

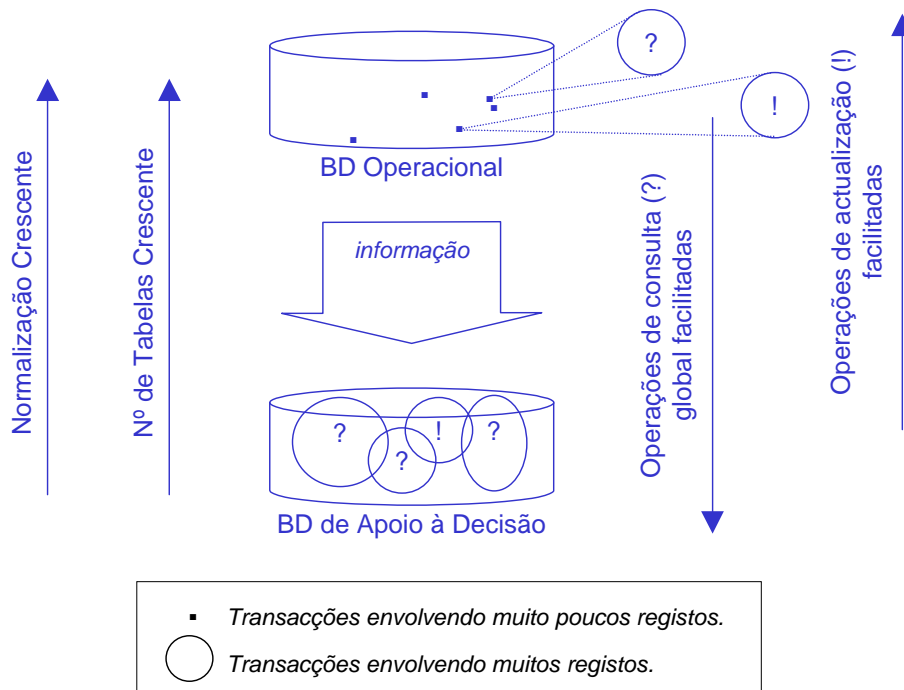
4. Projecto de Bases de Dados

Normalização Funcional de Dados para Concepção de Bases de Dados Relacionais

Tipos de operações numa base de dados

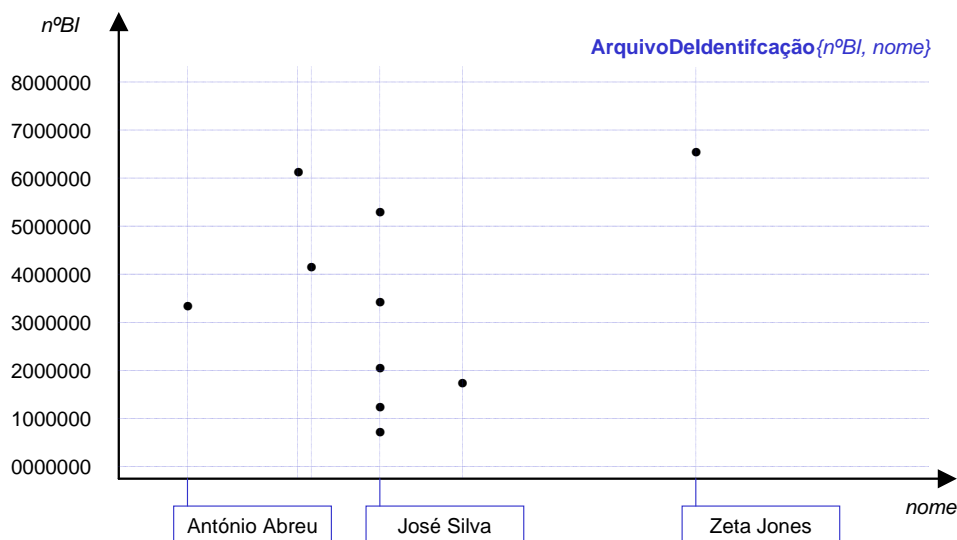
- num problema de gestão de informação existem normalmente dois tipos de operações ou processos:
 - os que modificam a informação (inserem, alteram ou eliminam)
 - os que pesquisam ou consultam informação
- a organização das tabelas da base de dados é fortemente condicionada
 - pelos tipos de operações a realizar
 - pela respectiva eficiência e tempo máximo de resposta permitido pelo utilizador
- qualquer organização de tabelas resulta de compromissos de engenharia envolvendo decisões sobre custos de equipamentos e sistemas, tempos de desenvolvimento e tempos de resposta.

