Banco de Dados – Ciência da Computação

Projeto de Banco de Dados

Prof. Mariane Moreira

mariane.souza@unifal-mg.edu.br

Universidade Federal de Alfenas <u>Instituto de Ciências Exatas</u> <u>Departamento de Ciência da Computação</u>

Projeto de banco de dados

- Com relação ao projeto de banco de dados, até agora vimos como:
 - Construir um modelo conceitual de bd (MER).
 - Mapear o modelo conceitual em um modelo representacional (MR).
- Até o momento não foi definido explicitamente nenhum critério de qualidade para o projeto.
 - A definição/agrupamento de atributos nas entidades foi feito através de regras naturais e intuitivas.
- Em nenhum momento nos preocupamos com a melhor configuração dos atributos nas relações.
 - Veremos isso na aula de hoje ☺.

Problemas com projeto de BD

cliente(codigo, numero, rua, cidade, descricao)

- O que é numero?
- O que é codigo?
- O que é descricao?

Campos com significados ambíguos

 cliente(idC, nomeC, rua, numero, cidade, descPerfil)

Problemas com projeto de BD

 emprestimoCliente(idC, nomeC, endC, numEm, quantia)

(1, Joao, Rua A, 120, Centro, E001, 20.000)

Repetição desnecessária de informações...

 (2, Joao, Rua A, 120, Centro, E002, 30000)

Problemas com projeto de BD

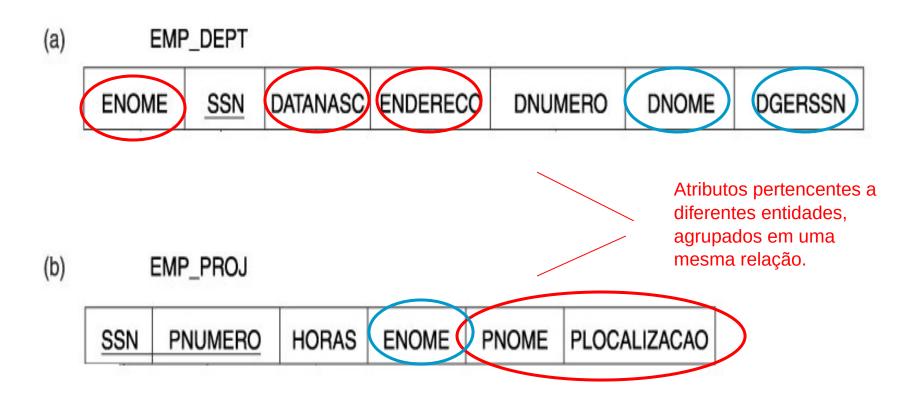
- Pessoa
 - ID
 - Nome
 - Endereco
 - DataNascimento
 - Matricula
 - Curso
 - Titulacao
 - Salario

Relações com baixo nível de coesão...

Diretrizes informais de qualidade de projeto

- Semântica dos atributos e relações
 - Um esquema deve ser de fácil entendimento.
 - Os atributos de entidades reais diferentes devem ser agrupados em relações diferentes.
- Redução de informações redundantes
 - **Evitar repetição** de informações nas tabelas.
- Impedimento de anomalias de atualização
 - Evitar problemas causados por certas operações de inserção, atualização, exclusão.
- Redução de valores nulos nos atributos
 - Evitar desperdício de espaço.
 - Evitar problemas no tratamento de funções agregadas (ex: sum, count...).
 - Evitar ambiguidade no significado de um valor nulo (desconhecido, qusente).

Exemplos de esquemas com problemas



redundância

_			
		DEDT	
_	MP	DEPT	
_			

ENOME	SSN	DATANASC	ENDERECO	DNUMERO	DNOME	DGERSSN
Smith,John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Pesquisa	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Pesquisa	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administracao	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administracao	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Pesquisa	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Pesquisa	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administracao	987654321
Borg,James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Sede Administrativa	888665555

