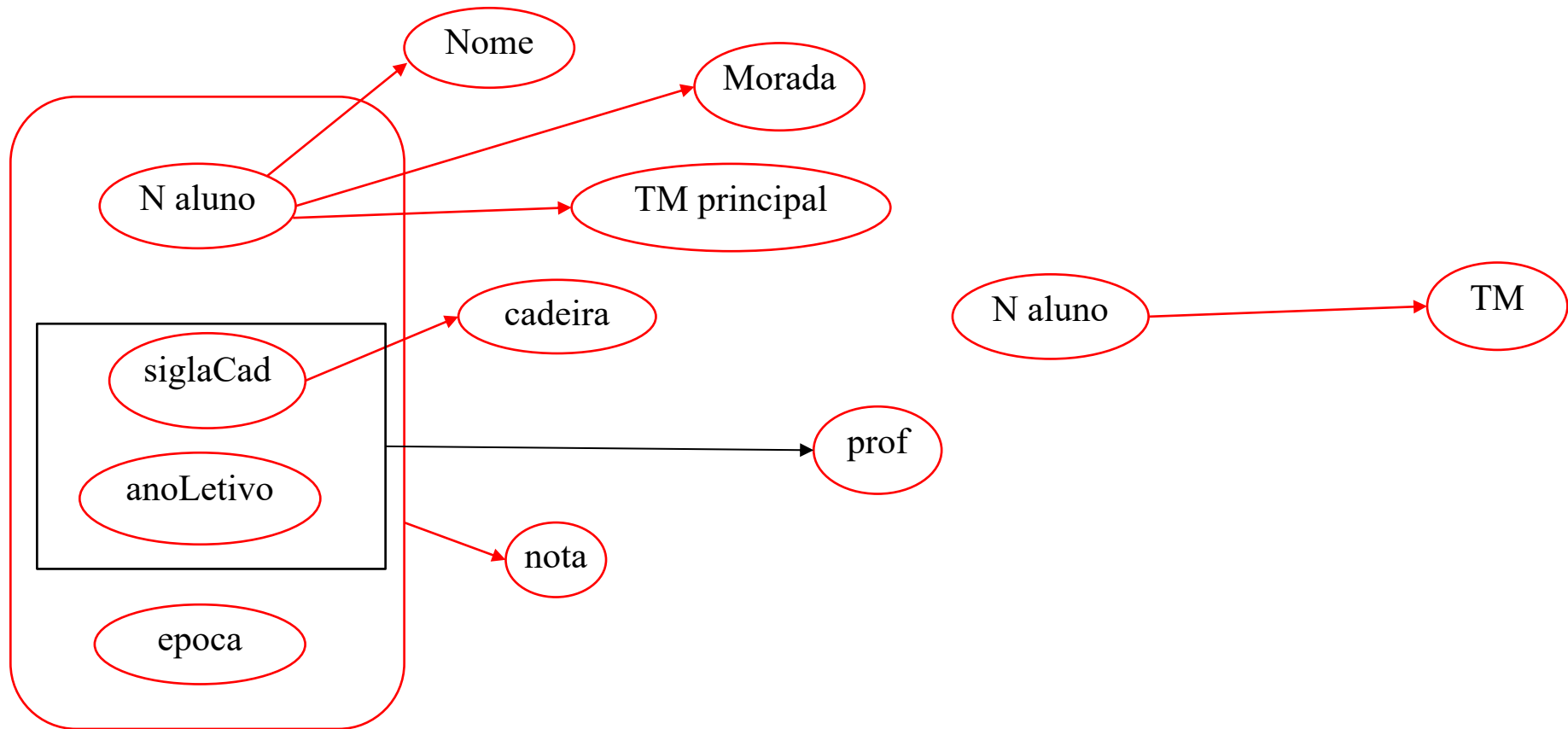
# Exemplo

Notas: naluno, nome, morada, TMs, siglaCad, cadeira, prof, ano, epoca, nota



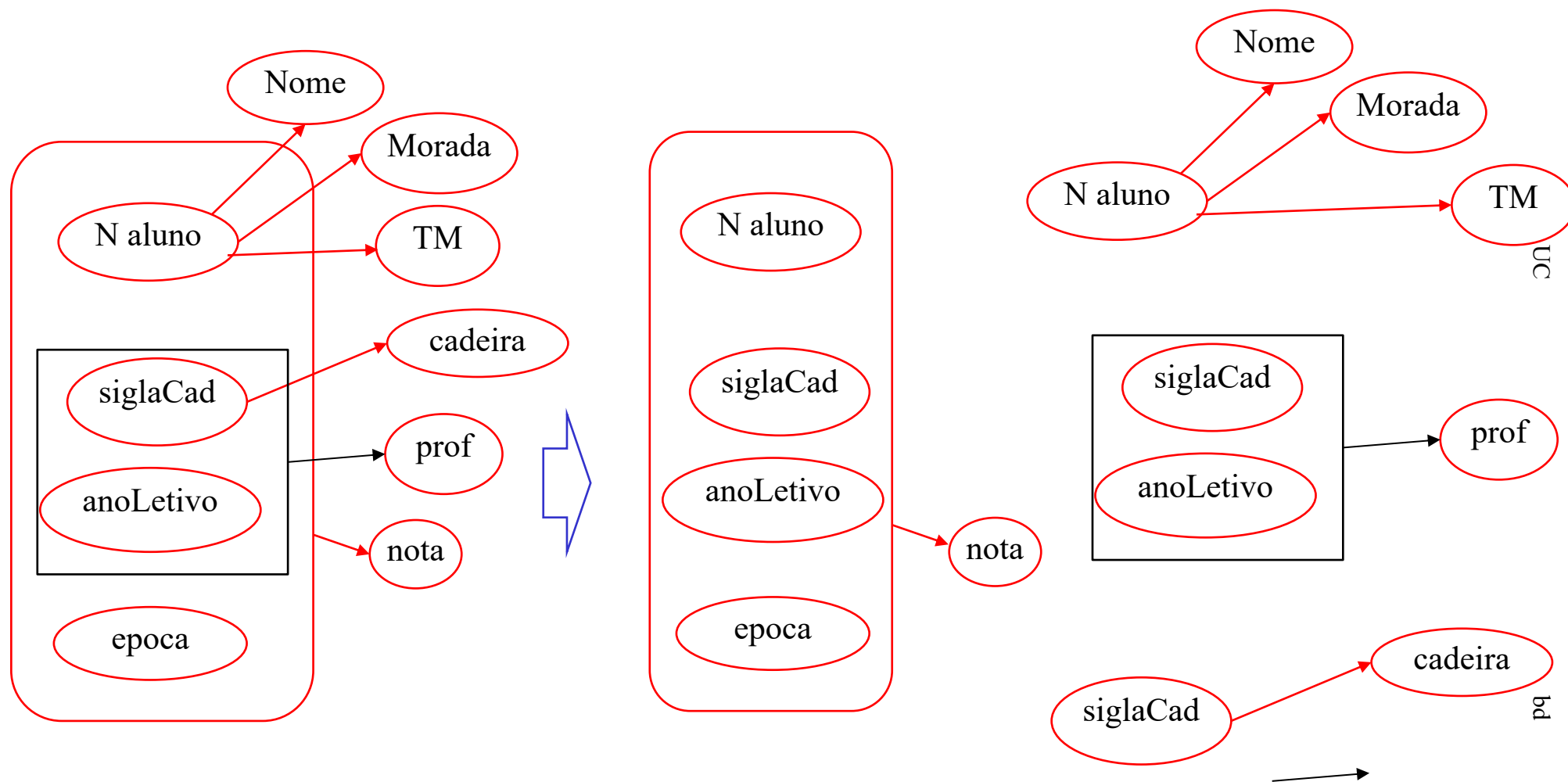
# Exemplo: 1 FN attrs atomicos

Notas: naluno, nome, morada, TMs, siglaCad, cadeira, prof, ano, epoca, nota



# Exemplo: 2 FN non-key depende totalmente de key

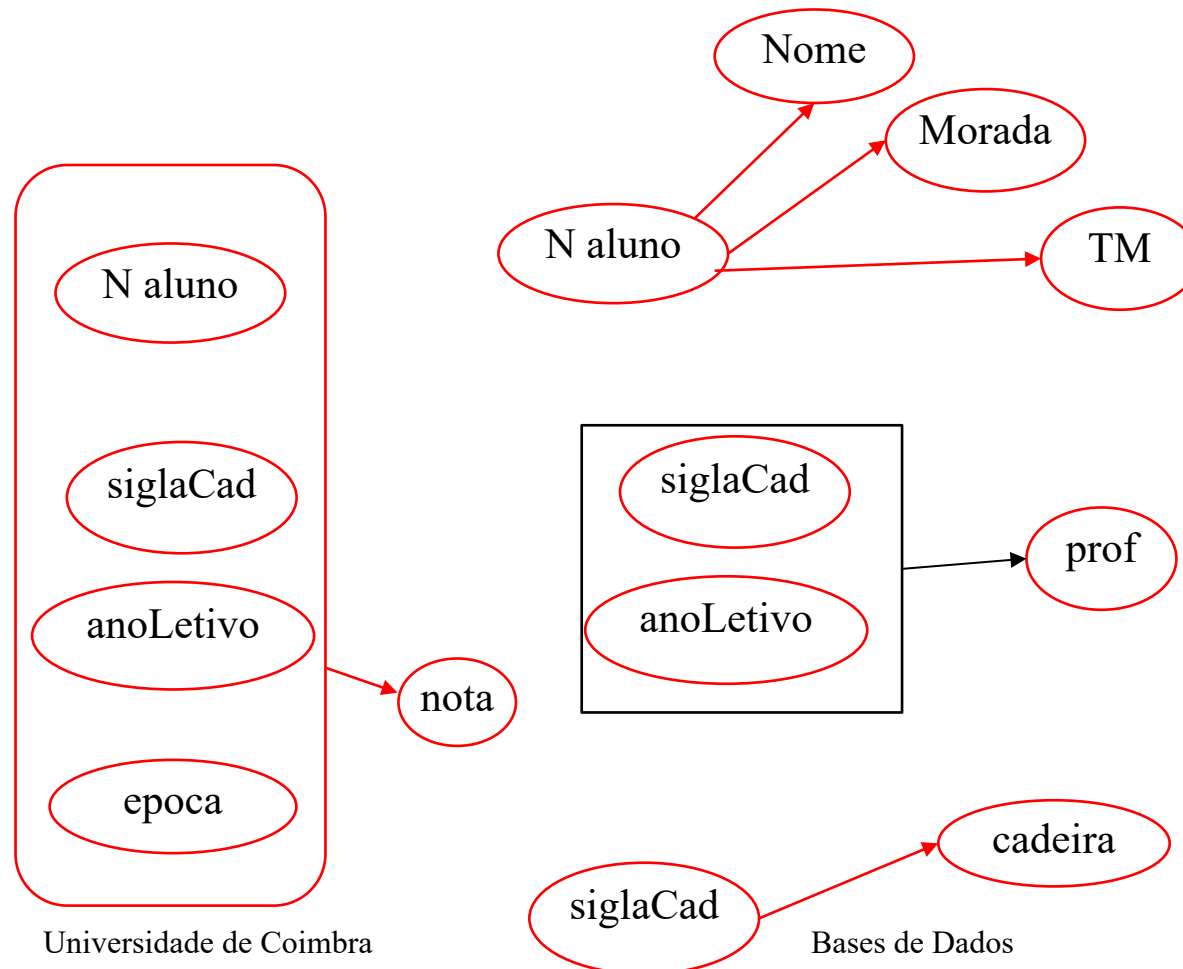
