

DCC011: Introdução a Banco de Dados

Rodrygo Santos

rodrygo@dcc.ufmg.br

Departamento de Ciência da Computação
Universidade Federal de Minas Gerais

Mapeamento ER - Relacional

1. Transformações entre Modelos
2. Algoritmo de mapeamento ER → Relacional

Regras gerais de tradução

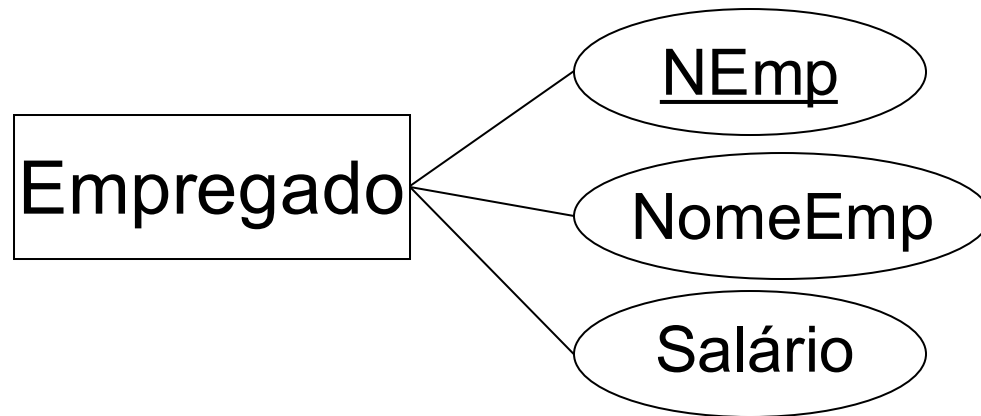
- Evitar junções
 - Acesso randômico
- Diminuir o número de chaves
 - Espaço extra + gerenciamento
- Evitar campos opcionais
 - Ambiguidade no tratamento de NULLs

Algoritmo de Mapeamento

Elmasri & Navathe

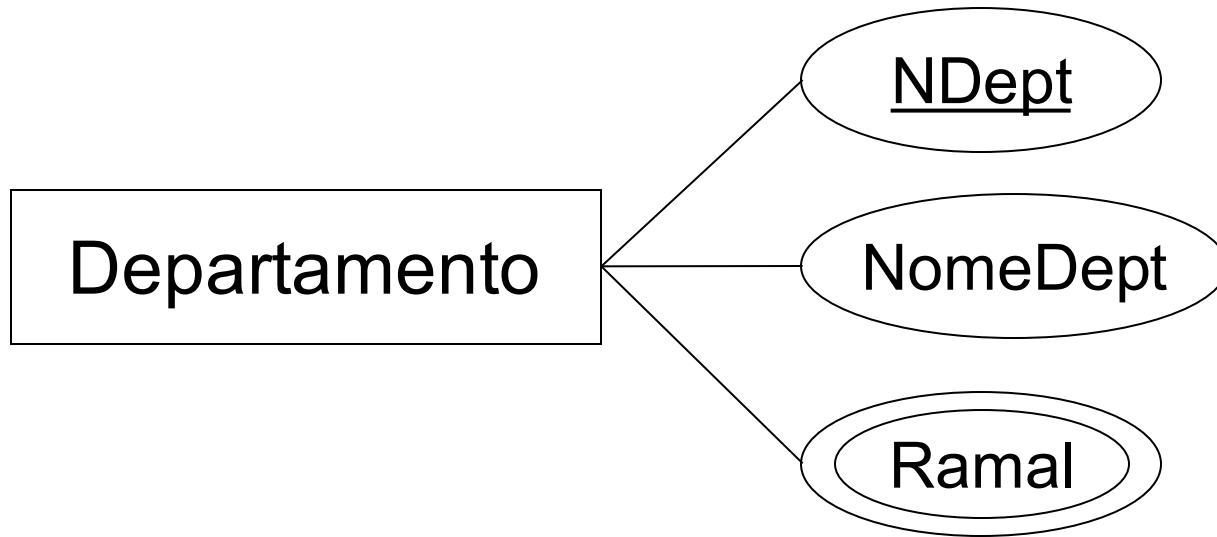
- a. Entidades regulares
- b. Atributos multivalorados
- c. Entidades fracas
- d. Relacionamentos
 - d.1 Relacionamentos binários 1:1
 - d.2 Relacionamentos binários 1:N
 - d.3 Relacionamentos binários N:M
 - d.4 Relacionamentos N-ários
- e. Hierarquias
(especialização/generalização)

a. Entidades regulares
(sem atributos multivalorados)