redundância

EMP_PROJ

RO HORAS ENOME PNOME PLOCALIZAÇÃO

SSN	PNUMERO	HOHAS	ENOME	PNOME	PLOCALIZACAO
123456789	1	32.5	Smith,John B.	ProdutoX	Bellaire
123456789	2	7.5	Smith, John B.	ProdutoY	Sugarland
666884444	3	40.0	Narayan,Ramesh K.	ProdutoZ	Houston
453453453	1	20.0	English, Joyce A.	ProdutoX	Bellaire
453453453	2	20.0	English, Joyce A.	ProdutoY	Sugarland
333445555	2	10.0	Wong, Franklin T.	ProdutoY	Sugarland
333445555	3	10.0	Wong, Franklin T.	ProdutoZ	Houston
333445555	10	10.0	Wong, Franklin T.	Automação	Stafford
333445555	20	10.0	Wong, Franklin T.	Reorganização	Houston
999887777	30	30.0	Zelaya, Alicia J.	NovosBenefícios	Stafford
999887777	10	10.0	Zelaya,Alicia J.	Automação	Stafford
987987987	10	35.0	Jabbar, Ahmad V.	Automação	Stafford
987987987	30	5.0	Jabbar, Ahmad V.	Novos benefícios	s Stafford
987654321	30	20.0	Wallace, Jennifer S.	Novos benefícios	Stafford
987654321	20	15.0	Wallace, Jennifer S.	Reorganização	Houston
888665555	20	null	Borg,James E.	Reorganização	Houston

redundância

As informações de departamento serão repetidas redundantemente para cada empregado relacionado com o mesmo.

O mesmo acontece com as informações de empregado e projeto na tabela ao lado.

Inserção de um novo departamento?

Inserção de um novo empregado em um departamento existente?

Exclusão do empregado?

EMP DEPT

redundância

						,
ENOME	SSN	DATANASC	ENDERECO	DNUMERO	DNOME	DGERSSN
Smith,John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Pesquisa	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Pesquisa	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administracao	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administracao	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Pesquisa	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Pesquisa	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administracao	987654321
Borg,James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Sede Administrativa	888665555

Como inserir um novo departamento que a princípio não tem se quer um empregado trabalhando?

EMP DEPT

(2.5)				,		
ENOME	SSN	DATANASC	ENDERECO	DNUMERO	DNOME	DGERSSN
Smith,John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Pesquisa	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Pesquisa	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administracao	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administracao	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Pesquisa	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Pesquisa	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administracao	987654321
Borg,James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Sede Administrativa	

redundância

Como inserir um novo departamento que a princípio não tem se quer um empregado trabalhando?

Para **inserir** um novo empregado, as informações relativas ao seu departamento devem estar consistentes com as informações já definidas para este departamento no cadastro de outro empregado.

EMP DEPT

					201	
ENOME	SSN	DATANASC	ENDERECO	DNUMERO	DNOME	DGERSSN
Smith,John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Pesquisa	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Pesquisa	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administracao	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administracao	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Pesquisa	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Pesquisa	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administracao	987654321
Borg,James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Sede Administrativa	888665555

redundância

Como inserir um novo departamento que a princípio não tem se quer um empregado trabalhando?

Para inserir um novo empregado, as informações relativas ao seu departamento devem estar consistentes com as informações já definidas para este departamento no cadastro de outro empregado.

No momento de **excluir**um empregado, pode ser
que ele seja o último
relacionado ao departamento.
Com isso, as informações
deste departamento são
excluídas junto com o registro
do empregado.

redundância

EMP_DEPT

2336				33		
ENOME	SSN	DATANASC	ENDERECO	DNUMERO	DNOME	DGERSSN
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Pesquisa	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Pesquisa	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administracao	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administracao	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Pesquisa	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Pesquisa	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administracao	987654321
Borg,James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Sede Administrativa	888665555

Ex. de anomalias de atualização

Se eu atualizar o valor de algum atributo de departamento?

