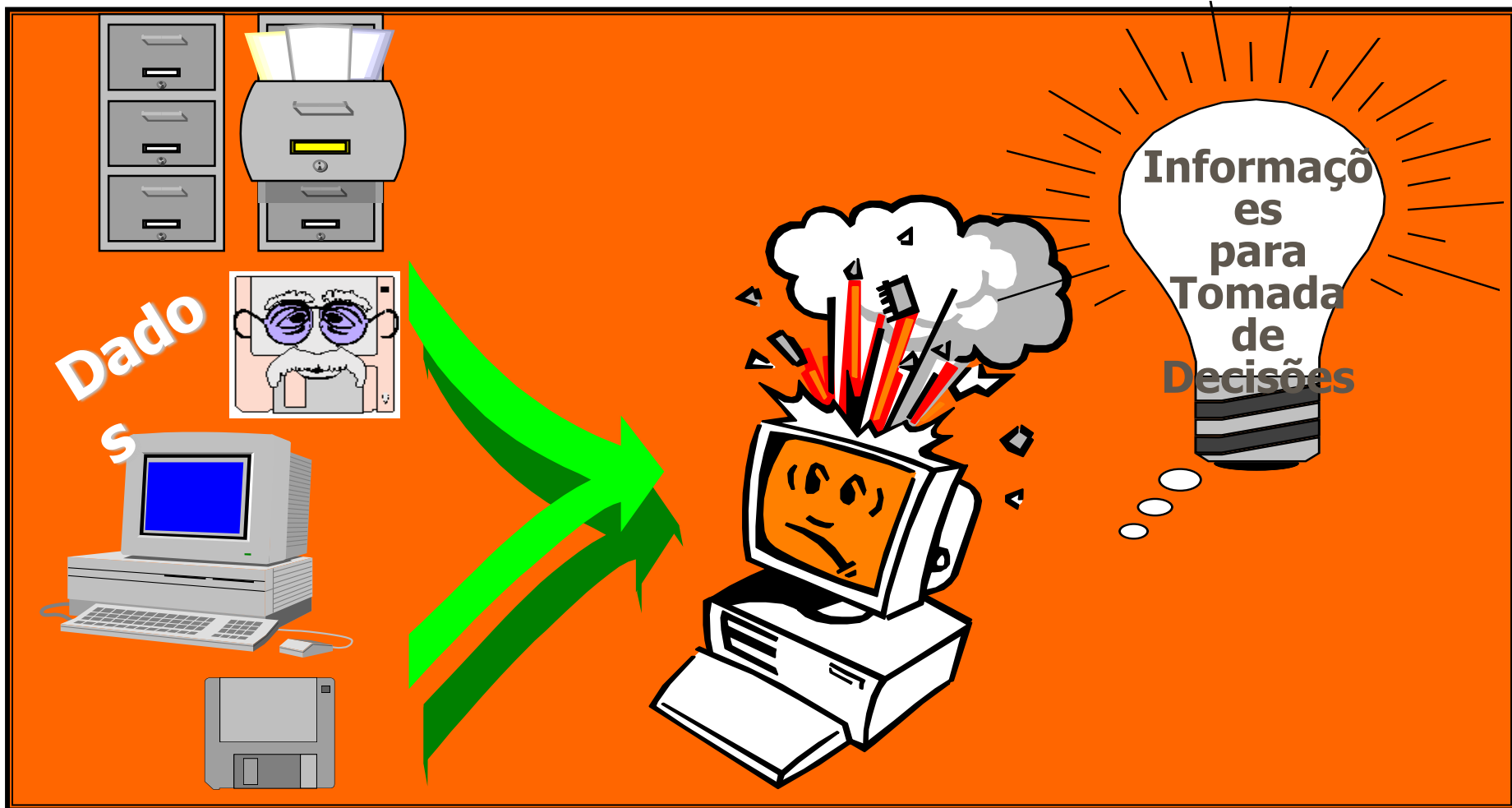
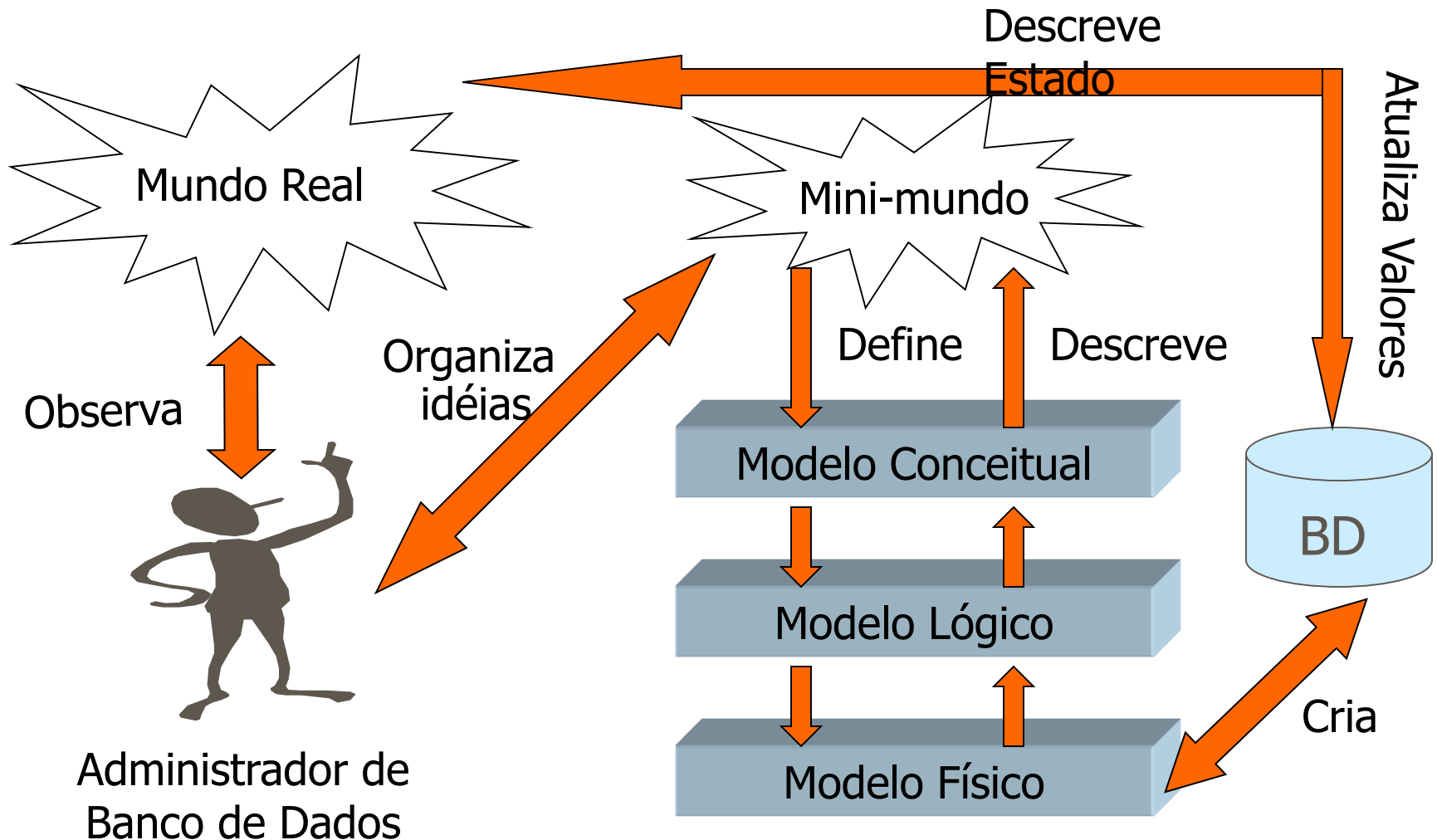


