

Gerenciamento de Dados e Informação

Ana Carolina Salgado
Fernando Fonseca
Robson Fidalgo



Cin.ufpe.br

Modelo Relacional

- Definido por E. F. Codd em 1970, teve sua grande aceitação comercial a partir de meados da década de 1980
- Razões da grande aceitação
 - Simplicidade dos conceitos básicos
 - Poder dos operadores de manipulação



2

Modelo Relacional

- Conceitos Básicos
 - Dada uma coleção de conjuntos D_1, D_2, \dots, D_n (não necessariamente disjuntos), R é uma Relação sobre estes n conjuntos se ela é um conjunto de n -uplas ordenadas $\langle d_1, d_2, \dots, d_n \rangle$ tal que d_1 pertence a D_1 , d_2 pertence a D_2 , ..., d_n pertence a D_n
 - D_1, D_2, \dots, D_n são Domínios
 - n é o grau de R



3

Modelo Relacional

- Sejam os domínios D_1 (D- Pessoa) e D_2 (D-Endereço)
 - | | |
|--------|----------|
| José | R. A, 30 |
| Maria | R. B, 45 |
| João | R. C, 17 |
| Thais | R. D, 67 |
| Branca | R. E, 55 |
- Seja a relação <Esposo, Esposa, Logradouro> em $D_1 \times D_1 \times D_2$

$\langle \text{José, Maria, R. A, 30} \rangle$
 $\langle \text{João, Thais, R. D, 67} \rangle$



4

Modelo Relacional

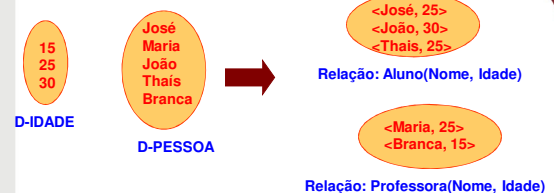
- O Domínio representa o conjunto de valores atômicos admissíveis de um componente de uma relação. Ele funciona como um conector semântico inter-relação a 2 níveis
 - Definição: todo valor de uma n -upla pertence a D_i
 - Manipulação: 2 valores só podem ser comparados se definidos sobre o mesmo domínio D
- Ex: D-IDADE: inteiro



5

Modelo Relacional

- Exemplo: Domínios e Relações abaixo



- É válido comparar a idade da professora com a dos alunos (mesmo domínio)



6

Modelo Relacional

- dupla: $\langle d_1, d_2 \rangle$, grau = 2
- tripla: $\langle d_1, d_2, d_3 \rangle$, grau = 3
- ...
- n-upla: $\langle d_1, d_2, \dots, d_n \rangle$, grau = n

➔ tupla

- Tupla

É uma n-upla $\langle a_1, a_2, \dots, a_n \rangle$ de uma relação $R(D_1, \dots, D_n)$ tal que a_i pertence a D_i ($1 \leq i \leq n$)

• $\langle \text{José}, 25 \rangle$ ➔ ∈ D-IDADE ➔ ∈ D-PESSOA

Centro de Informática
7

Modelo Relacional

- Atributo: Explicita o papel de um domínio em uma relação

• Ex: Fone-res: D-FONE } ➔ DOMÍNIO
 Fone-com: D-FONE }

- Os atributos de uma mesma relação devem ser diferentes (nomes diferentes)
- Um (ou vários) atributos identificam uma relação: **Chave Primária**

↙ ↘
Unicidade Minimalidade

Centro de Informática
8

Modelo Relacional

- Chave Candidata

Uma relação pode ter mais de um atributo como identificador único. Um deles é escolhido como chave primária e os outros são chaves candidatas

- Chave Estrangeira

Um atributo que corresponde a uma chave primária em outra relação

Centro de Informática
9

Modelo Relacional

Atributos

Chave Primária

PILOTO	Num_cad	Nome	CPF	Endereço
	0101	João	123456	Recife
	0035	José	234567	São Paulo
	...			
	0987	Pedro	567890	Recife

Tuplas

Piloto(Num_cad, Nome, CPF, Endereço)

Chave Candidata

Centro de Informática
10

Modelo Relacional

Chave Estrangeira

Voo	Num_voo	...	Num-pil	...
	