BD operacional e BD de apoio à decisão



Motivação: os Bilhetes de Identidade

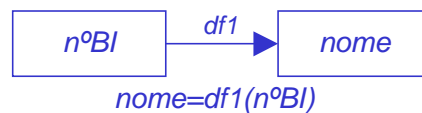
O conceito de **dependência funcional** tem uma importância crucial na discussão das questões de organização e projecto de bases de dados, em particular na teoria da normalização de informação.



Dependências Funcionais – Alguns exemplos

Exemplo: O Arquivo de Identificação - caso simples

"O Arquivo de Identificação gere os números de Bilhete de Identidade e os Nomes de todos os cidadãos".



Dependências Funcionais

Dependência Funcional (DF)

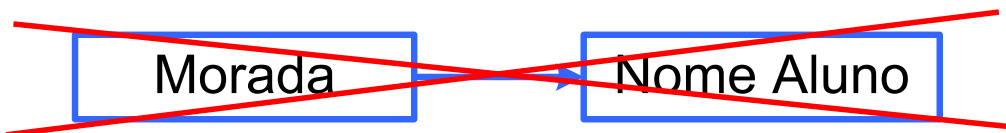
Seja R uma variável de relação e sejam X e Y dois subconjuntos do conjunto de atributos de

$$R: X=\{x_1, ..x_n\}, Y=\{y_1, ..y_m\}.$$

Diz-se que Y é funcionalmente dependente de X , ou que X determina funcionalmente Y , escrevendo-se $X \rightarrow Y$ (X seta Y), se e só se, para todo e qualquer valor válido de R , a cada valor distinto dos atributos de X estiver associado precisamente um só valor para cada atributo de Y . Nos casos de $n=1$ ou $m=1$ pode escrever-se apenas $x_1 \rightarrow Y$, $X \rightarrow y_1$ ou $x_1 \rightarrow y_1$.

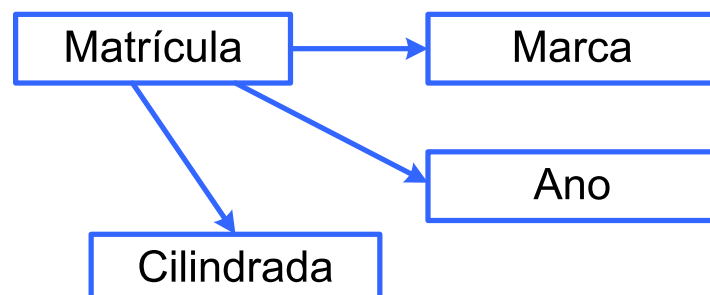
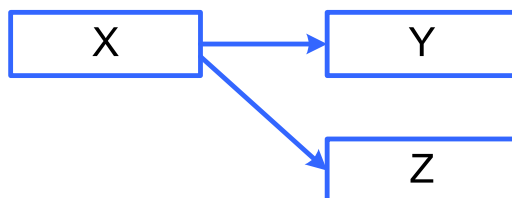
Dependências Funcionais

O atributo Y depende funcionalmente do atributo X



Dependências Funcionais

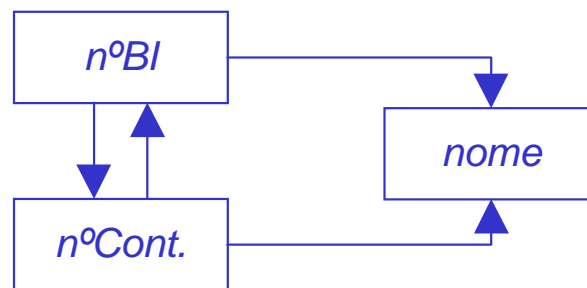
Os atributos Z e Y dependem funcionalmente do atributo X



Dependências Funcionais – Alguns exemplos

Exemplo: O Arquivo de Identificação - caso simples com códigos de contribuinte.