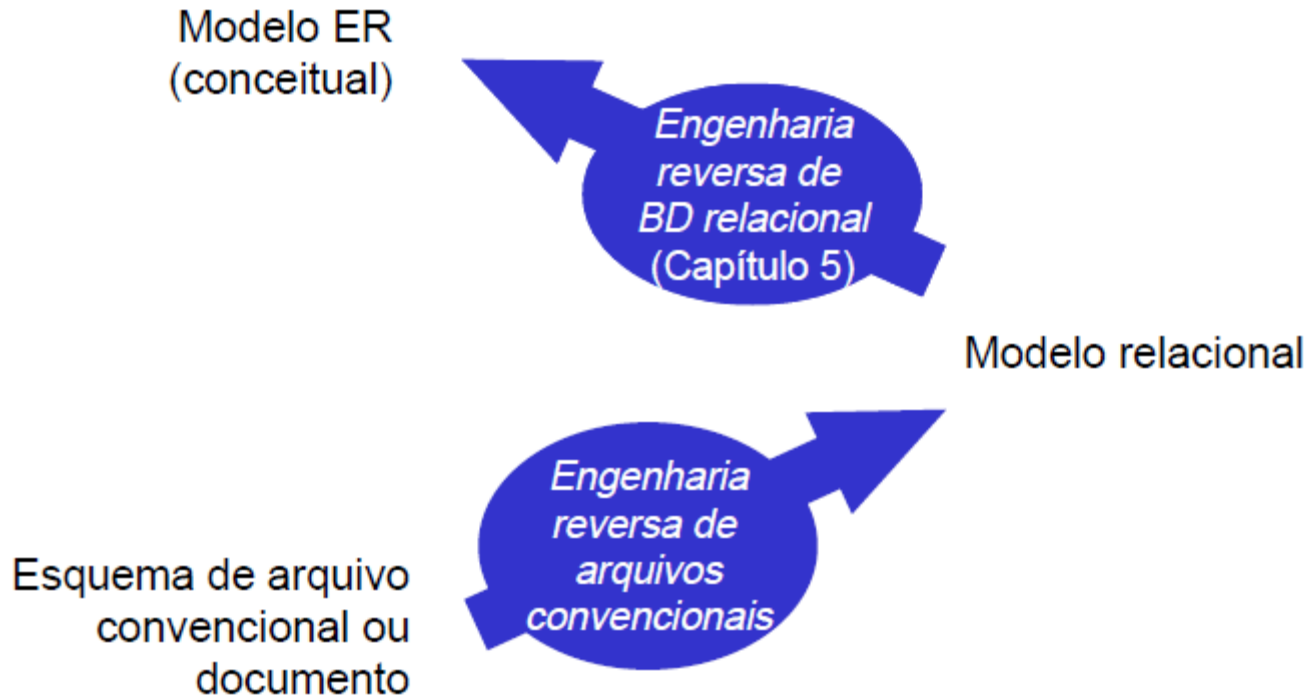
Bancos de Dados



Deborah Ribeiro Carvalho
2019

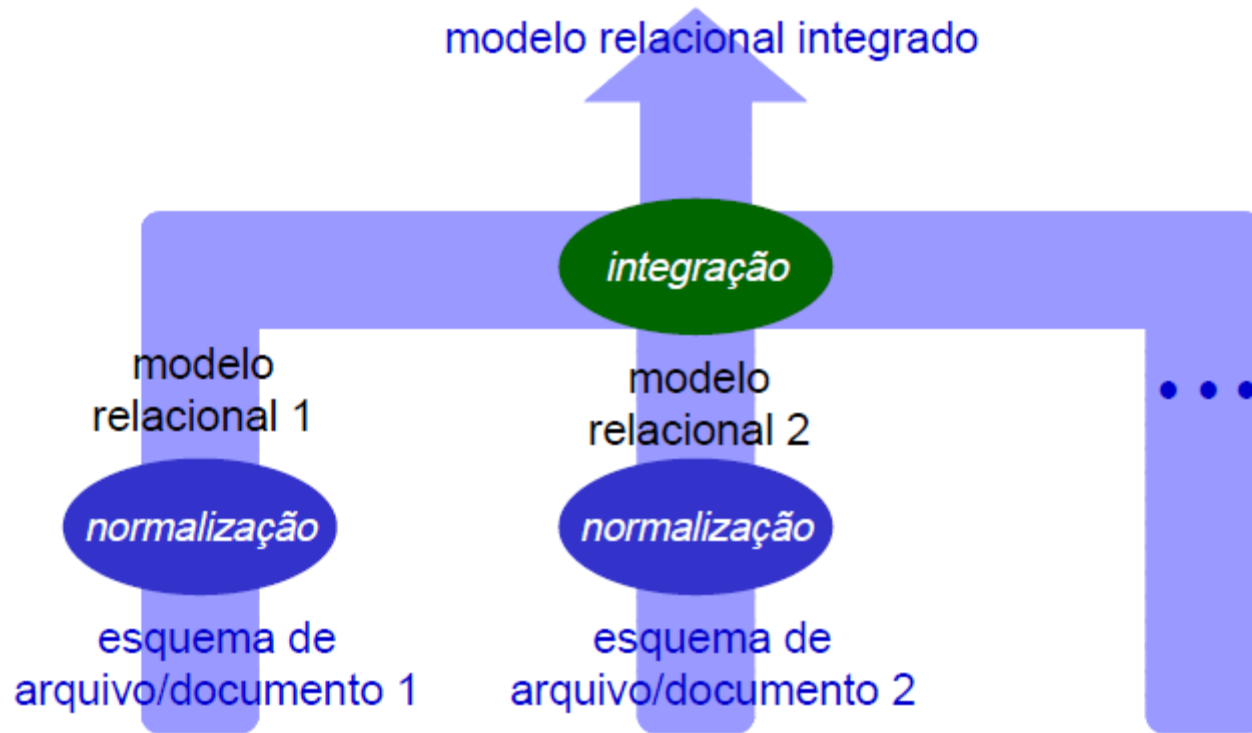
Visão Geral de SGBD





- **Entrada:**
 - **qualquer conjunto de dados para os quais se disponha de uma descrição**
 - documentos
 - arquivos manuais
 - arquivos convencionais em computador
 - bancos de dados gerenciados por **SGBD não relacional**

- **Sistemas legados**
- **Raramente documentados**
- **Necessidade de modelo ER**
 - **Manutenção**
 - **Migração para outro tipo de BD**
 - **Integração com outros BD**



- relação não normalizada: contém tuplos cujos valores são conjuntos, i.e., não atómicos