Notas: naluno, nome, morada, TM, siglaCad, cadeira, prof, ano, epoca, nota



# Exemplo: 3 FN non-key nao depende de nao key

Notas: naluno, nome, morada, TMs, siglaCad, cadeira, prof, ano, epoca, nota

**JA ESTAVA**



# Dependências transitivas

- Se  $A \rightarrow B$  e  $B \rightarrow C$ , então  $A \rightarrow C$  é uma dependência transitiva.

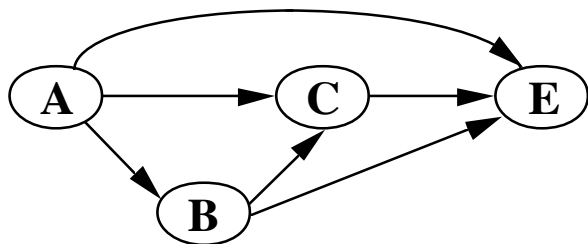
Notar que:

$A \rightarrow C$  é uma DF válida;

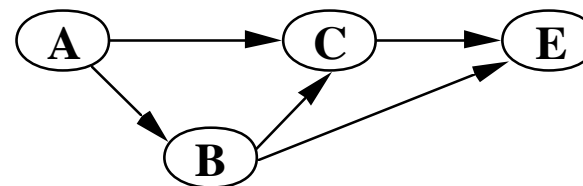
Se  $A \rightarrow B$ ,  $B \rightarrow C$  e  $A \rightarrow C$  pertencem todas a um mesmo conjunto de DFs, então  $A \rightarrow C$  é redundante.