	330	...	0101	...
	

Voo (Num_voo, ..., Num-pil, ...) Num-pil referencia Piloto

Centro de Informática
11

Modelo Relacional

- Esquema Relacional

- Conjunto de relações semanticamente ligadas por seus domínios de definição
- O conceito de relação permite ao mesmo tempo representar
 - Uma entidade
 - Uma relação semântica (relacionamento)

Centro de Informática
12


Modelo Relacional

- Restrições de Integridade
 - Integridade de Domínio: diz respeito ao controle sintático e semântico de um dado e faz referência ao tipo de definição do domínio
 - Integridade de Entidade: diz respeito aos valores de chave primária que devem ser únicos e não nulos
 - Integridade Referencial: diz respeito aos valores de um atributo chave estrangeira e os valores do atributo chave primária correspondente

 13

Modelo Relacional

- Linguagem Algébrica/Álgebra Relacional
 - Compreende dois tipos de operadores
 - Operadores clássicos sobre conjuntos
 - União, interseção e diferença entre relações compatíveis
 - Operadores relacionais
 - Operadores unários de restrição: seleção e projeção
 - Operadores binários de extensão: junção e divisão


 14

Álgebra Relacional

- Seleção: seleciona todas as tuplas que satisfazem à condição de seleção em uma relação R
 - $\sigma_{\langle \text{condição de seleção} \rangle}(\text{Nome-da-relação})$


↓

 - condição simples: =, /≠ ou <, >, <=, >=
 - condição booleana: conexão de condições simples por operadores booleanos: AND, OR, NOT

 15

Álgebra Relacional

- Projeção: produz uma nova relação com alguns dos atributos de R
 - $\pi_{\langle \text{lista de atributos} \rangle}(\text{Nome-da-relação})$


 16

Álgebra Relacional

Projeção

Num_cad	Nome	CPF	Endereço
0101	João	123456	Recife
0035	José	234567	São Paulo
...
0987	Pedro	567890	Recife

Seleção


 17

Álgebra Relacional

- Junção: produz todas as combinações de tuplas de R1 e R2 que satisfazem a condição de junção
 - $R1 \bowtie_{\langle \text{condição de junção} \rangle} R2$

Num_cad	Nome	CPF	Endereço	Num_vod	...	Num_pil	...
0101	João	12345	Recife	330	...	0101	...
0205	José	71265	São Paulo
0957	Pedro	56789	Recife

Num_cad	Nome	CPF	Endereço	Num_vod	...	Num_pil	...
0101	João	12345	Recife	330	...	0101	...

 18

Álgebra Relacional

- **Divisão:** produz a relação $R(X)$ incluindo todas as tuplas de $R1(A)$ que aparecem em $R1$, combinadas com cada tupla de $R2(B)$, onde $B \subseteq A$ e $X = A-B$
 - $R1 \div R2$
 - Quais os pilotos que conduzem todos os aviões?

V	Piloto	Avião	Divisor	Avião	R	Piloto
	0020	101				
	0020	105				
	0010	101				
	0010	104				
	0010	105				
	0010	103				
	0015	103				
	0015	104				

Avião
101
104
105
103

Piloto
0010

Centro de Informática
19

Álgebra Relacional

- **União:** produz uma relação que inclui todas as tuplas de $R1$ ou $R2$. $R1$ e $R2$ devem ser **união compatíveis**
 - $R1 \cup R2$