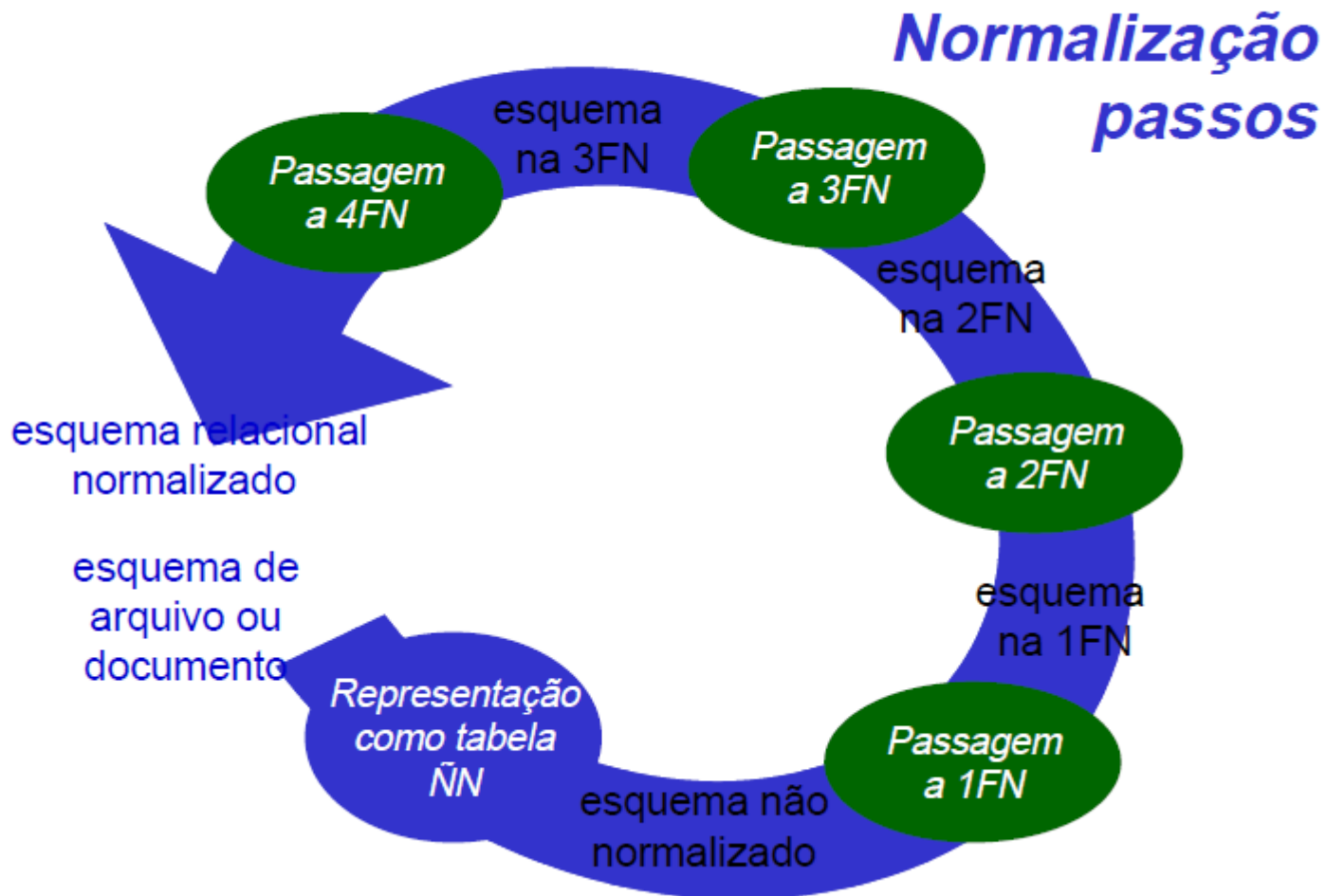
- não normalizada

$$R = \{ (r, \{b,a\}), (c, \{b,v\}), (t,n) \}$$

- normalizada correspondente

$$R = \{ (r,b), (r,a), (c,b), (c,v), (t,n) \}$$

- **Reagrupar informações para**
 - **eliminar redundâncias de dados**
- **Reagrupar informações para**
 - **eliminar estruturas inexistentes no modelo ER (atributos multivalorados)**



Documento exemplo

RELATÓRIO DE ALOCAÇÃO A PROJETO

CÓDIGO DO PROJETO: LSC001

TIPO: Novo Desenv.

DESCRIÇÃO: Sistema de Estoque

CÓDIGO DO EMPREGADO	NOME	CATEGORIA FUNCIONAL	SALÁRIO	DATA DE INÍCIO NO PROJETO	TEMPO ALOCADO AO PROJETO
2146	João	A1	4	1/11/91	24
3145	Silvio	A2	4	2/10/91	24
6126	José	B1	9	3/10/92	18
1214	Carlos	A2	4	4/10/92	18
8191	Mário	A1	4	1/11/92	12

CÓDIGO DO PROJETO: PAG02

TIPO: Manutenção

DESCRIÇÃO: Sistema de RH

CÓDIGO DO EMPREGADO	NOME	CATEGORIA FUNCIONAL	SALÁRIO	DATA DE INÍCIO NO PROJETO	TEMPO ALOCADO AO PROJETO
8191	Mário	A1	4	1/05/93	12
4112	João	A2	4	4/01/91	24
6126	José	B1	9	1/11/92	12

- Tabela *não-normalizada* ou
- tabela *não-primeira-forma-normal*
 - possui uma ou mais *tabelas aninhadas*
 - tabela aninhada (ou grupo repetido ou coluna multi-valorada oucoluna não atômica)
 - coluna que ao invés de conter valores atômicos, contém tabelas aninhadas
- Abreviatura: **ÑN**

Tabela aninhada exemplo

CódProj	Tipo	Descr	Emp					
			CodEmp	Nome	Cat	Sal	DataIni	TempAl
LSC001	Novo Desenv.	Sistema de Estoque	2146	João	A1	4	1/11/91	24
			3145	Sílvio	A2	4	2/10/91	24
			6126	José	B1	9	3/10/92	18
			1214	Carlos	A2	4	4/10/92	18
			8191	Mário	A1	4	1/11/92	12
PAG02	Manutenção	Sistema RH	8191	Mário	A1	4	1/05/93	12
			4112	João	A2	4	4/01/91	24
			6126	José	B1	9	1/11/92	12

Tabela aninhada em uma linha de projeto

***Tabela ÑN
Esquema***

Proj (CodProj, Tipo, Descr,
(CodEmp, Nome, Cat, Sal, DataIni, TempAl))

Arquivo em Pascal

```
type reg_aluno= record
  cod_al: integer;
  nome_al: char_60;
  ingressos_cursos_al: array [1..10] of record
    cod_curso: integer;
    semestre_ingresso: integer
  end;
  disciplinas_cursadas_al: array [0..200] of record
    cod_disc: integer;
    semestres_cursados: array [1..20] of record
      semestre_disc: integer;
      nota_disc: integer
    end
  end
end;
arq_aluno= file of reg_aluno;
```