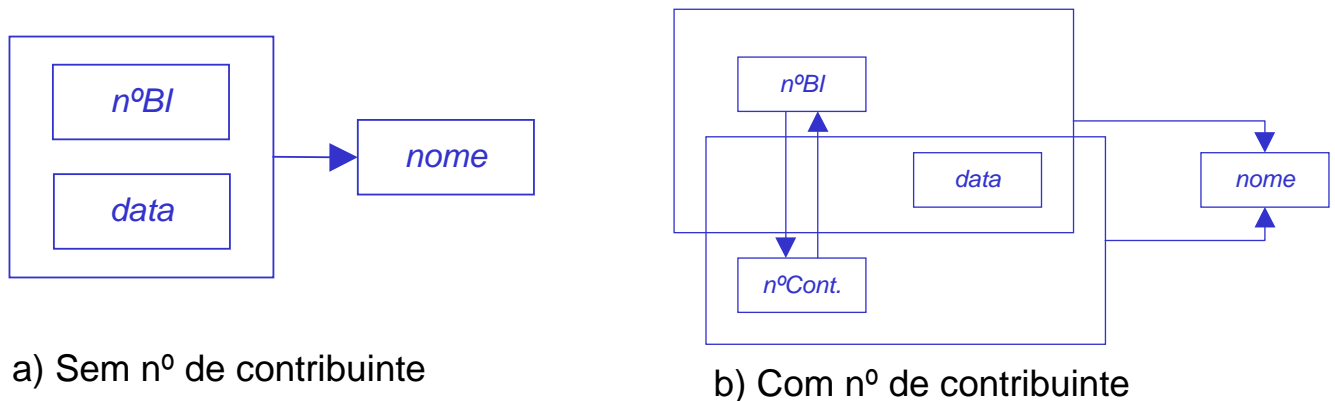
"O Arquivo de Identificação gere os números de Bilhete de Identidade e os Nomes de todos os cidadãos. Além disso o Arquivo mantém ainda os códigos de contribuinte de todos os cidadãos".



Dependências Funcionais – Alguns exemplos

Exemplo: O Arquivo de Identificação - caso histórico.

"O Arquivo de Identificação gere os números de Bilhete de Identidade e os Nomes de todos os cidadãos. O Arquivo de Identificação mantém um registo histórico com todas as alterações de nomes".



Dependências Funcionais – Alguns exemplos

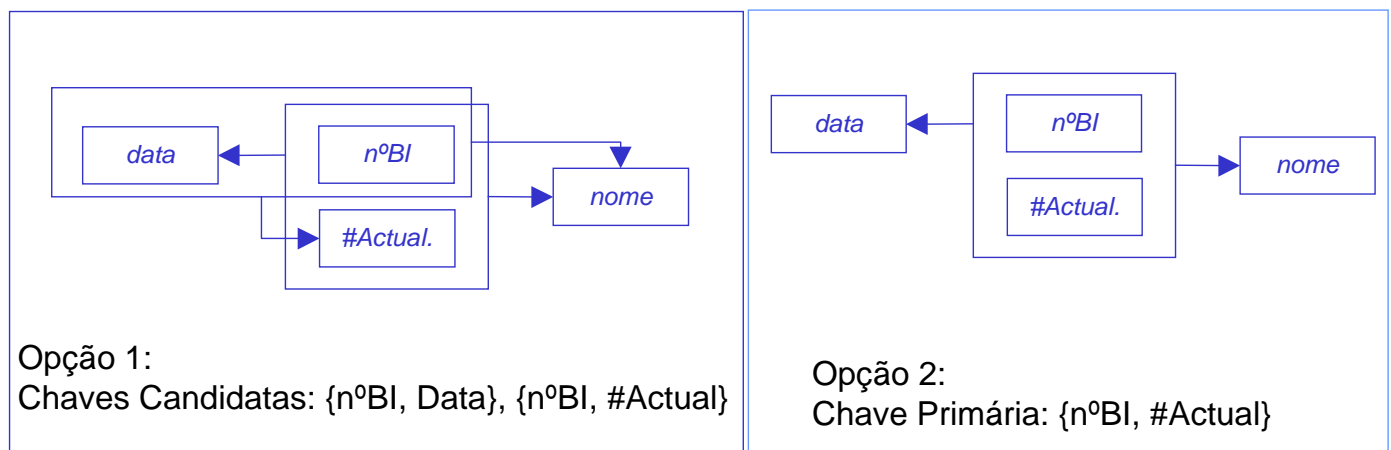
Exemplo: O Arquivo de Identificação - caso histórico com código de actualização e data.

"O Arquivo de Identificação gere os números de Bilhete de Identidade e os Nomes de todos os cidadãos. O Arquivo mantém um registo histórico com todas as alterações de nomes, e as respectivas datas de actualização. De cada vez que um BI sofre alteração do nome, é gerado um novo código de actualização sequencial para esse BI.

Por exemplo: em 1963-03-15 a Senhora Maria Elisa Silva obteve o seu primeiro Bilhete de Identidade com o número 2309927 (#Actual=1); em 1975-09-03, e após contrair matrimónio, Maria Elisa Silva alterou o seu nome para Maria Elisa Silva e Cunha (#Actual=2); em 1996-09-23 Maria Elisa Silva e Cunha voltou a casar e retomou o seu nome de solteira (#Actual=3)".