Empregado (NEmp, NomeEmp, Salário)

b. Atributos Multivalorados

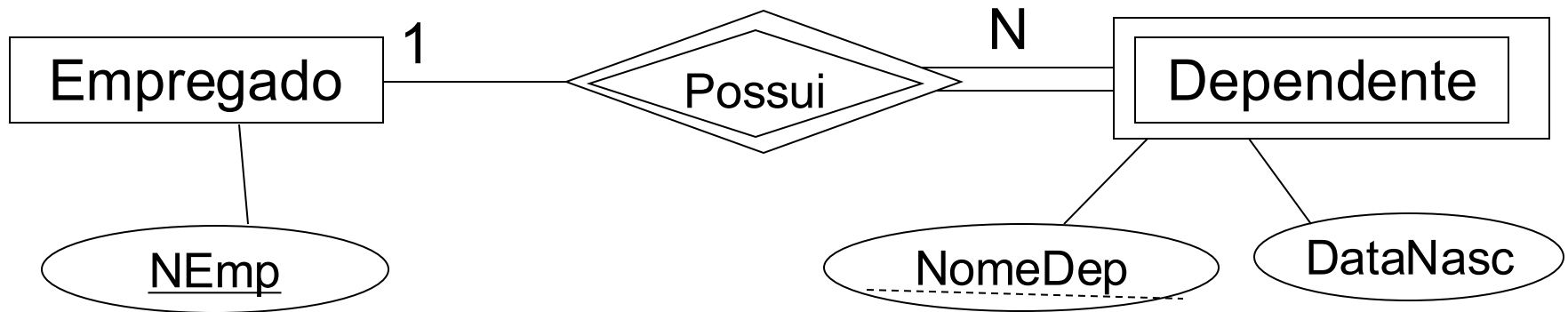


Departamento (NDept, NomeDept)

Ramal-Departamento (NDept, Ramal)

NDept referencia Departamento, por propagação

c. Entidade Fraca



Empregado (NEmp,...)

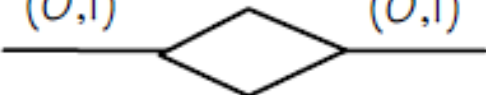
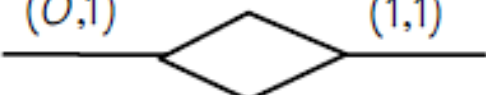
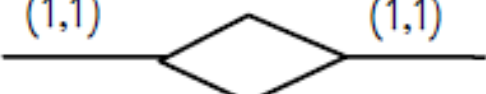
Dependente (NEmp, NomeDep, DataNasc)

NEmp referencia Empregado, **por propagação**

d. Relacionamentos

- Tabela própria
- Adição de colunas a uma das tabelas
- Fusão de tabelas
- Alternativa depende da cardinalidade (máxima e mínima) do relacionamento
 - d.1 Relacionamentos binários 1:1
 - d.2 Relacionamentos binários 1:N
 - d.3 Relacionamentos binários N:M
 - d.4 Relacionamentos N-ários

d.1. Relacionamento binário (1:1)



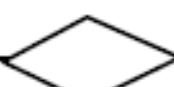
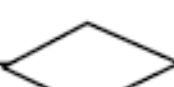
Tipo de relacionamento			Regra de implementação		
			Tabela própria	Adição coluna	Fusão tabelas
	(0,1)  (0,1)		±	✓	×
	(0,1)  (1,1)		×	±	✓
	(1,1)  (1,1)		×	×	✓