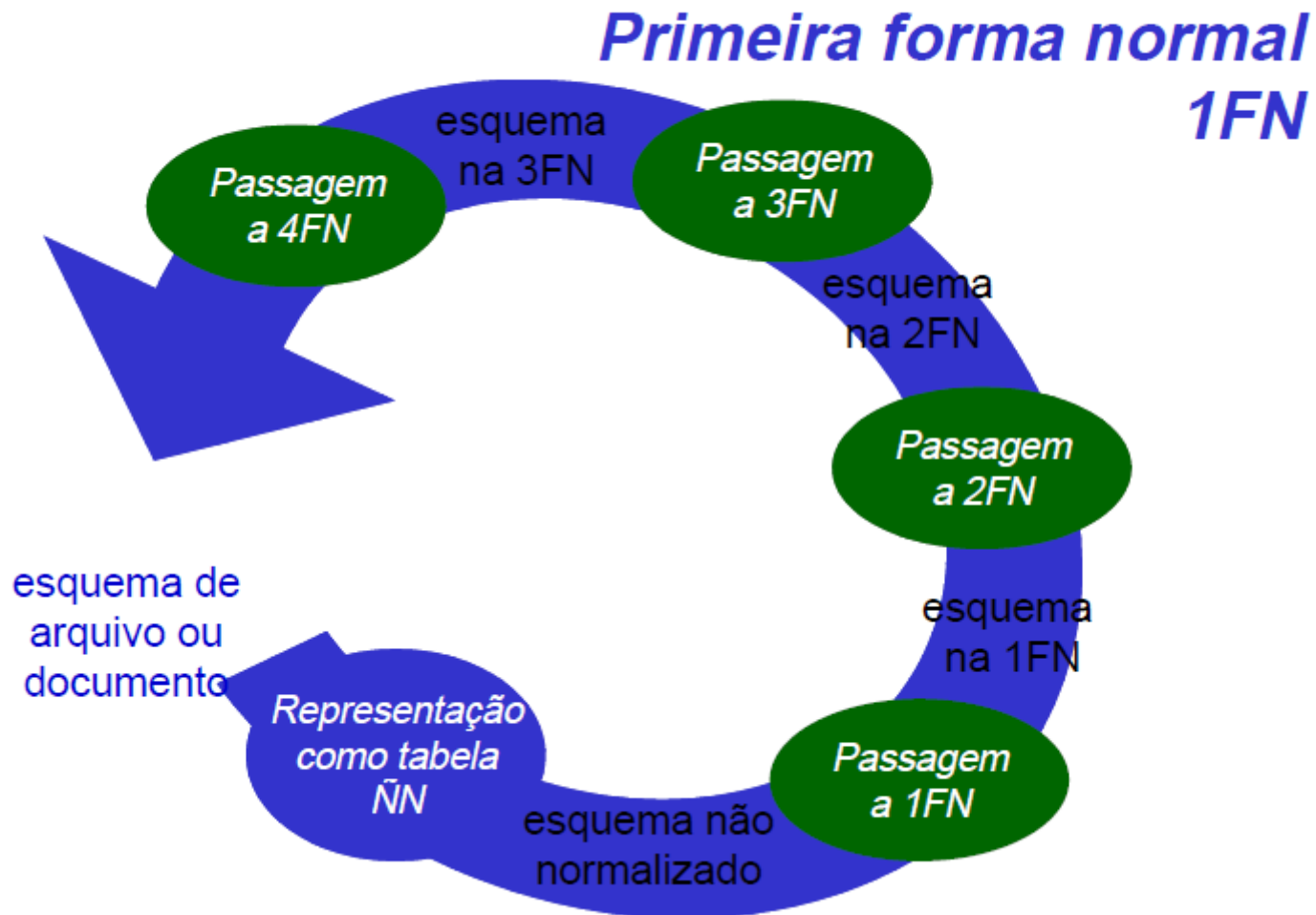
Arquivo em Cobol

```
FD   Arq-Alunos
01   Reg-Al.
      03   Cod-Al
      03   Nome-Al
      03   Ingr-Cursos-al OCCURS 1 TO 10
            05   Cod-Curso
            05   Sem-ingresso
      03   Disc-Curs-Al OCCURS 0 to 200
            05   Cod-Disc
            05   Sem-Cursado OCCURS 1 TO 20
                  07   Sem-Disc-Cursada
                  07   Nota-Disc
```

*Esquema ÑN
para arquivos exemplo*

Arq-Alunos (Cod-Al, Nome-Al,
 (Cod-Curso, Sem-ingresso)
 (Cod-Disc,
 (Sem-Disc-Cursada, Nota-Disc)))

- **Regra que uma tabela deve obedecer por para ser considerada “bem projetada”**
- **Há diversas formas normais, cada vez mais rígidas, para verificar tabelas relacionais**
- **Aqui tratadas**
 - primeira forma normal (1FN)
 - segunda forma normal (2FN)
 - terceira forma normal (3FN)
 - quarta forma normal (4FN)

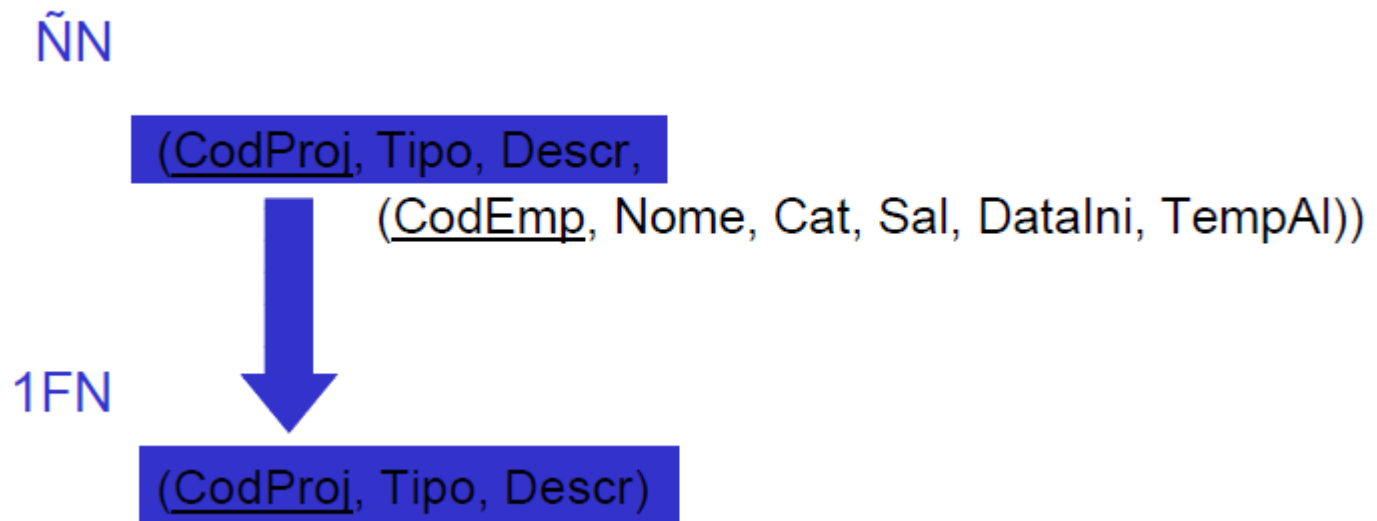


primeira forma normal (1FN)

=

**diz-se que uma tabela está na primeira
forma normal, quando ela não contém
tabelas aninhadas**

- Criar uma tabela na 1FN referente a tabela não normalizada
- A chave primária da tabela na 1FN é idêntica a chave da tabela ÑÑ



- **Para cada tabela aninhada**
 - **criar uma tabela na 1FN composta pelas seguintes colunas:**
 - a chave primária de cada uma das tabelas na qual a tabela em questão está aninhada
 - as colunas da própria tabela aninhada