redundância EMP DEPT **ENOME** DNUMERO SSN DATANASC **ENDERECO** DNOME **DGERSSN** Smith, John B. 123456789 1965-01-09 731 Fondren, Houston, TX 5 Pesquisa 333445555 Wong, Franklin T. 333445555 638 Voss, Houston, TX Pesquisa 333445555 1955-12-08 Zelaya, Alicia J. Administração 999887777 1968-07-19 3321 Castle, Spring, TX 987654321 Wallace, Jennifer S. 1941-06-20 291 Berry, Bellaire, TX Administração 987654321 987654321 975 FireOak, Humble, TX Narayan, Ramesh K. 666884444 1962-09-15 Pesquisa 333445555 English, Joyce A. 453453453 1972-07-31 5631 Rice, Houston, TX Pesquisa 333445555 Jabbar, Ahmad V. 1969-03-29 980 Dallas, Houston, TX Administração 987987987 987654321 Borg, James E. 450 Stone, Houston, TX Sede Administrativa 888665555 888665555 1937-11-10

Ex. de anomalias de atualização

Se **atualizarmos** o valor de algum atributo de um dado departamento em particular, teremos que atualizar essa mesma informação em todos os empregados que trabalham naquele departamento.

redundância EMP DEPT **ENOME** DNUMERO SSN DATANASC **ENDERECO** DNOME **DGERSSN** Smith, John B. 123456789 1965-01-09 731 Fondren, Houston, TX Pesquisa 333445555 5 Wong, Franklin T. 638 Voss, Houston, TX Pesquisa 333445555 333445555 1955-12-08 3321 Castle, Spring, TX Zelaya, Alicia J. Administração 999887777 1968-07-19 987654321 Wallace, Jennifer S. 1941-06-20 291 Berry, Bellaire, TX Administração 987654321 987654321 975 FireOak, Humble, TX Narayan, Ramesh K. 666884444 1962-09-15 Pesquisa 333445555 English, Joyce A. 453453453 1972-07-31 5631 Rice, Houston, TX Pesquisa 333445555 Jabbar, Ahmad V. 1969-03-29 980 Dallas, Houston, TX Administração 987987987 987654321 Borg, James E. 450 Stone, Houston, TX Sede Administrativa 888665555 888665555 1937-11-10

Ex. de esquema que segue as diretrizes de qualidade de projeto

EMPREGADO

DEPARTAMENTO

|--|

DEPTO_LOCALIZACOES

DNUMERO	DLOCALIZACAO

PROJETO

PJNOME	PNUMERO	PLOCALIZACAO	DNUM
--------	---------	--------------	------

TRABALHA_EM

PNO	HORAS
	PNO

Processo Formal para um bom projeto de BD

- Normalização
 - Processo formal de análise de determinados esquemas de relações para:
 - Minimizar redundâncias
 - Minimizar anomalias de atualização

- Consiste em projetar o esquema, de modo a atender formas normais.
 - Formas normais definem, em conjunto, a maneira mais correta que um esquema deve ser projetado para atender os objetivos já definidos.
 - Existem 5 formas normais (As 3 primeiras são as mais aplicadas

Primeira forma normal

- Domínio atômico
 - Elementos são considerados unidades indivisíveis.

- Um esquema relacional R está na primeira forma normal se os domínios de todos os atributos de R forem atômicos.
 - Atributos não podem ser compostos.
 - Atributos não podem ser multivalorados.

- Representa uma *restrição* entre dois subconjuntos de atributos α e β de uma relação r, denotada por $\alpha \rightarrow \beta$.
- A dependência funcional

$$\alpha \rightarrow \beta$$

se aplica na relação r se, e somente se, sempre que quaisquer duas tuplas de r concordarem nos atributos α , elas também concordem nos atributos β . Ou seja,:

$$t_i[\alpha] = t_j[\alpha] \Rightarrow t_i[\beta] = t_j[\beta]$$
 para quaisquer i,j

 \blacksquare Considere r(A,B,C,D) com a seguinte instância:

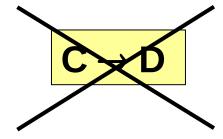
a1	b1 c1	d1	
a1	b2 c1	d2	
a2	b2 c2	d2	
a2	b3 c2	d3	
a3	b3 c2	d4	

A → C é satisfeita

A DF (propriedade da relação) pode ser observada em instâncias da relação

Uma DF é definida segundo a semântica dos atributos.





- Uma dependência funcional pode ser vista como uma restrição a ser assegurada no modelo criado.
- Isso ocorre quando a semântica de dois conjuntos de atributos indicam a necessidade de assegurar a dependência funcional.
 - {NumMat} → {Nome, Idade, Curso}
 - {Sigla, Sala, Hora} → {CódigoTurma, Professor}
 - {Sigla} → {NomeDisciplina, NCréditos}
- Podemos então testar se uma relação, do jeito que foi criada, satisfaz restrições definidas por DFs.