- As DFs redundantes devem ser eliminadas antes de iniciar o processo de decomposição.

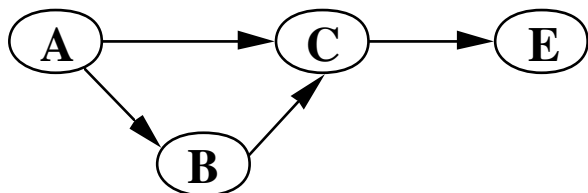
# Exemplo de remoção DF-transitivas



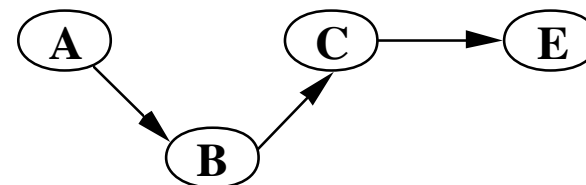
1) Conjunto original de DFs



2)  $A \rightarrow E$  é removida, uma vez que  $A \rightarrow C$  e  $C \rightarrow E$ .



3)  $B \rightarrow E$  é removida, uma vez que  $B \rightarrow C$  e  $C \rightarrow E$ .

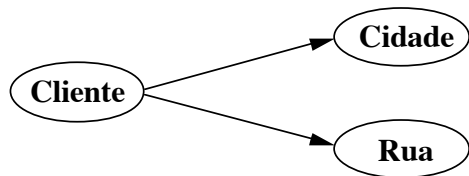


4)  $A \rightarrow C$  é removida, uma vez que  $A \rightarrow B$  e  $B \rightarrow C$ .

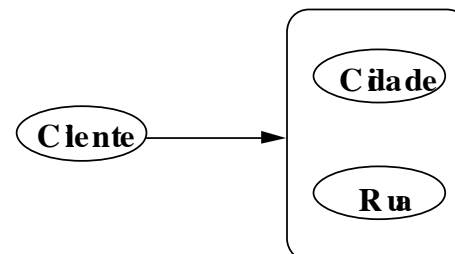
# Junção e decomposição de DFs

Se  $A \rightarrow B$  e  $A \rightarrow C$ , então  $A \rightarrow B, C$ .

Se:

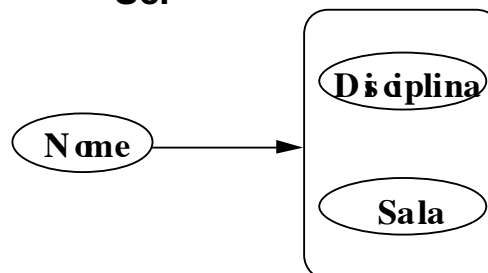


Então:

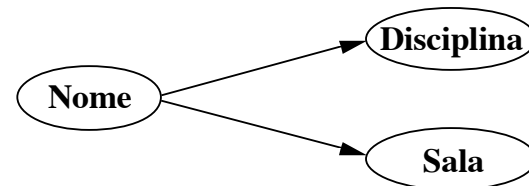


Se  $A \rightarrow B, C$  então  $A \rightarrow B$  e  $A \rightarrow C$ .

Se:



Então:



# Como se usa DF e normalização na prática?

- Normalmente vocês projectam as tabelas (relações)
- Pensam de forma a garantir 1FN, 2FN, 3FN
- Verificam se está na 3FN
- Se não estiver, normalizam para a 3FN
- No exame, se o professor pedir, têm de explicitar o processo de obtenção das 3 FN a partir de relação universal...

... para provarem que perceberam o conceito essencial de DF e normalização



# Exemplo para fazermos

Paciente( NúmeroDeCartãoCidadão, primeiroNome, ultimoNome, morada, localidade, DataDeVisitaHospital, sintoma, diagnostico, medicoNumero, medicoNome, especialidadeDoMedico)

- Paciente( **NúmeroDeCartãoCidadão**, primeiroNome, ultimoNome, morada)
- Morada(**morada**, localidade)
- VisitaHospital(**NúmeroDeCartãoCidadão**, **DataHoraDeVisitaHospital**, sintoma, diagnostico, medicoNumero)
- Medico(**medicoNumero**, medicoNome, especialidadeDoMedico)

DF

FIM