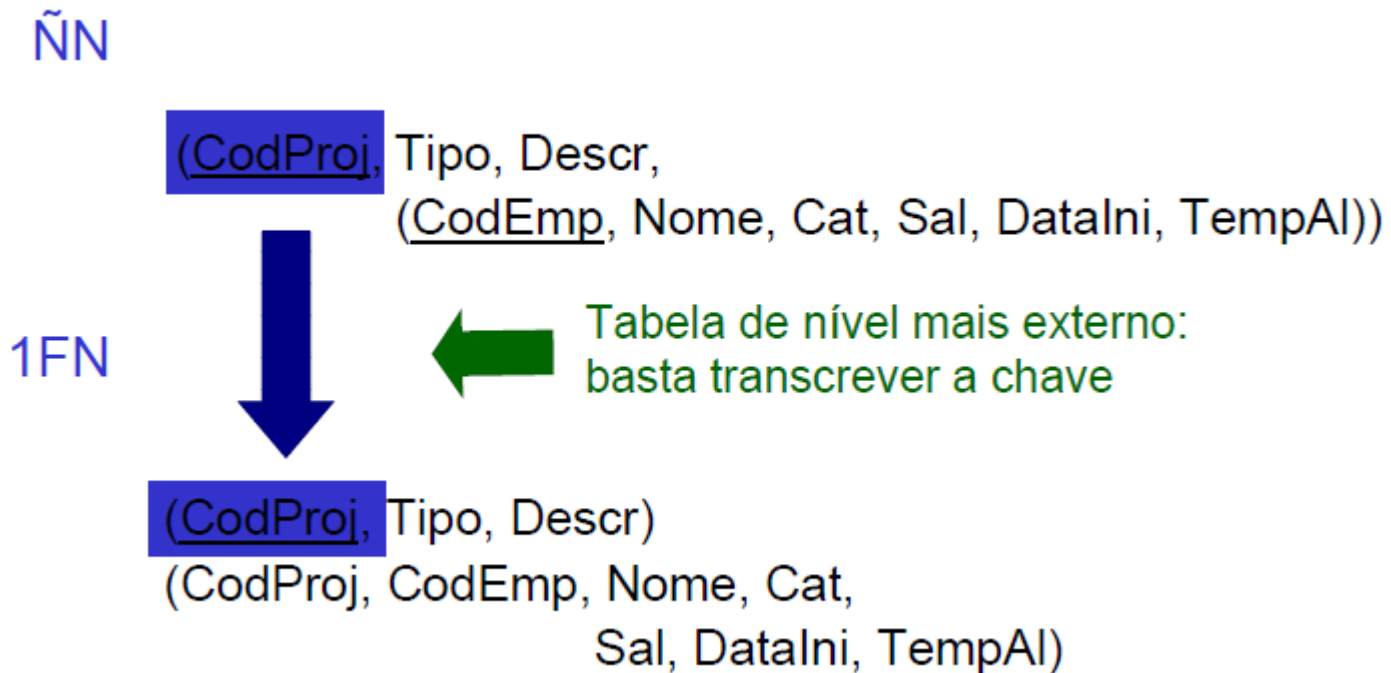
ÑN

(CodProj, Tipo, Descr,
(CodEmp, Nome, Cat, Sal, DataIni, TempAl))

1FN

(CodProj, Tipo, Descr)
(CodProj, CodEmp, Nome, Cat,
Sal, DataIni, TempAl)

- Definir as **chaves primárias** das tabelas na 1FN que correspondem a tabelas aninhadas.



1FN

(CodProj, Tipo, Descr)

(CodProj, CodEmp, Nome, Cat,
Sal, DataIni, TempAl)



Qual é a chave primária desta tabela?
Pergunta a fazer:

“um valor de **CodEmp** (chave da tabela origem) aparece uma vez só no documento ou várias?”

1FN

(CodProj, Tipo, Descr)

(CodProj, CodEmp, Nome, Cat,
Sal, DataIni, TempAl)



Como um valor de **CodEmp** aparece várias vezes, é necessário **CodProj** para distinguir as várias aparições

Modelagem de BD

Exemplo de passagem de ÑÑ para 1FN

Proj:

CódProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

ProjEmp:

CódProj	CodEmp	Nome	Cat	Sal	DataIni	TempAl
LSC001	2146	João	A1	4	1/11/91	24
LSC001	3145	Sílvio	A2	4	2/10/91	24
LSC001	6126	José	B1	9	3/10/92	18
LSC001	1214	Carlos	A2	4	4/10/92	18
LSC001	8191	Mário	A1	4	1/11/92	12
PAG02	8191	Mário	A1	4	1/05/93	12
PAG02	4112	João	A2	4	4/01/91	24
PAG02	6126	José	B1	9	1/11/92	12

Modelagem de BD

Exercício 1 de passagem de ÑÑ para 1FN

ÑÑ

Arq-Candidatos (Cod-Curso, Nome-Curso, Numero-Vagas-Curso,
(Cod-Cand, Nome-Cand, Escore-Cand))

Modelagem de BD

Exercício 1 de passagem de ÑÑ para 1FN

Decomposicao das tabelas

ÑÑ

Arq-Candidatos (Cod-Curso, Nome-Curso, Numero-Vagas-Curso,
(Cod-Cand, Nome-Cand, Escore-Cand))

1FN

Cursos (Cod-Curso, Nome-Curso, Numero-Vagas-Curso)
Candidatos (Cod-Curso, Cod-Cand, Nome-Cand, Escore-Cand)

Modelagem de BD

Exercício 1 de passagem de ÑÑ para 1FN

ÑÑ

Arq-Candidatos (Cod-Curso, Nome-Curso, Numero-Vagas-Curso,
(Cod-Cand, Nome-Cand, Escore-Cand))

Tabela de nível mais externo:
basta transcrever a chave

1FN

Cursos (Cod-Curso, Nome-Curso, Numero-Vagas-Curso)
Candidatos (Cod-Curso, Cod-Cand, Nome-Cand, Escore-Cand)

Modelagem de BD

Exercício 1 de passagem de ÑÑ para 1FN

ÑÑ

Arq-Candidatos (Cod-Curso, Nome-Curso, Numero-Vagas-Curso,
(Cod-Cand, Nome-Cand, Escore-Cand))



1FN

Cursos	(<u>Cod-Curso</u> , Nome-Curso, Numero-Vagas-Curso)
Candidatos	(Cod-Curso, <u>Cod-Cand</u> , Nome-Cand, Escore-Cand)

Qual é a chave primária desta tabela?

Pergunta a fazer:

“um valor de Cod-Cand (chave da tabela origem) aparece uma só vez no documento ou várias?”

Modelagem de BD

Exercício 2 de passagem de ÑN para 1FN

ÑN

Arq-Alunos (Cod-AI, Nome-AI,
(Cod-Curso, Sem-ingresso)
(Cod-Disc,
(Sem-Disc-Cursada, Nota-Disc)))

Modelagem de BD

Exercício 2 de passagem de ÑN para 1FN

ÑN

Arq-Alunos (Cod-Al, Nome-Al,
(Cod-Curso, Sem-ingresso)
(Cod-Disc,
(Sem-Disc-Cursada, Nota-Disc)))

1FN

Alunos	(Cod-Al, Nome-Al)
AlunoCurso	(Cod-Al, Cod-Curso, Sem-ingresso)
AlunoDisc	(Cod-Al, Cod-Disc)
AlunoDiscSem	(Cod-Al, Cod-Disc, Sem-Disc-Cursada, Nota-Disc)

Modelagem de BD

Exercício 2 de passagem de ÑN para 1FN

ÑN

Arq-Alunos (Cod-Al, Nome-Al,
(Cod-Curso, Sem-ingresso)
(Cod-Disc,
(Sem-Disc-Cursada, Nota-Disc)))

1FN

Alunos	(<u>Cod-Al</u> , Nome-Al)
AlunoCurso	(<u>Cod-Al</u> , <u>Cod-Curso</u> , Sem-ingresso)
AlunoDisc	(<u>Cod-Al</u> , <u>Cod-Disc</u>)
AlunoDiscSem	(<u>Cod-Al</u> , <u>Cod-Disc</u> , <u>Sem-Disc-Cursada</u> , Nota-Disc)

- **Para entender 2FN e 3FN**
 - é necessário compreender o conceito de *dependência funcional*.
- **Em uma tabela relacional, diz-se que**
 - uma coluna C_2 *depende funcionalmente* de uma coluna C_1 (ou que a coluna C_1 *determina* a coluna C_2) quando,
em todas linhas da tabela, para cada valor de C_1 que aparece na tabela, aparece o mesmo valor de C_2 .