✓ Alternativa preferida

± Pode ser usada

× Não usar

d.2. Relacionamentos binários (1:N)

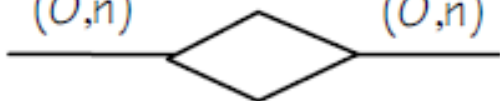
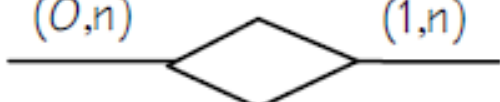
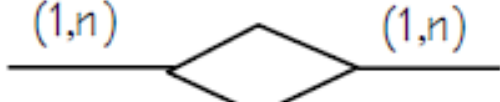
Tipo de relacionamento			Regra de implementação		
			Tabela própria	Adição coluna	Fusão tabelas
	(O,1)  (O,n)		±	✓	×
	(O,1)  (1,n)		±	✓	×
	(1,1)  (O,n)		×	✓	×
	(1,1)  (1,n)		×	✓	×

✓ Alternativa preferida

± Pode ser usada

× Não usar

d.3 Relacionamento binário (N:M)

Tipo de relacionamento				Regra de implementação		
				Tabela própria	Adição coluna	Fusão tabelas
	(0,n) 			✓	×	×
	(0,n) 			✓	×	×
	(1,n) 			✓	×	×

✓ Alternativa preferida

× Não usar

d.4. Relacionamento N-ario

- Não são definidas regras específicas
 - O relacionamento é transformado em uma entidade
- Nova entidade Rel
 - Colunas = chaves primárias das tabelas relacionadas

e. Hierarquias

- e.1. Relações : superclasse e subclasses
 - Funciona para total/partial + disjoint/overlapping
- e.2. Relações : subclasses
 - Funciona somente para total + disjoint
 - Precisa de OUTER UNION (ou FULL OUTER JOIN) para obter todas as instâncias da superclasse
- e.3. Relação única (disjoint)
- e.4. Relação única + tipos (overlapping)
 - Trade-off esparsidade vs. eficiência

Normalização

- Processo através do qual esquemas de relação são sucessivamente decompostos até que satisfaçam determinadas propriedades ou **formas normais**.
- Medidas informais de qualidade do projeto de um esquema relacional:
 - Atributos das relações semanticamente compatíveis
 - Ausência de valores redundantes nas relações (inexistência de anomalias de atualização)
 - Número reduzido de valores nulos nas relações
 - Sem possibilidade de gerar tuplas espúrias/falsas

Primeira Forma Normal (1NF)

- Primeira Forma Normal (1FN):

Um esquema de relação R está na 1FN se todos os seus atributos forem **atômicos**.