Exemplo (continuação)

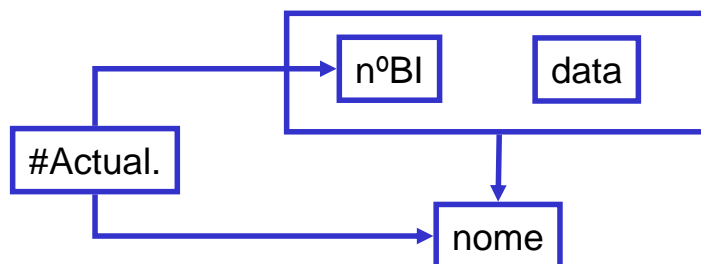
Por exemplo: em 1963-03-15 a Senhora Maria Elisa Silva obteve o seu primeiro Bilhete de Identidade com o número 2309927 (#Actual=1); em 1975-09-03, e após contrair matrimónio, Maria Elisa Silva alterou o seu nome para Maria Elisa Silva e Cunha (#Actual=2); em 1996-09-23 Maria Elisa Silva e Cunha voltou a casar e retomou o seu nome de solteira (#Actual=3)".



Dependências Funcionais – Alguns exemplos

Exemplo: O Arquivo de Identificação - caso histórico com código de actualização único.

"O Arquivo de Identificação gere os números de Bilhete de Identidade e os Nomes de todos os cidadãos. O Arquivo mantém um registo histórico com todas as alterações de nomes, e as respectivas datas de actualização. De cada vez que um BI sofre alteração do nome, é gerado um novo código de actualização sequencial. Os códigos de actualização são únicos em todo o arquivo, permitindo identificar não só o número do Bilhete de Identidade como a alteração sofrida no nome".



Chaves candidatas

Seja R uma relação definida num conjunto de atributos $X=\{x_1, \dots, x_n\}$, sujeita a um conjunto Cdf de dependências funcionais.

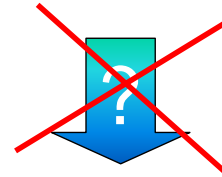
Uma **chave candidata** para R é qualquer conjunto de atributos $C_i=\{c_1, \dots, c_m\}$, $C_i \subseteq X$, que satisfaça simultaneamente as duas propriedades seguintes:

Unicidade (não repetição ou valores únicos): Dois n -uplos distintos de um dado R não podem ter os mesmos valores para os atributos do conjunto $C_i \subseteq X$. Para um dado valor de R os valores dos m -uplos de C_i nesse R são únicos.

Irreduzibilidade: Nenhum subconjunto próprio A do conjunto C_i , $A \subset C_i$, pode ter a propriedade anterior. Se existir um tal A então C_i não seria uma chave candidata para R . Designava-se então C_i como uma superchave.

CHAVES CANDIDATAS?

Unicidade e Irreduzibilidade



Num Aluno	Num BI	Nome Aluno	Morada	Data Nascimento
11002	7899342	João Gomes	Rua X	11-03-1978
11003	8766565	José Gomes	Rua Y	11-06-1984
11004	3254781	Joaquim Gomes	Rua X	23-08-1980
11005	6577904	João Santos	Rua X	25-04-1974
11006	2333234	Duarte Costa	Rua Z	25-04-1974
11007	9980987	Joaquina Gomes	Rua W	13-03-1969
11008	9976654	Alberto Santos	Rua XPTO	07-08-1967

Chave primária

Seja R uma relação para a qual está identificado o conjunto

$CC = \{C_1, C_2, \dots, C_k, \dots\}$ de chaves candidatas.

- Uma destas chaves candidatas é designada a **chave primária**, ou C_p , da relação.
- As restantes chaves candidatas são designadas **chaves alternativas**.

(A chave primária deverá ser escolhida como a chave candidata mais simples, mais estável ou mais apropriada em cada caso particular.)

Integridade de entidades

Nenhum atributo que participe numa chave primária de uma relação pode aceitar valores nulos.

Chave primária

Chave alternativa

Num Aluno	Num BI	Nome Aluno	Morada	Data Nascimento
11002	7899342	João Gomes	Rua X	11-03-1978
11003	8766565	José Gomes	Rua Y	11-06-1984
11004	3254781	Joaquim Gomes	Rua X	23-08-1980
11005	6577904	João Santos	Rua X	25-04-1974
11006	2333234	Duarte Costa	Rua Z	25-04-1974
11007	9980987	Joaquina Gomes	Rua W	13-03-1969
11008	9976654	Alberto Santos	Rua XPTO	07-08-1967

Chave Alheia ou Estrangeira

Uma **chave alheia** C_a de uma relação R_2 é um atributo, ou conjunto de atributos, cujos valores sejam do mesmo tipo dos valores da chave primária C_p de uma outra relação R_1 distinta (ou da própria relação R_2).