↓

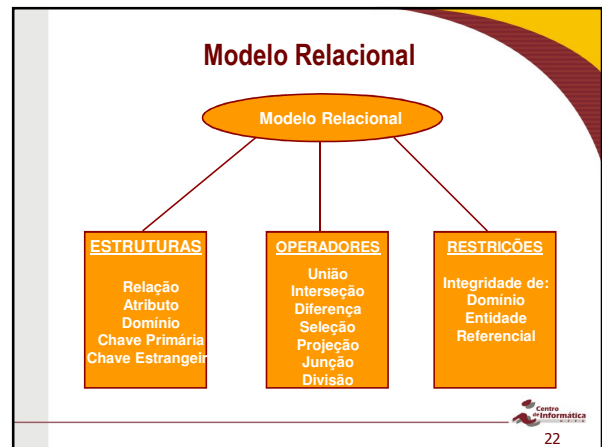
Duas relações $R(a_1, a_2, \dots, a_n)$ e $S(b_1, b_2, \dots, b_n)$ são **união compatíveis** se elas têm o mesmo grau n e se o $Dom(a_i) = Dom(b_i)$, $1 \leq i \leq n$

Centro de Informática
20

Álgebra Relacional

- **Interseção:** produz uma relação que inclui as tuplas comuns de $R1$ e $R2$. $R1$ e $R2$ devem ser **união compatíveis**
 - $R1 \cap R2$
- **Diferença:** produz uma relação que inclui todas as tuplas de $R1$ que não estão em $R2$. $R1$ e $R2$ devem ser **união compatíveis**
 - $R1 - R2$

Centro de Informática
21



Teoria das dependências Normalização

- No projeto de um banco de dados devemos
 - Identificar dados
 - Fazer com que estes dados representem eficientemente o mundo real
- Como proceder? Por intuição?
 - O processo de identificar e estruturar dados

Decomposição

→

Modelo relacional

←

Normalização

Centro de Informática
23


Normalização

- **Normalização**
 - Método permitindo identificar a existência de problemas potenciais (anomalias de atualização) no projeto de um BD relacional
 - Converte progressivamente uma tabela em tabelas de grau e cardinalidade menores até que pouca ou nenhuma redundância de dados exista

Centro de Informática
24

Normalização

- Se a normalização for bem sucedida
 - O espaço de armazenamento dos dados diminui
 - A tabela pode ser atualizada com maior eficiência
 - A descrição do BD será imediata



25


Teoria das Dependências

- Dependência Funcional

Sejam $R(A_1, A_2, \dots, A_n)$ e X, Y contidos em $\{A_1, A_2, \dots, A_n\}$ diz-se que existe uma Dependência Funcional (DF) de X para Y ($X \rightarrow Y$) se somente se, em R , a um valor de X corresponde um e um só valor de Y

Ex: Num-cad \rightarrow Nome
- DF Total: Se $X \rightarrow Y$ e $Y \rightarrow X$ ($X \leftrightarrow Y$)

Ex: Num-cad \leftrightarrow CPF




26

Teoria das Dependências

- DF Plena: quando um atributo é dependente de dois (ou mais) outros

Ex: Num-pil \rightarrow Trajeto
 Num-av \rightarrow Trajeto




27

Teoria das Dependências

- Chave Primária

Um atributo A (ou uma coleção de atributos) é a chave primária de uma relação R , se

 - Todos os atributos de R são funcionalmente dependentes de A
 - Nenhum subconjunto de atributos de A também tem a propriedade 1



28

Normalização

- Definição

Uma relação está na Primeira Forma Normal (1NF) se todos os atributos que a compõem são atômicos

Piloto	Num_cad	Nome	CPF	Telefone
	0010	José	123456	81 324569
	0015	João	234567	83 456785
	0020	Manuel	345678	45 768439
	0028	Josué	987654	21 347654



29

Normalização

- Como corrigir?
 - Separar os componentes em atributos distintos

Piloto	Num_cad	Nome	CPF	COD	TELEFONE
	0010	José	123456	81	324569
	0015	João	234567	83	456785
	0020	Manuel	345678	45	768439
	0028	Josué	987654	21	347654