- Para construir um modelo relacional consistente, é necessário conhecer todas as dependências funcionais existentes (oriundas da natureza do domínio).
 - As relações devem ser construídas de modo a satisfazer essas dependências.

 Algumas DFs podem ser deduzidas a partir de outras já descobertas.

Para isso existem regras de inferência...

Regras de Inferência - DFs

- Reflexiva: se $B \subseteq A \Rightarrow A \rightarrow B$
- Aumentativa: se A \rightarrow B \Rightarrow AC \rightarrow BC
- Decomposição: se A → BC ⇒ A → B, A → C
- Aditiva: se A \rightarrow B, A \rightarrow C \Rightarrow A \rightarrow BC
- Transitiva: se A \rightarrow B, B \rightarrow C \Rightarrow A \rightarrow C
- Pseudo-Transitiva: se A \rightarrow B, BC \rightarrow D \Rightarrow AC \rightarrow D

Ex. de dependências funcionais

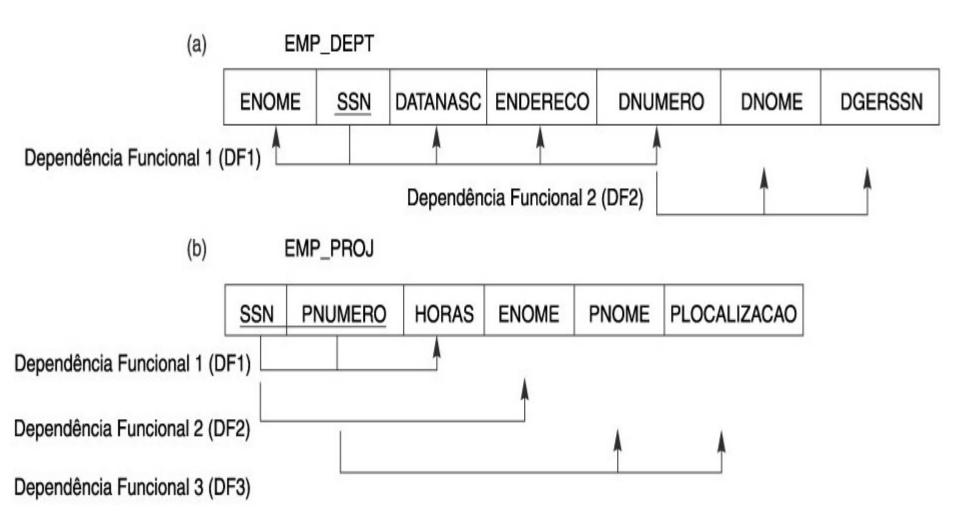
(a) EMP_DEPT

ENOME SSN DATANASC ENDERECO DNUMERO DNOME DGERSSN

(b) EMP_PROJ

SSN	PNUMERO	HORAS	ENOME	PNOME	PLOCALIZACAO
-----	---------	-------	-------	-------	--------------

Ex. de dependências funcionais



Dependências total, parcial, trivial e transitiva

- Dados os conjuntos de atributos X e Y, e um atributo A ∈ X :
 - X →Y é dependência funcional parcial se (X {A}) →Y
 - X →Y é dependência funcional total se (X {A}) →Y
 - X →Y é dependência funcional trivial se Y ⊆ X
 - X →Y é dependência funcional transitiva se existe X → Z e Z →Y, e Z não pode ser chave candidata, nem subconjunto da chave primária.

Segunda forma normal – 2FN

 Um esquema de relação R está na 2FN se estiver na 1FN e todo atributo A que não é chave primária em T tiver dependência funcional total da chave primária de lação abaixo...

> Ela está na segunda forma normal? Não!

