

# **1.264 – Aula 6**

**Diagramas de entidade/ relação e  
normalização de dados**

# Empresa Transportadora MIT

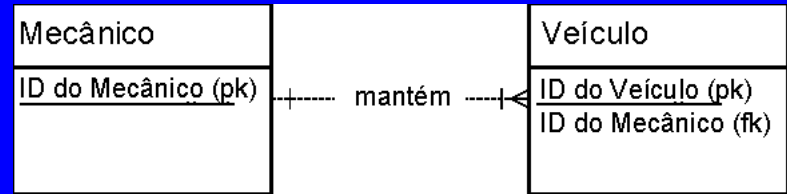
- **Exercício**
  - **Determinar e nomear tipos de entidades**
  - **Determinar cardinalidade (1:N, N:N) e o tipo de relações (obrigatória/opcional)**
  - **Adicionar identificadores e chaves externas**
  - **Rotular frase de relação**

## Empresa Transportadora MIT

- **A Transportadora emprega mecânicos para manter os veículos. Para cada mecânico se atribui pelo menos um veículo, e freqüentemente muitos. Um veículo está sempre sob a responsabilidade de 1 mecânico.**

# Empresa Transportadora MIT

- A Transportadora emprega mecânicos para manter os veículos. Para cada mecânico se atribui pelo menos um veículo, e freqüentemente muitos. Um veículo está sempre sob a responsabilidade de 1 mecânico.

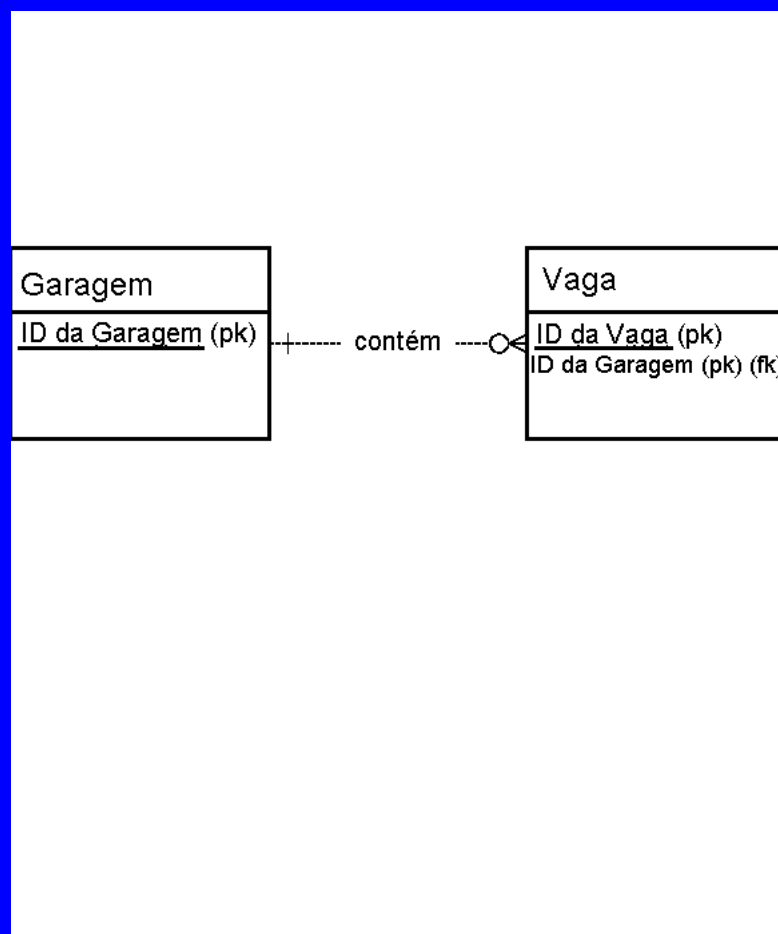


# Empresa Transportadora MIT

- A transportadora tem varias garagens. Uma garagem pode conter muitas vagas. Entretanto, nem todas as garagens têm vagas. Uma vaga pode estar em uma garagem.

# Empresa Transportadora MIT

- A transportadora tem varias garagens. Uma garagem pode conter muitas vagas. Entretanto, nem todas as garagens têm vagas. Uma vaga pode estar em uma garagem.