= simples e monovalorados

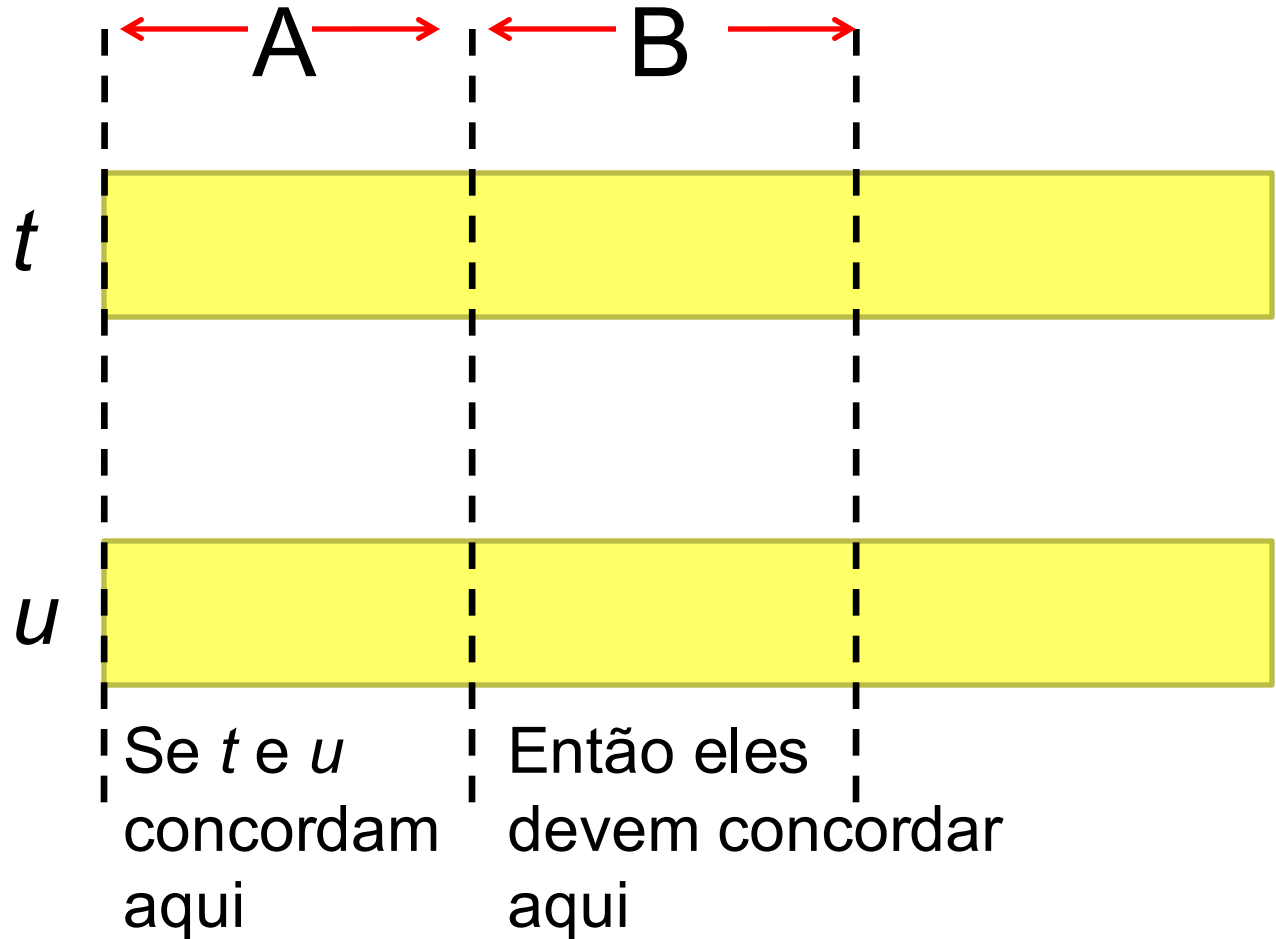
- Contra-exemplo

- Atributo composto
- Atributo multivalorado

Dependências Funcionais

$A \rightarrow B$

Se duas tuplas em R concordam nos atributos $A1, A2, \dots, AN$ (i.e. as tuplas têm os mesmos valores), então eles devem concordar em um outro atributo, B



Segunda Forma Normal (2FN)

- Um esquema de relação R está na 2FN se todo atributo de R não pertencente a uma de suas chaves for totalmente dependente da chave primária
- Contra-exemplo
 - Dependências parciais

Terceira Forma Normal (3FN)

- Um esquema de relação R está na 3FN se estiver na 2FN e nenhum atributo não-chave de R depender de outro atributo não-chave
- Contra-exemplo
 - Dependência transitiva

Análise Exploratória de Dados

- Entendimento de dados
 - Tipos de dados, dados faltantes, dados duplicados, dados incorretos
- Limpeza de dados
 - Imputação, descarte
- Análise univariada e bivariada
 - Distribuições, correlações
 - Dependências funcionais