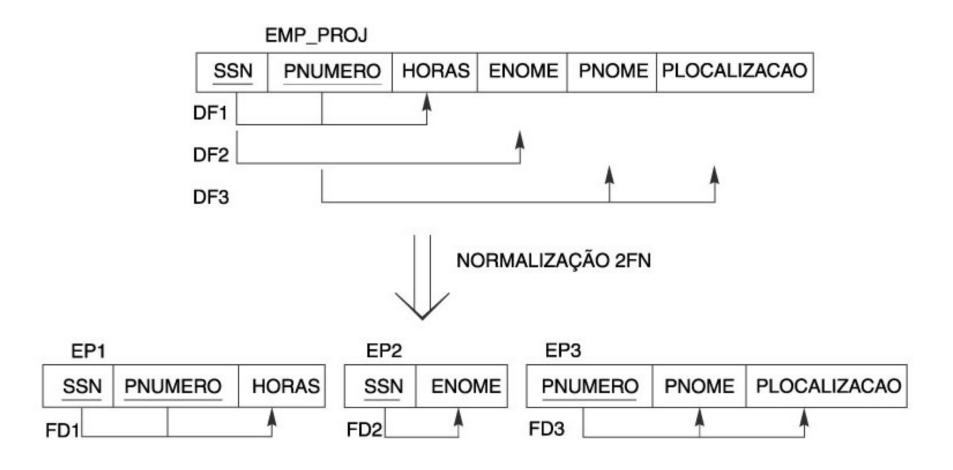
SSN PNUMERO HORAS ENOME PNOME PLOCALIZACAC
--

Solução: Decomposição

Como decompor?

Decompor a relação em relações menores, de modo que a restrição da 2FN seja respeitada.

Normalização para 2FN



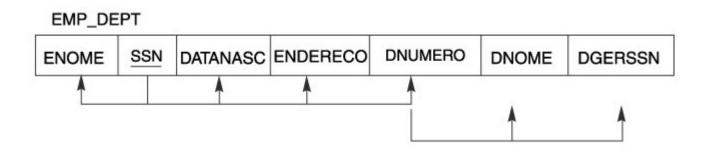
Terceira Forma Normal

 Um esquema de relação está na 3FN se estiver na 2FN e nenhum atributo que não for chave primária for transitivamente dependente da chave primária.

OU (em outras palavras)...

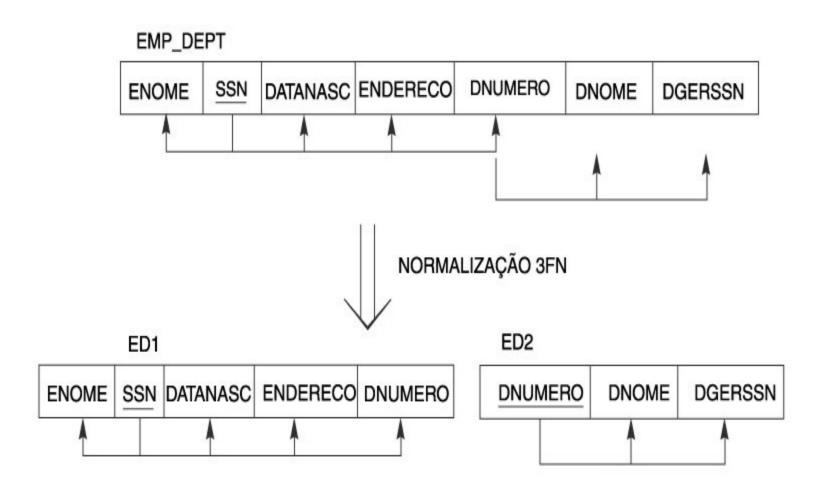
 Todos os atributos não primários devem possuir dependência total, não transitiva, da chave primária

Terceira Forma Normal



- Emp_Dept está na 2FN pois nenhum atributo não primário depende parcialmente da chave primária.
- Mas não está na 3FN pois DGERSSN e DNOME dependem transitivamente de SSN (DGERSSN e DNOME dependem de outro atributo que não pertence a qualquer chave (dnumero)).
 - SSN → DNUMERO → DNOME,DGERSSN

Normalização para 3FN



Generalização da segunda e terceira forma normal

2FN

 Um esquema de relação R está na 2FN se cada atributo não primário de R não for parcialmente dependente de nenhuma chave de R (chave primária ou candidata).