Diz-se, quando tal acontece, que as relações se podem **cruzar** por essas chaves, e que C_a de R_2 **liga com** C_p de R_1 (ou da própria R_2).

Num Aluno	Num BI	Nome Aluno	Morada	Data Nascimento	Cod Curso
11002	7899342	João Gomes	Rua X	11-03-1978	MIEM
11003	8766565	José Gomes	Rua Y	11-06-1984	MIEM
11004	3254781	Joaquim Gomes	Rua X	23-08-1980	MIEIG
11005	6577904	João Santos	Rua X	25-04-1974	MIEIG
11006	2333234	Duarte Costa	Rua Z	25-04-1974	MIEIG
11007	9980987	Joaquim Gomes	Rua W	13-03-1969	MIEIG
11008	9976654	Alberto Santos	Rua XPTO	07-08-1967	MIEIC

Chave Primária

Chave Candidata

Chave Alheia ou Estrangeira

Num Aluno	Cod Disciplina	Classificação
11002	SI	15
11003	IO	18
11004	SI	17
11002	AM1	12
11003	MP	16
11008	IO	13

Cod Curso	Nome Curso
MIEM	Mestrado Integrado Engenharia Mecânica
MIEIG	Mestrado Integrado Engenharia Industrial e Gestão
MIEIC	Mestrado Integrado Engenharia Informática e Computação

Integridade Referencial

Integridade referencial

se numa tabela existe uma referência com valor **não nulo** a informação adicional que está definida numa outra relação ou tabela (referência mantida através da ligação - igualdade de valores - entre a chave alheia e a chave primária), então nessa outra tabela tem de existir em alguma linha pelo menos uma vez esse valor.

Num Aluno	Num BI	Nome Aluno	Morada	Data Nascimento	Cod Curso
11002	7899342	João Gomes	Rua X	11-03-1978	MIEM
11003	8766565	José Gomes	Rua Y	11-06-1984	MIEM
11004	3254781	Joaquim Gomes	Rua X	23-08-1980	MIEIG
11005	6577904	João Santos	Rua X	25-04-1974	MIEIG
11006	2333234	Duarte Costa	Rua Z	25-04-1974	MIEIG
11007	9980987	Joaquim Gomes	Rua W	13-03-1969	MIEIG
11008	9976654	Alberto Santos	Rua XPTO	07-08-1967	MIEIC

Num Aluno	Cod Disciplina	Classificação
11002	SI	15
11003	IO	18
11004	SI	17
11002	AM1	12
11003	MP	16
11008	IO	13

Integridade Referencial

Cod Curso	Nome Curso
MIEM	Mestrado Integrado Engenharia Mecânica
MIEIG	Mestrado Integrado Engenharia Industrial e Gestão
MIEIC	Mestrado Integrado Engenharia Informática e Computação

Exercício

Proponha as DF consistentes com cada uma das quatro relações apresentadas, respeitando os valores e as chaves candidatas apresentadas.