...	Código	...	Salário	...
	E1		10	
	E3		10	
	E1		10	
	E2		5	
	E3		10	
	E2		5	
	E1		10	

Código → Salário

Modelagem de BD

Dependencia funcional

A	B	C	D
B	5	2	20
C	4	2	15
B	6	7	20
B	5	2	20
C	2	2	15
C	4	2	15
A	10	5	18
A	12	3	18
A	10	5	18
B	5	2	20
C	4	2	15
A	10	5	18
C	4	2	15

Dependência funcional
inexistente na tabela

$A \not\rightarrow B$

Modelagem de BD

Dependencia funcional

A	B	C	D
B	5	2	20
C	4	2	15
B	6	7	20
B	5	2	20
C	2	2	15
C	4	2	15
A	10	5	18
A	12	3	18
A	10	5	18
B	5	2	20
C	4	2	15
A	10	5	18
C	4	2	15

Dependência funcional
existente na tabela

$A \rightarrow D$

Modelagem de BD

Dependencia funcional

A	B	C	D
B	5	2	20
C	4	2	15
B	6	7	20
B	5	2	20
C	2	2	15
C	4	2	15
A	10	5	18
A	12	3	18
A	10	5	18
B	5	2	20
C	4	2	15
A	10	5	18
C	4	2	15

Uma coluna pode
depender funcionalmente
de uma combinação de
mais de uma coluna

$$(A,B) \rightarrow C$$

- **Objetiva eliminar um certo tipo de redundância de dados**
- **Exemplo**

(CodProj, CodEmp, Nome, Cat,, Sal, DataIni, TempAl)

- **Dados referentes a empregados (Nome, Cat e Sal)**
 - Redundantes, para os empregados que trabalham em mais de um projeto

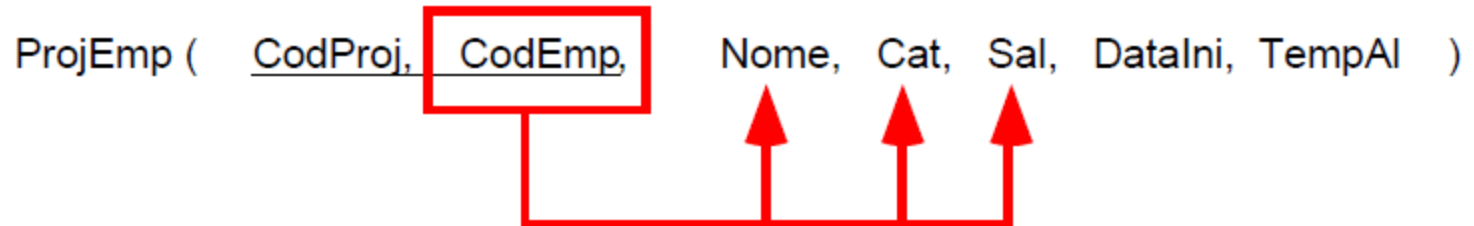
ProjEmp:

CódProj	CodEmp	Nome	Cat	Sal	DataIni	TempAl
LSC001	2146	João	A1	4	1/11/91	24
LSC001	3145	Sílvio	A2	4	2/10/91	24
LSC001	6126	José	B1	9	3/10/92	18
LSC001	1214	Carlos	A2	4	4/10/92	18
LSC001	8191	Mário	A1	4	1/11/92	12
PAG02	8191	Mário	A1	4	1/05/93	12
PAG02	4112	João	A2	4	4/01/91	24
PAG02	6126	José	B1	9	1/11/92	12

segunda forma normal (2FN)
=
uma tabela encontra-se na segunda forma normal,
quando, além de estar na 1FN, não contém
dependências parciais

dependência parcial
=
uma dependência (funcional) parcial ocorre
quando uma coluna depende apenas *de parte*
de uma chave primária composta

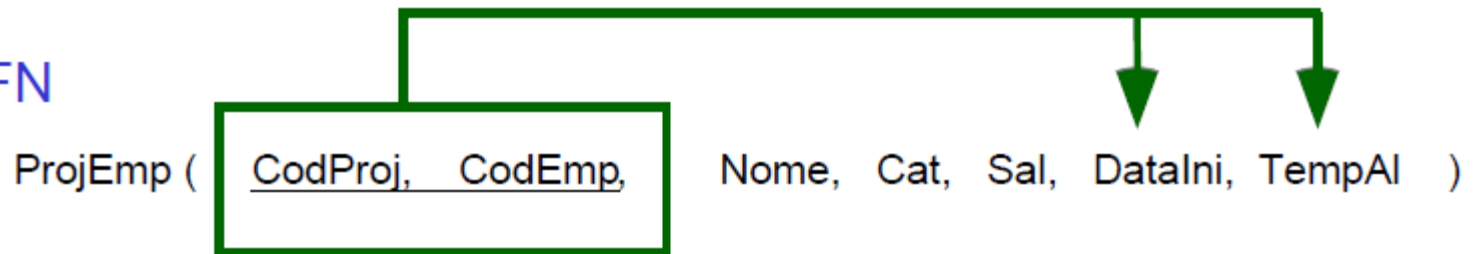
1FN



Modelagem de BD

2FN – Dependencia não parcial

1FN



- Tabela 1FN e que possui **apenas uma coluna como chave primária**
 - não contém dependências parciais
- É impossível uma coluna depender de uma parte da chave primária, quando a chave primária não é composta por partes
- Conclusão
 - Toda tabela 1FN que possui apenas uma coluna como chave primária já está na 2FN

1FN

(CodProj, Tipo, Descr)

(CodProj, CodEmp, Nome, Cat,
Sal, DataIni, TempAl)

2FN

(CodProj, Tipo, Descr)

- **Também**
 - Tabela que contenha apenas colunas chave primária
 - Impossível atributo não chave depender de parte da chave (tabela não tem colunas não chave)
 - Tabela sem colunas não chave já está na 2FN

1FN

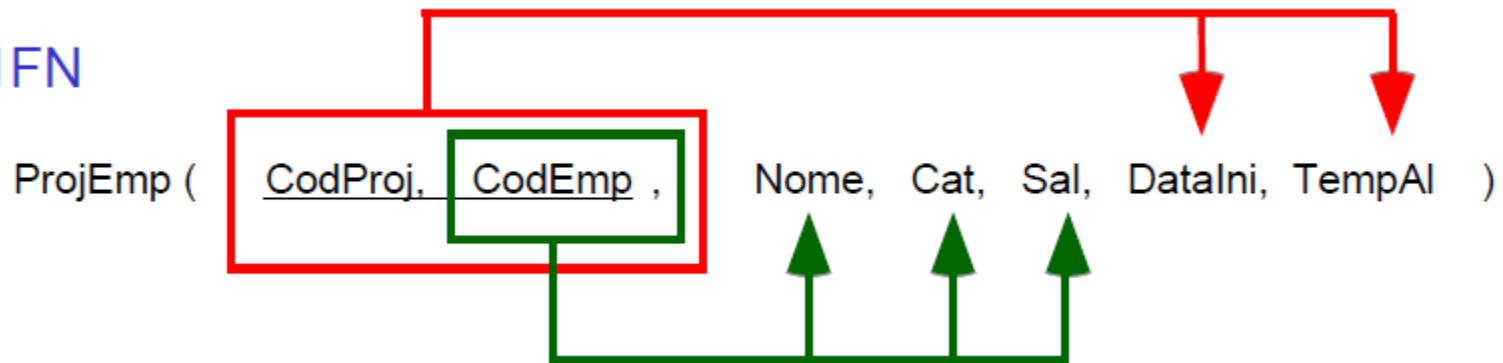


Tabela que possui chave primária com várias colunas e possui colunas não chave deve ser examinada

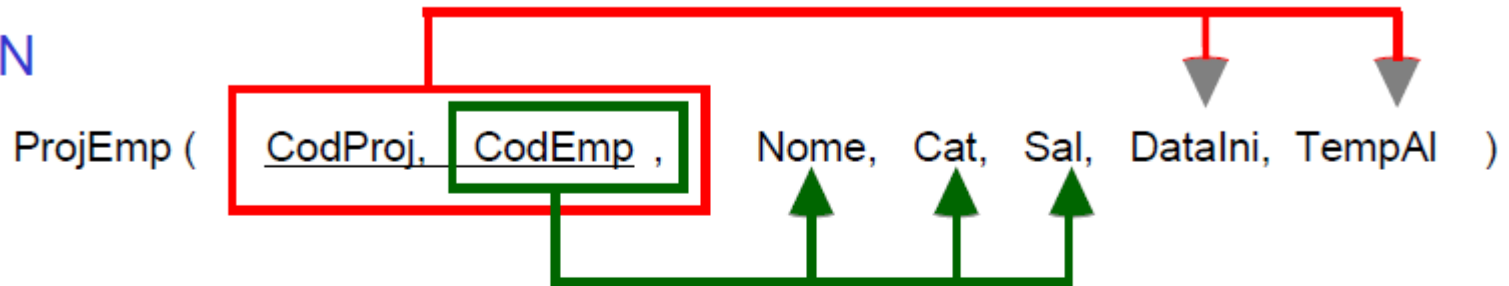
Pergunta a fazer, para cada coluna não chave:

“a coluna depende de toda a chave ou só de parte?”

ou

“para identificar um valor da coluna necessita de toda chave ou só de parte dela?”