3FN

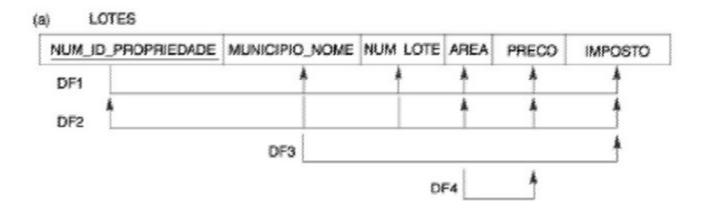
- Um esquema de relação R está na 3FN se cada atributo não primário de R apresentar as seguintes condições:
 - Ter dependência funcional total para todas as chaves (primária ou candidata).
 - Não ser dependente transitivamente de nenhuma chave de R (primária ou candidata).

Exemplo

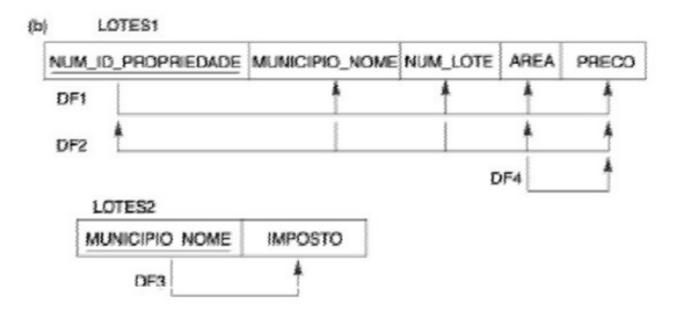
Generalização das formas normais

Domínio de Lotes

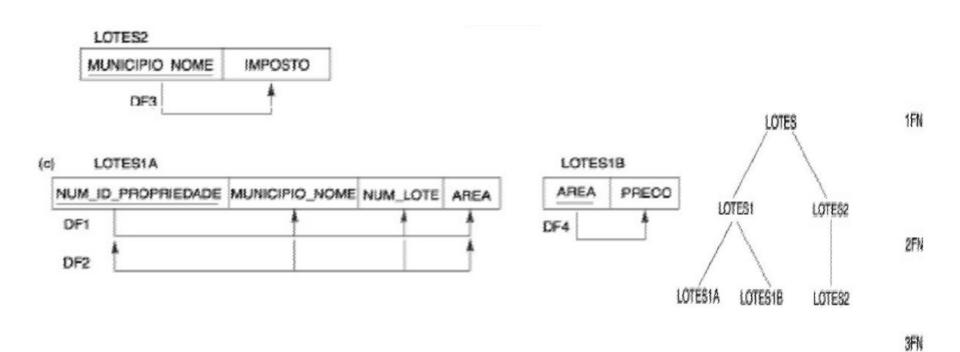
Ex. domínio de lotes

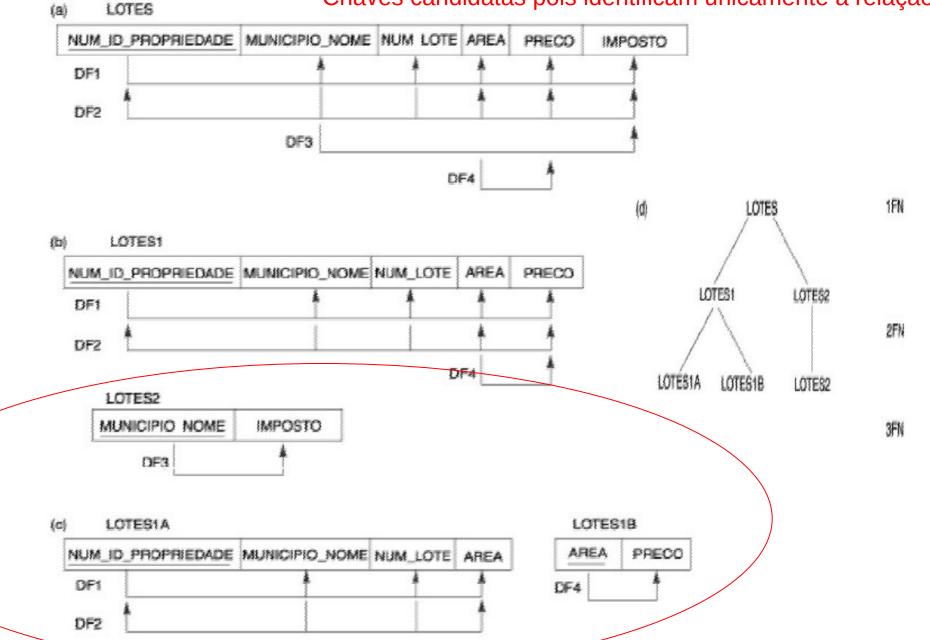


Ex. domínio de lotes



Ex. domínio de lotes





Algoritmo de decomposição 3FN

```
seja F o conjunto de dependências funcionais existente
i := 0;
                                                                    como chave primária
                                                                    ou candidata
para cada dependência funcional \alpha \to \beta em F faca
   se nenhum dos esquemas R_{i}, j = 1, 2, ..., i contém \alpha \beta
   entao
      i := i + 1;
       R_i := \alpha \beta;
   fimse
fimpara
se nenhum dos esquemas R_i, i = 1, 2, ..., i contém uma chave candidata para R_i
entao
       i := i + 1;
       R_i := qualquer chave candidata para R;
fimse
retorne (R_1, R_2,..., R_i)
```

Algoritmo de decomposição 3FN - exemplo

- R = (nomeAg,nomeC, nomeG,númeroS)
- F = {

 Df1: nomeG → númeroS

 Df2: nomeC, nomeAg → nomeG}
- Chave: {nomeC, nomeAg}
 - i = 0
 - Loop
 - para Df1
 - R1 = {nomeG, númeroS } (i=1)
 - i=1, para Df2
 - R2 = {nomeC, nomeAg, nomeG}
- Teste de chave candidata
- {nomeC, nomeAg} ⊆ R2, então não cria outra relação

Considerações finais

- 1FN, 2FN, 3FN são consideradas para cada relação:
 - O BD é considerado normalizado para uma determinada FN quando todas as suas relações estiverem nessa FN

- Normalização: decomposição de relações:
 - ② aumenta consistência;