Data modeling goals

- Data-oriented modeling
 - Driven by the structure of the data

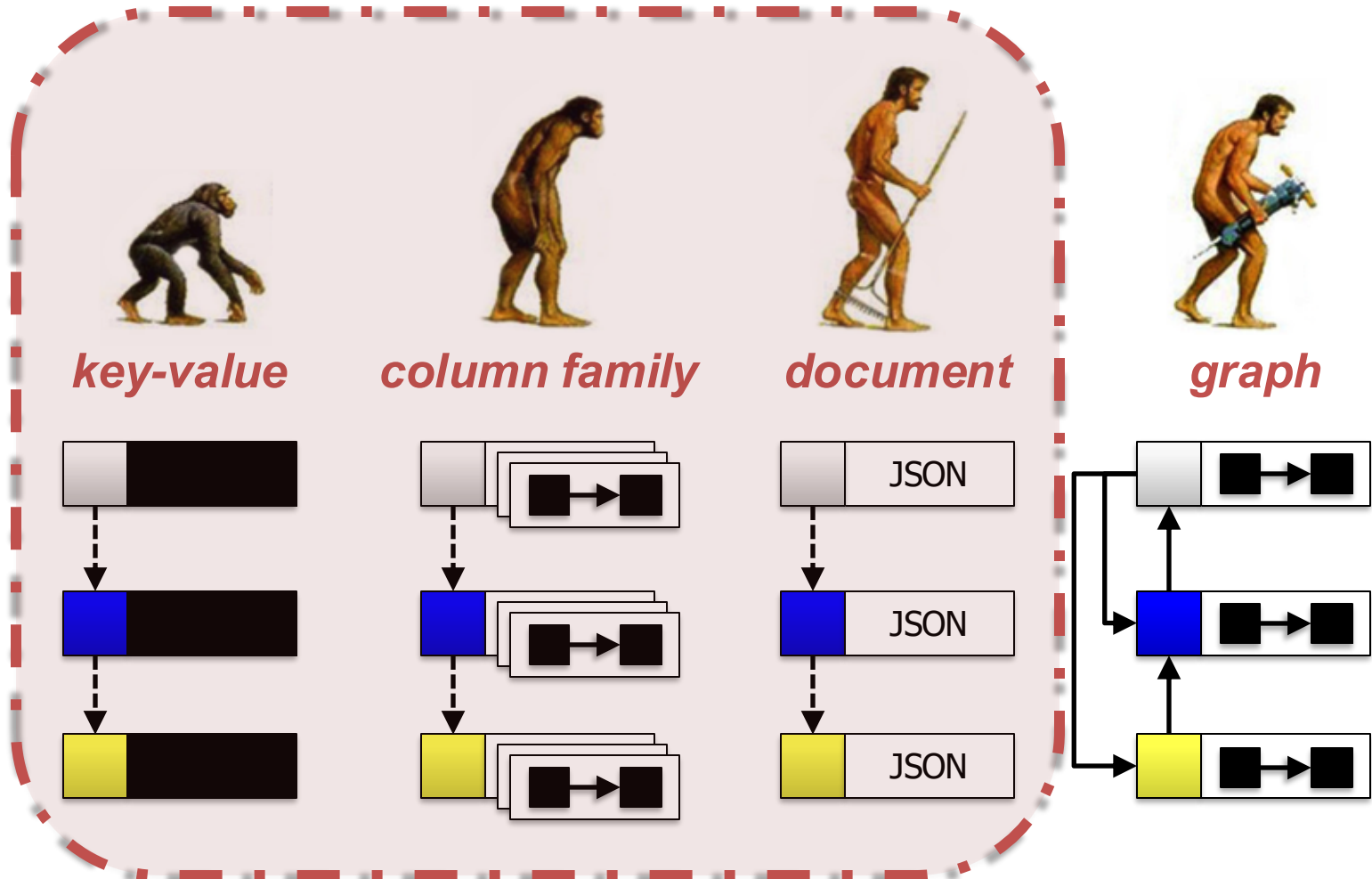
"What answers do I have?"

- Query-oriented modeling
 - Driven by access patterns

"What questions do I have?"

NoSQL models

aggregate models



Cronograma

- 16/06: Revisão
- **18/06: Prova 3**
- 23/06: Apresentação TP2
- 25/06: Apresentação TP2
- **30/06: Prova Substitutiva**
- **02/07: Exame Especial**