30

Normalização

- Seja a relação

Piloto	NumCad	Nome	CPF	Salário	Diploma	Descrição
0010	José	123456	5.000,00	D1	Helicópteros	
0010	José	123456	5.000,00	D2	Aviões a jato	
0015	João	234567	3.000,00	D3	Bi-motor	
0020	Manuel	345678	8.000,00	D1	Helicópteros	
0020	Manuel	345678	8.000,00	D2	Aviões a jato	
0020	Manuel	345678	8.000,00	D4	Concorde	
0018	Josué	987654	4.000,00	D2	Aviões a jato	

Normalização

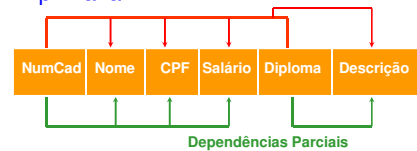
- Conseqüências da normalização na 1NF
 - Extensão da chave primária
 - Dependência funcional de parte da chave primária
 - Anomalias de atualização
 - Atualização de todas as tuplas com mesmo valor de atributo

Normalização

- Anomalias de atualização(Cont.)
 - Inconsistência: se a atualização não for feita em todos os níveis
 - Inclusões: de um item que não tem correspondente para os outros campos da chave primária
 - Remoções: de um item da chave provoca a remoção de informações adicionais

Normalização

- Definição
 - Uma relação está na Segunda Forma Normal (2NF) se ela está na 1NF e todo atributo não-chave é plenamente dependente da chave primária



Normalização

- Como corrigir?
 - Para cada subconjunto de atributos que compõem a chave primária, criar uma relação com este subconjunto como chave primária
 - Colocar cada um dos outros atributos com o subconjunto mínimo do qual ele depende

Normalização

- Relações criadas
 - (Num-cad, Nome, CPF, Salário)
 - (Diploma, Descrição)
 - (Num-cad, Diploma)
- Dar nome às novas relações, por exemplo:
 - Piloto, Diploma e Formação (respectivamente).

Normalização

Piloto	Num-cad	Nome	CPF	Salário	Formação	Num-cad	Diploma
	0010	José	12345	5.000,00		0010	D1
	0015	João	23456	3.000,00		0010	D2
	0020	Manuel	34567	8.000,00		0015	D3
	0018	José	98765	4.000,00		0020	D1
						0020	D2
						0020	D4
						0018	D2

Diploma	Diploma	Descrição
	D1	Helicópteros
	D2	Aviões a jato
	D3	Bi-motor
	D4	Concorde

OBS: Anomalias foram eliminadas - Não houve perda de informação

37

Teoria das Dependências

- Dependência Transitiva
 - Ocorre quando Y depende de X e Z depende de Y
 - Logo, Z também depende de X

$$X \rightarrow Y \rightarrow Z$$

No-avião	Tipo	Capacidade	Local

38

Normalização

- Definição
 - Uma relação está na Terceira Forma Normal (3NF) se ela está na 2NF e nenhum atributo não-chave é transitivamente dependente da chave primária

39

Normalização

- Como corrigir?
 - Para cada determinante que não é chave candidata, remover da relação os atributos que dependem dele para criar uma nova relação onde o determinante será chave primária

40

Normalização

Avião	No_av	Tipo	Capacidade	Local
	101	A320	320	Rio
	104	B727	250	S.Paulo
	105	DC10	350	Rio
	103	B727	250	Recife
	110	B727	250	Rio

Avião1	No_av	Tipo	Local
	101	A320	Rio
	104	B727	S.Paulo
	105	DC10	Rio
	103	B727	Recife
	110	B727	Rio

Tipo-av	Tipo	Capacidade
	A320	320
	B727	250
	DC10	350

41

Normalização

- Definição
 - Uma relação está na Forma Normal de Boyce/Codd (BCNF) se todo determinante é uma chave candidata

42

Normalização

- Relações com mais de uma chave candidata
- Considere a relação

ADP	Aluno	Disc	Prof.
	Maria	BD	Fernando
	Maria	ES	Paulo
	José	BD	Fernando
	José	ES	André

para cada disciplina, cada estudante tem um único professor
 - cada professor ensina uma única disciplina
 - cada disciplina é ensinada por vários professores