1FN



2FN

ProjEmp (CodProj, CodEmp, DataIni, TempAl)

Emp (CodEmp, Nome, Cat, Sal)

Proj:

CódProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

ProjEmp:

CódProj	CodEmp	DataIni	TempAl
LSC001	2146	1/11/91	24
LSC001	3145	2/10/91	24
LSC001	6126	3/10/92	18
LSC001	1214	4/10/92	18
LSC001	8191	1/11/92	12
PAG02	8191	1/05/93	12
PAG02	4112	4/01/91	24
PAG02	6126	1/11/92	12

Emp:

CodEmp	Nome	Cat	Sal
2146	João	A1	4
3145	Sílvio	A2	4
6126	José	B1	9
1214	Carlos	A2	4
8191	Mário	A1	4
8191	Mário	A1	4
4112	João	A2	4
6126	José	B1	9

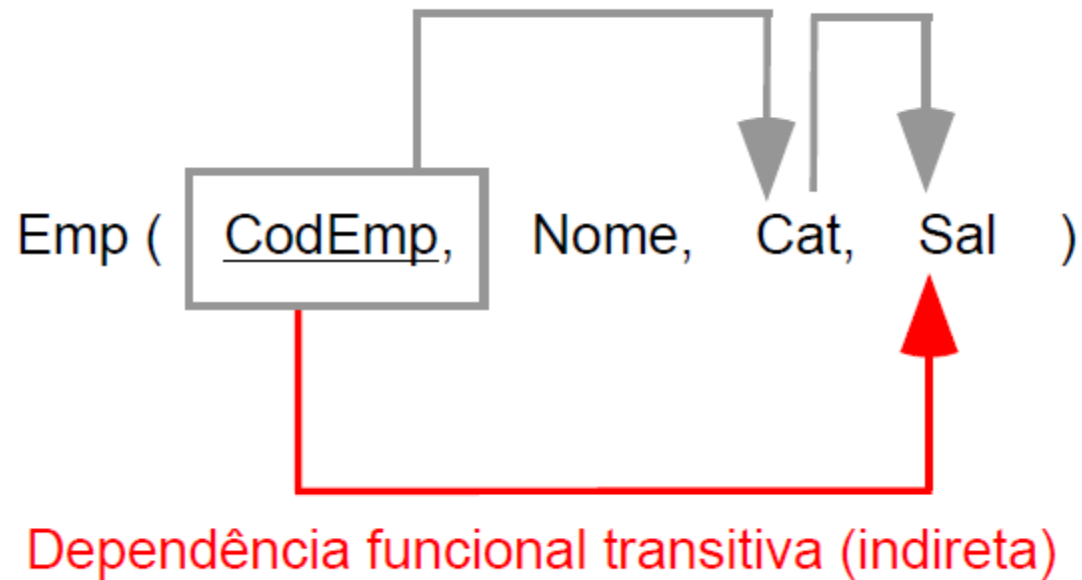
- Trata de um outro tipo de redundância
- Exemplo

Emp (CodEmp, Nome, Cat, Sal)

- Considerar
 - salário (coluna Sal) é determinado pela categoria funcional (coluna Cat)
- Salário que é pago a uma categoria funcional é armazenado tantas vezes quantos empregados possui a categoria funcional

Emp:

CodEmp	Nome	Cat	Sal
2146	João	A1	4
3145	Sílvio	A2	4
6126	José	B1	9
1214	Carlos	A2	4
8191	Mário	A1	4
8191	Mário	A1	4
4112	João	A2	4
6126	José	B1	9



terceira forma normal (3FN)

=

uma tabela encontra-se na terceira forma normal, quando, além de estar na 2FN, não contém dependências transitivas

3FN

Proj (CodProj, Tipo, Descr)

ProjEmp (CodProj, CodEmp, DataIni, TempAl)

Emp (CodEmp, Nome, Cat)

Cat (Cat, Sal)

2FN

ProjEmp (CodProj, CodEmp , DataIni, TempAl)

Emp (CodEmp , Nome, Cat, Sal)

3FN

Proj (CodProj, Tipo, Descr)

ProjEmp (CodProj, CodEmp, DataIni, TempAl)

Emp (CodEmp, Nome, Cat)

Cat (Cat, Sal)

1FN

Proj (CodProj, Tipo, Descr,
 (CodEmp, Nome, Cat, Sal, DataIni, TempAl))

2FN

Proj (CodProj, Tipo, Descr)
ProjEmp (CodProj, CodEmp, Nome, Cat, Sal, DataIni, TempAl)

3FN

Proj (CodProj, Tipo, Descr)
ProjEmp (CodProj, CodEmp, DataIni, TempAl)
Emp (CodEmp, Nome, Cat, Sal)
Cat (Cat, Sal)

Proj:

CódProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

ProjEmp:

CódProj	CodEmp	DataIni	TempAl
LSC001	2146	1/11/91	24
LSC001	3145	2/10/91	24
LSC001	6126	3/10/92	18
LSC001	1214	4/10/92	18
LSC001	8191	1/11/92	12
PAG02	8191	1/05/93	12
PAG02	4112	4/01/91	24
PAG02	6126	1/11/92	12

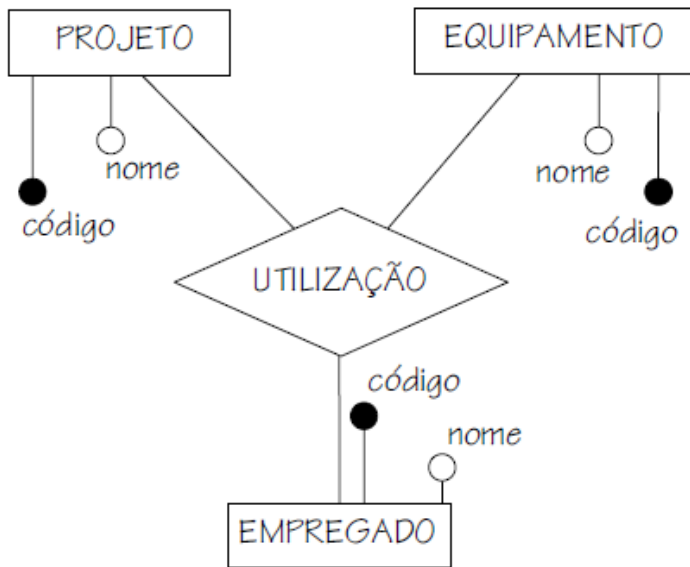
Emp:

NúmEmp	Nome	Cat
2146	João	A1
3145	Sílvio	A2
6126	José	B1
1214	Carlos	A2
8191	Mário	A1
8191	Mário	A1
4112	João	A2
6126	José	B1