 - ⊗ reduz desempenho ⇒ operações de junção.

EXERCÍCIO: Liste todas as dependências funcionais satisfeitas pela relação da figura:

Α	В	С
a ₁	b ₁	C ₁
a ₁	b ₁	C 2
a ₂	b ₁	C ₁
a ₂	b ₁	C 3

• Considere a seguinte relação:

VENDA_CARRO (Num_carro, Data_venda, Num_vendedor, Comissão_porc, Desconto_tempo)

 Suponha que um carro possa ser vendido por vários vendedores. Dependências adicionais são:

Data_venda → Desconto_tempo Num_vendedor → Comissão_porc

Essa relação está na 1FN, 2FN, 3FN? Por quê? Caso não esteja na 3FN, aplique o algoritmo de decomposição para a 3FN e mostre as relações resultantes.

 Nos exercícios seguintes, normalize as relações de forma que todas as relações resultantes estejam na forma normal mais restrita.

Considere a 1FN, a 2FN e a 3FN.

- Para cada FN:
- Se necessário, identifique quais as dependências funcionais que se aplicam sobre R;
- Identifique e justifique se R encontra-se ou não na forma normal em questão; e
- Caso R sendo analisada não se encontre na forma normal em questão, normalize-a, especificando as relações originadas.

- 1)vendedor (nro_vend, nome_vend, {clientetel (nro_cli, nome_cli)})
- As seguintes dependências funcionais devem ser garantidas na normalização:
 - nro_vend→nome_vend nro cli→nome cli
- Observação: considere que um vendedor pode atender diversos clientes, e um cliente pode ser atendido por diversos vendedores.

```
2)aluno ( nro_aluno, cod_depto, nome_depto, sigla_depto, cod_orient, nome_orient, fone_orient, cod_curso )
```

 As seguintes dependências funcionais devem ser garantidas na normalização:

Observações adicionais: -> próximo slide

- um aluno somente pode estar associado a um departamento;
- um aluno cursa um único curso;
- um aluno somente pode ser orientado por um único orientador.

Próxima aula...

Correção exercícios Normalização