Centro de Informática
43

Normalização

A relação está na 3NF mas não está na BCNF

AP	Aluno	Professor
	Maria	Fernando
	Maria	Paulo
	José	Fernando
	José	André

PD	Professor	Disciplina
	Fernando	BD
	Paulo	ES
	André	ES

Centro de Informática
44

Normalização

Seja a relação

Vôo	Piloto	Avião	Trajeto
	0020	101	Rec-Rio
		105	Rio- Spa
			Spa-Rec
	0010	101	Rec-For
		104	
	0015	103	Rio-Spa

?

Vôo	Piloto	Avião	Trajeto
	0020	101	Rec-Rio
	0020	105	Rio- Spa
	0020	105	Spa-Rec
	0020	101	Spa-Rec
	0020	101	Rio- Spa
	0020	105	Rec-Rio
	0010	101	Rec-For
	0010	104	Rec-For
	0015	103	Rio-Spa

Centro de Informática
45

Normalização

- Dependência Multivalorada (DMV)

Dada uma relação R com atributos A, B, C, existe uma dependência multivalorada do atributo A no atributo B ($A \twoheadrightarrow B$) se um valor de A é associado a uma coleção específica de valores de B, independente de quaisquer valores de C

> Dada $R(A, B, C)$, a DMV $A \twoheadrightarrow B$ existe se $A \twoheadrightarrow C$ também existir. Daí a notação: $A \twoheadrightarrow B|C$

Centro de Informática
46

Normalização

- Definição

Uma relação está na Quarta Forma Normal (4NF) se ela está na 3NF (BCNF) e não existem dependências multivaloradas

- Como corrigir?

Separar a relação em relações, cada uma contendo o atributo (A) que multidetermina os outros (B, C), ou seja, $R_1(A, B)$ e $R_2(A, C)$

Centro de Informática
47

Normalização

Vôo	Piloto	Avião	Trajeto
	0020	101	Rec-Rio
	0020	105	Rio- Spa
	0020	105	Spa-Rec
	0020	101	Spa-Rec
	0020	101	Rio- Spa
	0020	105	Rec-Rio
	0010	101	Rec-For
	0010	104	Rec-For
	0015	103	Rio-Spa

Vôo1

Piloto	Avião
0020	101
0020	105
0010	101
0010	104
0015	103

Vôo2

Piloto	Trajeto
0020	Rec-Rio
0020	Rio-Spa
0020	Spa-Rec
0010	Rec-For
0015	Rio-Spa

Centro de Informática
48

Normalização

- OBS: Para evitar relações não 4NF
 - Quando existir mais de um atributo multivalorado, o processo de normalização para 1NF é o seguinte:
 - Criar uma relação para cada atributo multivalorado (AMV) (e os que ele determina).
 - Incluir a chave primária da relação original (CPO)
 - A chave primária será: CPO + AMV

Centro de Informática
49

Normalização

- Notas
 - Até agora o processo de decomposição substituía uma relação por duas de suas projeções
 - Pode acontecer que uma relação não possa ser decomposta sem perdas em duas relações mas possa ser decomposta em três ou mais. Este fenômeno é chamado n-decomposição ($n > 2$)

Centro de Informática
50

Normalização

Seja a relação Voo :

Voo	Piloto	Avião	Trajeto
0020	101	Rec-Rio	
0020	105	Rec-Rio	
0020	101	Rio-Spa	
0010	101	Rec-Rio	

É possível aplicar a operação de projeção de várias maneiras

Projeção em duas relações

Piloto	Avião
0020	101
0020	105
0010	101

Avião	Trajeto
101	Rec-Rio
105	Rec-Rio
101	Rio-Spa

Centro de Informática
51

Normalização

Projeção em três relações

Piloto	Avião
0020	101
0020	105
0010	101

Avião	Trajeto
101	Rec-Rio
105	Rec-Rio
101	Rio-Spa

Trajeto	Piloto
Rec-Rio	0020
Rio-Spa	0020
Rec-Rio	0010

Projeção ...