R1		
a	b	c
1	A	L1
1	A	L2
1	B	L1
1	B	L2
2	A	L1
2	A	L2
2	B	L1
2	B	L2

R2		
a	b	c
1	A	L1
1	A	L2
1	A	L3
1	A	L4
2	B	L1
2	B	L2
2	B	L3
2	B	L4

R3		
a	b	c
1	A	L1
1	A	L2
1	A	L3
1	A	L4
2	A	L1
2	A	L2
2	A	L3
2	A	L4

R4		
a	b	c
1	A	L1
2	B	L2
3	C	L3
4	D	L5
5	E	L4
6	F	L8
7	G	L7
8	H	L9

CC1={{a, b, c}} CC2={{a, c}, {b, c}} CC3={{a, c}} CC4={{a}, {b}, {c}}

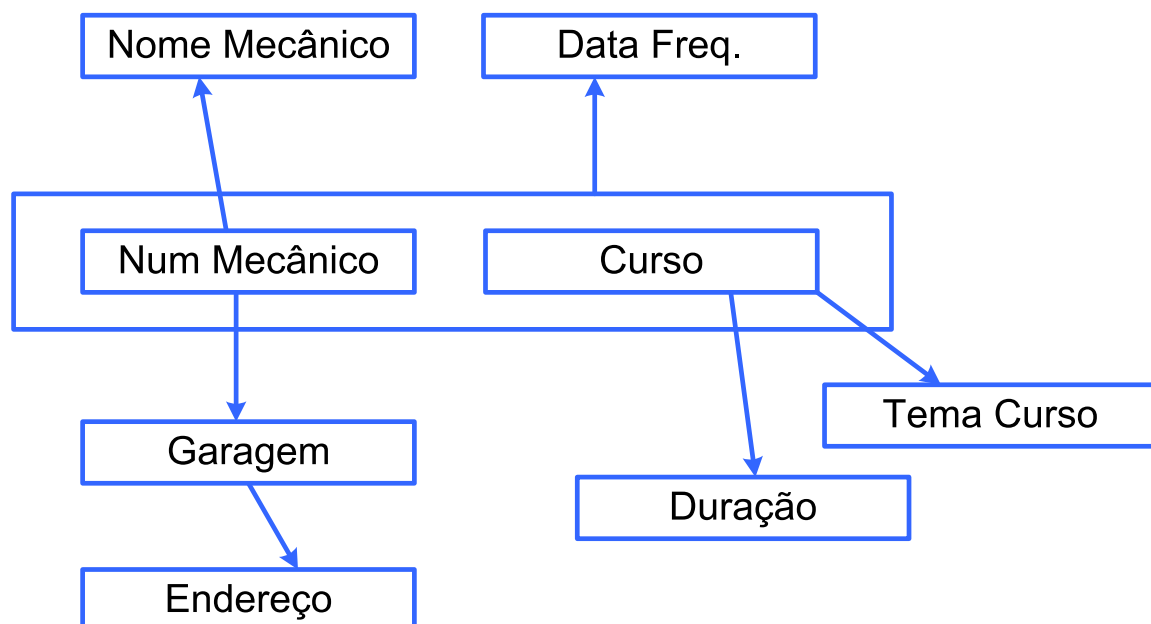
1ª, 2ª e 3ª Formas Normais

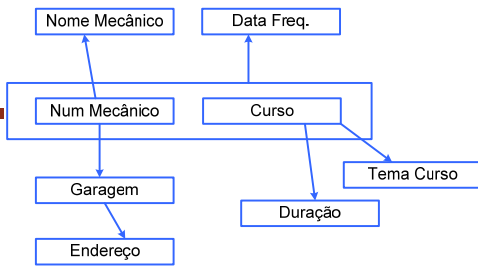
Primeira Forma Normal - 1FN: Uma relação R está em 1FN se e só se todos os domínios de definição dos atributos têm apenas **valores atômicos** ou escalares.

Segunda Forma Normal - 2FN: Uma relação R está em 2FN se e só se está em 1FN e todos os atributos não-chave (os atributos que não participam na chave primária da relação) são dependentes **irreduzivelmente** da chave primária (isto é, não dependem funcionalmente de um subconjunto próprio de atributos da chave primária).

Terceira Forma Normal - 3FN: Uma relação R está em 3FN se e só se está em 2FN e todos os atributos não-chave são dependentes **não transitivamente** da chave primária (isto é, não dependem funcionalmente de outros atributos que não os da chave primária).

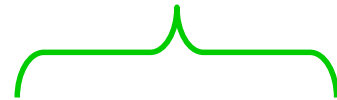
Diagrama de Dependências Funcionais



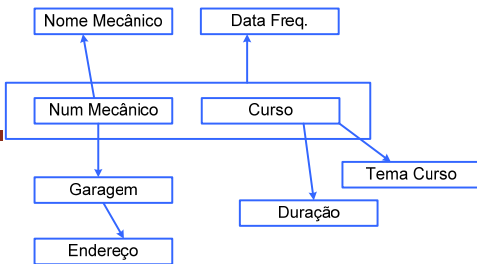


Primeira Forma Normal –
 todos os atributos têm
 apenas **valores atômicos**
 ou escalares.

Chave primária



Num Mecânico	Curso	Endereço	Garagem	Tema Curso	Duração	Data Freq.	Nome Mecânico
--------------	-------	----------	---------	------------	---------	------------	---------------



Segunda Forma Normal - está em 1FN e todos os atributos não-chave não dependem funcionalmente de um subconjunto próprio de atributos da chave primária.