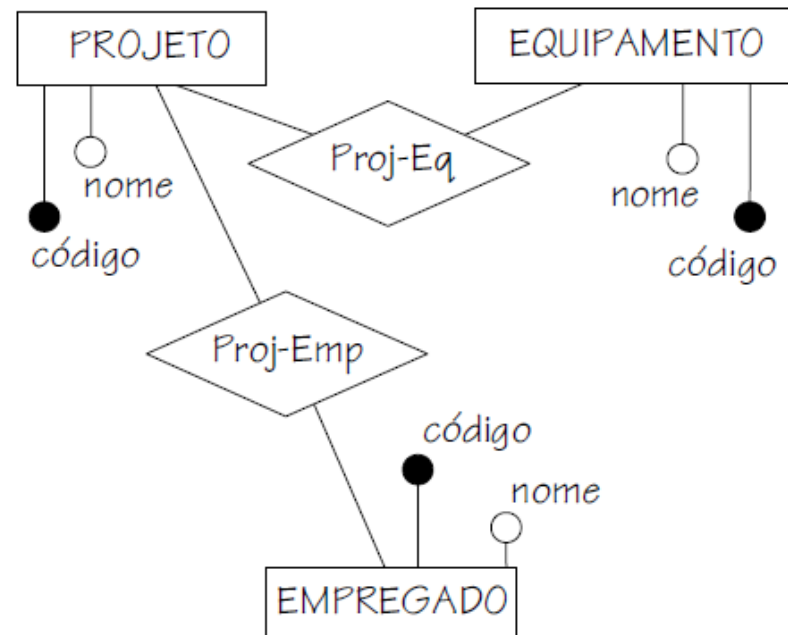
Cat:

Cat	Sal
A1	4
A2	4
B1	9

- **Maioria dos documentos e arquivos**
 - a decomposição até a 3FN é suficiente
- **Na literatura aparecem outras formas normais**
 - forma normal de Boyce/Codd
 - a 4FN
 - a 5FN



Requisitos alterados



CodProj	CodEmp	CodEquip
1	1	1
1	2	1
1	3	1
1	1	2
1	2	2
1	3	2
2	2	2
2	2	4
3	3	1
3	4	1
3	3	3
3	4	3
3	3	5
3	4	5
4	2	5

Verificar:

Quantas vezes cada empregado do projeto 1 é informado?

Quantas vezes cada equipamento usado no projeto 1 é informado?

CodProj	CodEmp	CodEquip
1	1	1
1	2	1
1	3	1
1	1	2
1	2	2
1	3	2
2	2	2

Dependencias
Multivaloradas

- **CodProj \twoheadrightarrow CodEmp**
- **CodProj \twoheadrightarrow CodEquip**

quarta forma normal (4FN)

=

uma tabela encontra-se na quarta forma normal, quando, além de estar na 3FN, não contém mais de uma dependência multi-valorada

CodProj	CodEmp	CodEquip
1	1	1
1	2	1
1	3	1
1	1	2
1	2	2
1	3	2
2	2	2

- **CodProj $\rightarrow\rightarrow$ CodEmp**
- **CodProj $\rightarrow\rightarrow$ CodEquip**

ProjEmp (CodProj,CodEmp)

ProjEquip (CodProj,CodEquip)

- Uma tabela não normalizada (NN) contém valores de atributos não atômicos, isto é, contém tabelas embutidas (grupos repetidos, arrays).
- Representação não normalizada do documento exemplo:
PROJ (CODPROJ, TIPOPROJ, DESCR,
 (NOEMP, NOME, CAT, SAL, DATAINÍCIO,TEMPOALOC))
- Deve-se observar a representação do embutimento de tabelas através de parênteses e a indicação das chaves primárias em cada nível de embutimento.

- **"Uma fábrica de móveis vai informatizar os pedidos de produtos especificados por funcionários para revendedores. Cada pedido contém um número, data, prazo de entrega, nome do cliente (revendedor), endereço, cidade, U.F., CGC e inscrição estadual do revendedor. Além disso, um pedido contém vários itens de pedido. Cada item de pedido registra o código do produto, descrição, setor, quantidade, valor unitário, desconto e subtotal."**

- "Uma fábrica de móveis vai informatizar os pedidos de produtos especificados por funcionários para revendedores. Cada pedido contém um número, data, prazo de entrega, nome do cliente (revendedor), endereço, cidade, U.F., CGC e inscrição estadual do revendedor. Além disso, um pedido contém vários itens de pedido. Cada item de pedido registra o código do produto, descrição, setor, quantidade, valor unitário, desconto e subtotal."

- Forma Não-Normalizada:

**PEDIDO (NUM_PEDIDO, DATA, PRAZO, NOME_REV, ENDEREÇO, CIDADE,
UF, CGC, INSCR_EST, NUM_F
(COD_PROD, DESCR, SETOR, QUANTIDADE, VALOR_UNIT,
DESCONTO, SUBTOTAL))**

- **“Uma empresa de ônibus possui um número, nome (empresa), endereço e nome do gerente. Cada empresa de ônibus possui vários ônibus e motoristas. Um ônibus pertence a somente uma empresa e um motorista pode trabalhar em mais de uma empresa de ônibus. Um ônibus possui um número, placa, ano de fabricação, tipo e capacidade (número de passageiros de acordo com o tipo do ônibus). Um motorista possui um número (especificado pela Secretaria Municipal de Transportes), nome e tempo de experiência. Uma linha de ônibus contém um número, itinerário, é executada por vários ônibus, pode ser percorrida várias vezes (cada uma delas com um horário de partida e de chegada) ao dia em sentidos opostos ou no mesmo sentido. Por exemplo, a linha 31 sai às 8 horas no sentido centro/bairro e também às 8 horas no sentido bairro/centro.”**

ÑN

Empregado (Número Empregado, Nome do Empregado, Número do Departamento, Nome do Departamento, Número do Gerente, Nome do Gerente, Número do Projeto, Nome do Projeto, Dia de Início do Projeto, Número de horas trabalhadas no projeto).

ÑÑ

Inscrição (Código do Aluno, Nome do Aluno, Telefone para contato, Ano de Admissão, Código da Disciplina, Nome da Disciplina, Nome do Curso, Data da Matricula).

ÑÑ

Paciente (num_paciente, nome_paciente, num_quarto, descrição_quarto, num_cômodos_quarto, {cód_médico, nome_médico, fone_médico}).

A vídeo Center of Europe Ltda., é uma cadeia de locadoras de DVDs. Ela precisa manter dados sobre os DVDs que têm para locação, os filmes dos DVDs, seus clientes e locações. Cada DVD para locação tem um número de série único. Os títulos de filme e números de cliente também são identificadores únicos. Suponha que cada filme tenha exatamente uma “estrela”. Observe a diferença no ano em que o filme foi originalmente filmado em oposição à data em que um DVD – o disco real – foi fabricado.

Alguns dos atributos e dependências funcionais neste ambiente são os seguintes:

Atributos: Número do DVD, Data de Fabricação, Título do Filme, Estrela, Ano de Filmagem, Duração (em minutos), Número do Cliente, Nome do Cliente, Endereço do Cliente, Data da Locação, Data da Devolução, Taxa Paga.

Dependências Funcionais:

Número do DVD → Título do Filme

Número do DVD → Estrela

Número do DVD → Data de Fabricação

Título do Filme → Estrela

Título do Filme → Duração

Título do Filme → Ano de Filmagem

Número do Cliente → Nome do Cliente

Número do Cliente → Endereço do Cliente

Número do DVD, Número do Cliente, Data de Locação → Data Devolução, Valor Pago