Centro de Informática
52

Normalização

Piloto	Avião
0020	101
0020	105
0010	101

Avião	Trajeto
101	Rec-Rio
105	Rec-Rio
101	Rio-Spa

Trajeto	Piloto
Rec-Rio	0020
Rio-Spa	0020
Rec-Rio	0010

Avião

Piloto	Avião	Trajeto
0020	101	Rec-Rio
0020	105	Rec-Rio
0020	101	Rio-Spa
0010	101	Rec-Rio
0010	101	Rio-Spa

(Trajeto, Piloto)

Relação Original

Centro de Informática
53

Normalização

- Definição
 - Uma relação R satisfaz a Dependência de Junção * (X, Y, ..., Z) se somente se R é igual à junção de suas projeções em X, Y, ..., Z onde X, Y, ..., Z são subconjuntos do conjunto de atributos de R.
 - No exemplo, a relação Voo satisfaz à DJ:
 - * ({Piloto, Avião}, {Avião, Trajeto}, {Trajeto, Piloto})

Centro de Informática
54

Normalização

Problemas na atualização

Voo	Piloto	Avião	Trajeto
	P1	A1	T2
	P1	A2	T1
	P2	A1	T1
	P1	A1	T1

- Se t3 for inserida => t4
- Se t4 for removida, uma das t também tem que ser

Exemplo:
 Se João pilota um B737 e o trajeto Rec-Rio é feito por um B737 e João pilota no trajeto Rec-Rio
 Então João pilota um B737 no trajeto Rec-Rio
 O que não é necessariamente verdadeiro

Centro de Informática
55

Normalização

Teorema de Fagin

$R(A, B, C)$ satisfaz à $DJ \rightarrow (AB, AC)$ se e somente se satisfaz o par de DMV $A \rightarrow B \mid C$

Este teorema pode ser visto como uma definição de DMV, logo

DJ

↓

DMV

↓

DF

: Generalização

Centro de Informática
56

Normalização

OBS: A DJ é a forma mais geral de dependências que tratam relações sendo decompostas via projeção e reconstruídas via junção

- Definição

Uma relação está na Quinta Forma Normal (5NF), também chamada de Forma Normal Projeção/ Junção (PJNF), se somente se cada dependência de junção em R é uma consequência de chaves candidatas de R

Centro de Informática
57

Normalização

Notas:

- Qualquer relação 5NF está na 4NF
- Resultados demonstrados por Fagin
- Qualquer relação dada pode ser decomposta sem perdas em uma coleção equivalente de relações em 5NF, ou seja, a 5NF é sempre realizável

Centro de Informática
58

Normalização

- Resultados demonstrados por Fagin (Cont.)
 - Dada uma relação R, podemos dizer se R está na 5NF se conhecermos suas chaves candidatas e todas as DJ em R
 - Embora seja fácil identificar DF e DMV, o mesmo não pode ser dito da DJ porque seu significado intuitivo não é óbvio

Centro de Informática
59

Resumo da Normalização

Eliminar atributos não atômicos

1NF

Eliminar DF não plenas

2NF

Eliminar dependências transitivas

3NF

Eliminar DF cujo determinante não é chave candidata

BCNF

Eliminar DMV

4NF

Eliminar DJ (se encontradas)

5NF

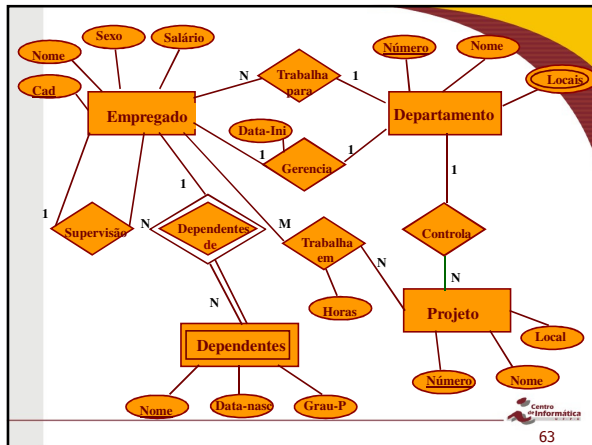
Para evitar problemas com a 4NF: Remover para outras relações atributos multivalorados quando da aplicação da 1NF