Chave primária

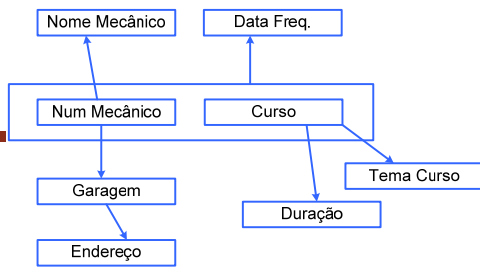
Curso	Tema Curso	Duração
-------	------------	---------

Chave alheia

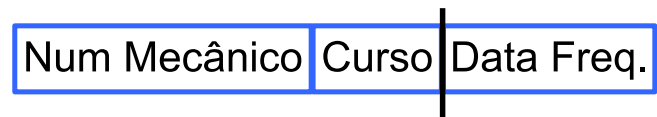
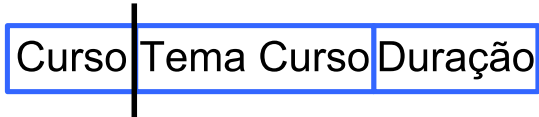
Num Mecânico	Nome Mecânico	Garagem	Endereço
--------------	---------------	---------	----------

D. F.

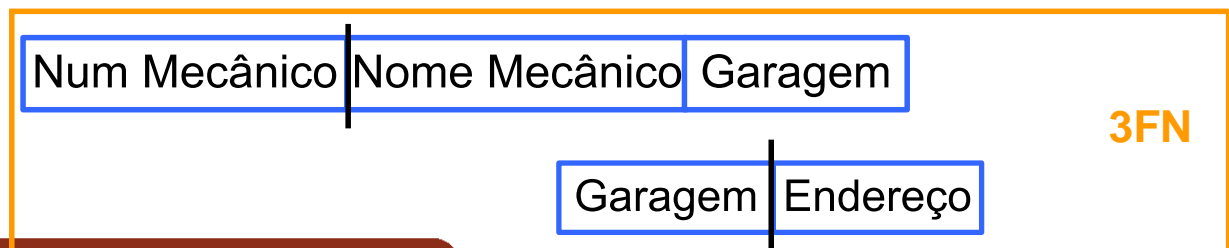
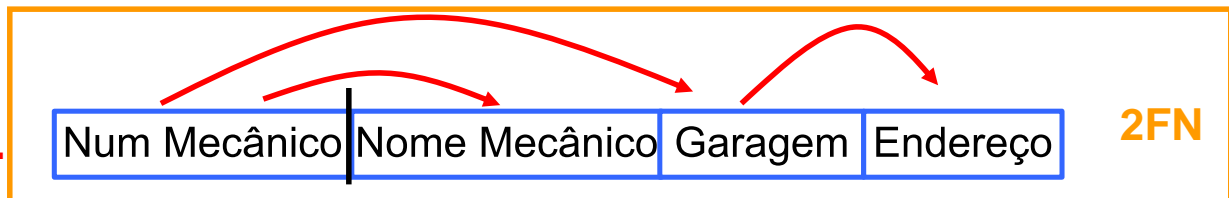
Num Mecânico	Curso	Data Freq.
--------------	-------	------------



Terceira Forma Normal – está em 2FN e todos os atributos não-chave são dependentes **não transitivamente** da chave primária



D. F.



Exercício (Dependências Funcionais)

Considere um projecto de desenvolvimento de uma base de dados para um portal de negócios para o sector da construção civil e obras públicas.

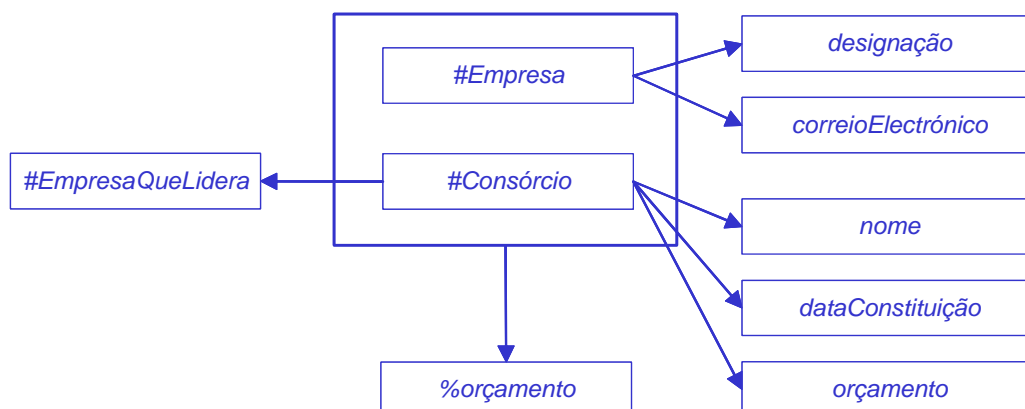
- As **empresas** deste sector **integram** muitos **consórcios**, por exemplo para executarem projectos ou para outras iniciativas conjuntas, tendo cada empresa em cada consórcio responsabilidade por uma certa **percentagem do orçamento** envolvido nas suas actividades. Cada consórcio é **liderado** por apenas uma única empresa.
- Para cada empresa interessa manter informação sobre a sua **designação** e sobre o endereço de **correio electrónico**.
- Sobre cada consórcio interessa manter informação sobre o **nome**, a **data da sua constituição** e o valor do **orçamento total** disponível para o consórcio (em euros).

Proponha uma base de dados para gerir a informação relevante. Derive a base de dados através de um processo de normalização funcional e depois compare com a solução obtida a partir de um processo de modelação conceptual de classes.

Dicionário de dados

Atributo	Tipo de dados	Descrição e observações	Exemplos
<i>#Empresa</i>		Identificador da empresa interno à BD	
<i>designação</i>		Designação da empresa	
<i>correioElectrónico</i>		Correio electrónico da empresa	
<i>#Consórcio</i>		Identificador do consórcio interno à BD	
<i>nome</i>		Nome do consórcio	
<i>dataConstituição</i>		Data de constituição do consórcio	
<i>orçamento</i>		Orçamento total do consórcio	
<i>%Orçamento</i>		Percentagem do orçamento total que cada empresa tem em cada consórcio	
<i>#EmpresaQueLidera</i>		Empresa que lidera um consórcio	

Diagrama de Dependências Funcionais



Chave primária: ?

Processo de Normalização

1ª FN

Empresas&Consórcios:

{#Empresa, #Consórcio, designação, correioElectrónico, nome, dataConstituição, orçamento, %Orçamento, #EmpresaQueLidera }

2FN e também 3FN

Empresa

{#Empresa, designação, correioElectrónico }

Consórcio

{#Consórcio, nome, dataConstituição, orçamento, #EmpresaQueLidera }

Empresa_Integra_Consórcio

{#Consórcio, #Empresa, %Orçamento }

Exercício 1 (Dependências Funcionais)

Considere a relação $R\{a, b, c, d, e, f, g\}$ com as seguintes dependências funcionais:

$b \rightarrow d$;

$\{a, c\} \rightarrow f$;

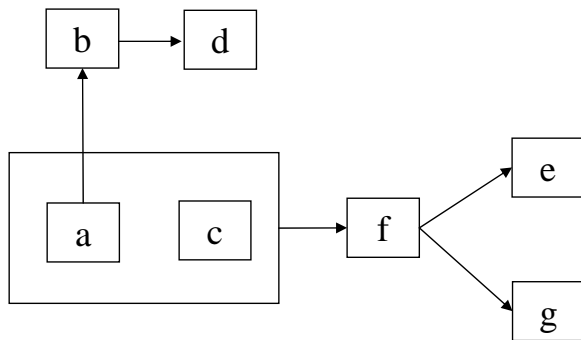
$f \rightarrow e$;

$a \rightarrow b$;

$f \rightarrow g$.

Represente o diagrama de dependências funcionais e indique quais as relações em terceira forma normal que preservam a informação de R (sublinhe as chaves primárias).

Resolução do Exercício 1 (Dependências Funcionais)



1FN R1{a, c, b, d, f, e, g}

2FN R2{a, b, d}
R3{a, c, f, e, g}

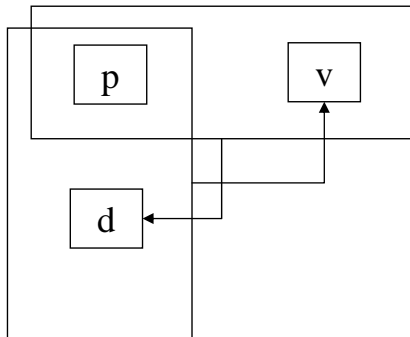
3FN R4{a, b}
R5{b, d}
R6{a, c, f}
R7{f, e, g}

Exercício 2 (Dependências Funcionais)

Uma relação **T** está definida nos atributos *p*, *v*, *d*. Uma n-upla desta relação significa que um piloto *p* reservou uma viatura *v* para a data *d*. Um piloto só pode reservar uma dada viatura para apenas um único dia, e num dado dia não pode reservar mais do que uma viatura.

Represente o diagrama de dependências funcionais implícito na informação. Indique em que forma normal está a relação **T**. Quais seriam as tabelas normalizadas que usava para manter a informação de reservas? Que problemas enfrenta para normalizar a relação **T**?

Resolução do Exercício 2 (Dependências Funcionais)



1FN $R1\{p, \underline{v}, d\}$ ou $R2\{p, \underline{d}, v\}$?

R1

p	v	d
A	X	1
A	Y	2
B	X	1
B	Y	2
A	Z	2

R2

p	d	v
A	1	X
A	2	Y
B	1	X
B	2	Y
A	2	Z

3FN $R1\{p, \underline{v}, d\}$
 $R2\{p, \underline{d}, v\}$