Centro de Informática
60

Mapeamento E/R - Relacional

E-R/Relacional

- Um esquema relacional pode ser facilmente derivado de um esquema conceitual desenvolvido usando o modelo E-R
- Seja o exemplo de modelagem E-R a seguir, utilizando a notação R. Elmasri & S. Navathe



E-R/Relacional

- Passo 1
Para cada entidade regular E no esquema E-R, criar uma relação R que inclui todos os atributos de R
- Exemplo

Empregado	Cad	Nome	Sexo	Salário
-----------	-----	------	------	---------

E-R/Relacional

- Passo 2
Para cada entidade fraca W no esquema E-R que tenha como entidade proprietária E
 - Criar uma relação R e incluir todos os atributos de W como atributos de R
 - Incluir o(s) atributo(s) chave primária da relação proprietária E
 A chave primária de R é a combinação da chave primária de E e da chave parcial de W

Dependente	Cad	Nome	Data-nasc	Grau-P
------------	-----	------	-----------	--------

E-R / Relacional


- Passo 3
Para cada relacionamento R de 1:1 no esquema ER
 - Identificar as relações das entidades participando do relacionamento R
 - Escolher uma das relações e incluir como chave estrangeira a chave primária da outra relação
 - Incluir todos os atributos do relacionamento na relação escolhida

E-R/Relacional

- Passo 3 (cont)
 - Exemplo

Departamento	Número	Nome	Cad-Ger	Data-Ini
--------------	--------	------	---------	----------

Note que um mapeamento alternativo de um relacionamento 1:1 é possível juntando as duas entidades em uma única relação. Isto é apropriado quando as entidades não participam de outros relacionamentos.




67

E-R / Relacional

- Passo 4
 - Para cada relacionamento regular (não fraco) 1:N
 - Identificar a relação S que representa a entidade do lado N
 - Incluir como chave estrangeira a chave primária da relação que representa a entidade do lado 1
 - Incluir os atributos do relacionamento em S

Empregado	Cad	Nome	Sexo	Salário	Num-Dep	Cad-Spv
-----------	-----	------	------	---------	---------	---------




68

E-R/Relacional

- Passo 5
 - Para cada relacionamento R de M:N
 - Criar uma nova relação para representar R
 - Incluir como chave estrangeira as chaves primárias das relações que participam em R, estas chaves combinadas formarão a chave primária da relação
 - Incluir também eventuais atributos de R

Trabalha-em	Cad-Emp	Num-Proj	Horas
-------------	---------	----------	-------




69

E-R/Relacional

- Passo 6
 - Para cada atributo multivalorado A, criar uma nova relação R, incluindo um atributo correspondendo a A mais a chave primária K da relação que tem A como atributo
 - A chave primária de R é a combinação de A e K


Locais	Num-Dep	Num-Loc
--------	---------	---------



70

E-R/Relacional


- Passo 7
 - Para cada relacionamento n-ário R (n > 2)
 - Criar uma nova relação S para representar R
 - Incluir como chaves estrangeiras as chaves primárias das relações que representam as entidades participantes
 - Incluir os eventuais atributos de R
 - A chave primária de S é normalmente a combinação das chaves estrangeiras



71

Esquema Relacional

Empregado	Cad	Nome	Sexo	Salário	Num-Dep	Cad-Spv
Departamento	Número	Nome	Cad-Ger	Data-Ini		
Locais	Num-Dep	Num-Loc				
Projeto	Número	Nome	Num-Dep			
Trabalha-em	Cad-Emp	Num-Proj	Horas			
Dependente	Cad	Nome	Data-nasc	Grau-P		



72