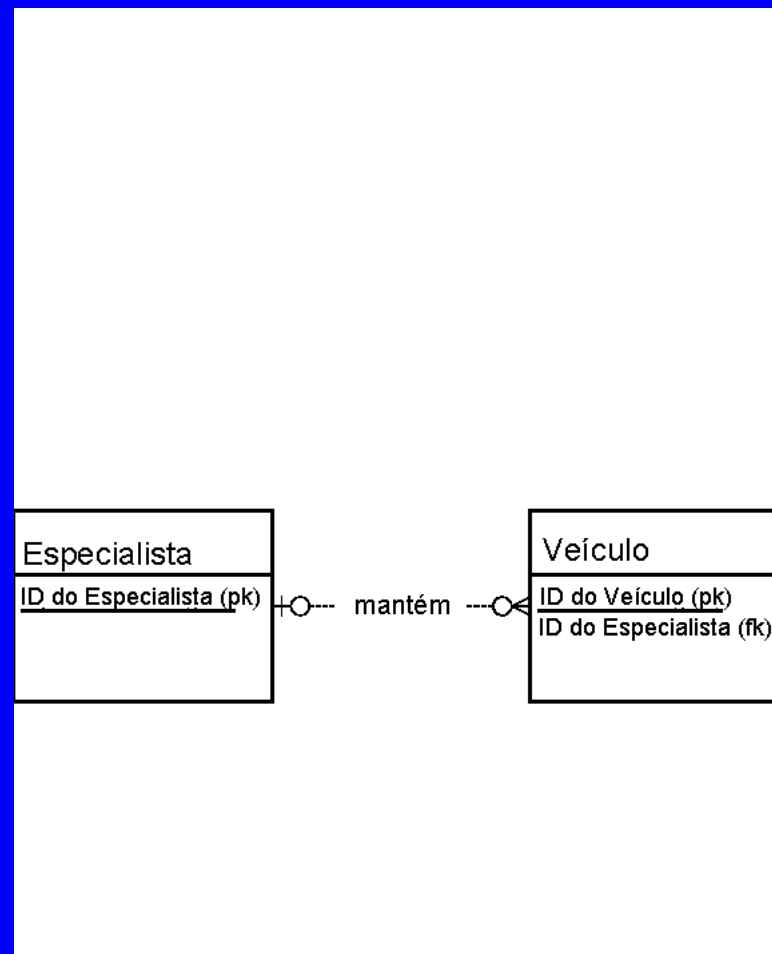


# Empresa Transportadora MIT

- **A garagem mantém uma lista de pessoal especializado em reparo, que é utilizado em caso de necessidade. Um especialista pode trabalhar em muitos veículos. Um veículo pode ou não ser reparado por um especialista.**

# Empresa Transportadora MIT

- A garagem mantém uma lista de pessoal especializado em reparo, que é utilizado em caso de necessidade. Um especialista pode trabalhar em muitos veículos. Um veículo pode ou não ser reparado por um especialista.

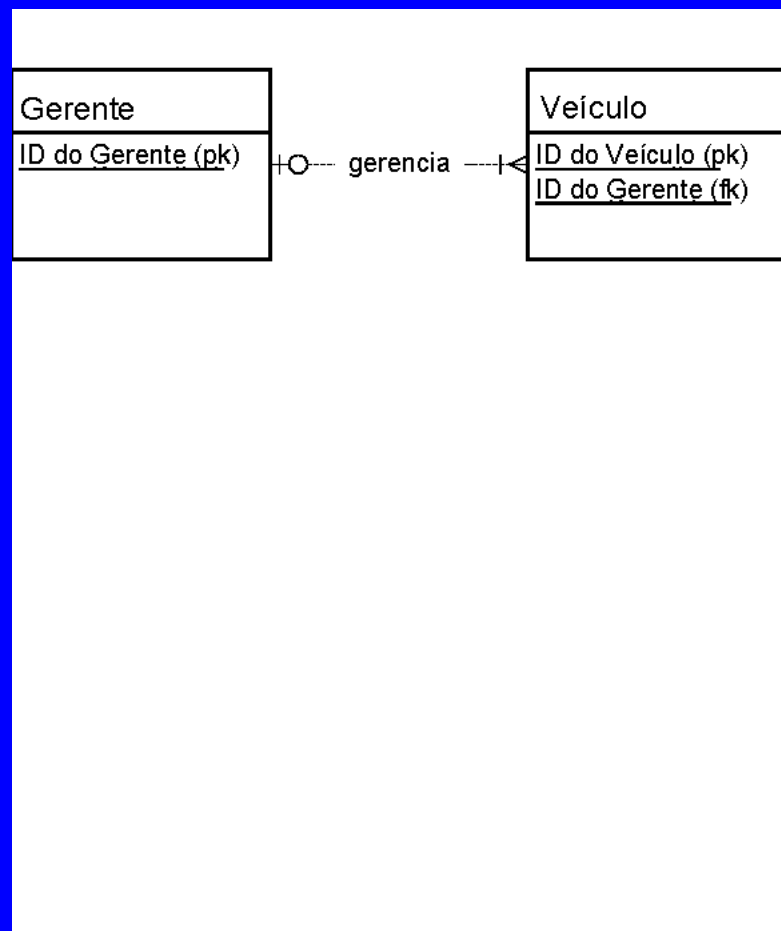


# Empresa Transportadora MIT

- A garagem emprega gerentes para monitorar os custos de disponibilidade e reparo. A cada veículo é atribuído pelo menos 1, e, freqüentemente, muitos veículos. Um veículo pode ou não ter um gerente responsável.

# Empresa Transportadora MIT

- A garagem emprega gerentes para monitorar os custos de disponibilidade e reparo. A cada veículo é atribuído pelo menos 1, e, freqüentemente, muitos veículos. Um veículo pode ou não ter um gerente responsável.



# Normalização

- **Regras de Normalização**
  - Evitar anomalias de atualização (equívocos) e inconsistência de dados
  - Degradar desempenho, em geral apenas levemente
    - Maior impacto nas leituras, em que várias linhas contra uma são lidas
    - Menor impacto nas gravações, o qual tende a ser de qualquer maneira um retardo de progresso
    - A desnormalização é comum em bancos de dados de somente leitura nos quais o alto desempenho é necessário (por exemplo, base de dados somente leitura da Web)
  - O projeto do banco de dados e a configuração de disco (arquitetura) interagem
    - Vamos estudar arquitetura mais para o final do semestre
  - Os dados iniciais de seu trabalho de casa 2 não são normalizados. A tarefa 2 pede que você normalize seus dados, com máxima precisão.

## **Cinco formas normais**

- **1: Todas as ocorrências de uma entidade devem conter o mesmo número de atributos. Sem listas.**
- **2: Todos os campos não-chave devem ser uma função da chave.**
- **3: Todos os campos não-chave não devem ser uma função de outros campos não-chave.**
- **4: Uma linha não deve conter dois ou mais fatos de valores múltiplos independentes sobre uma entidade.**
- **5: Um registro não pode ser reconstruído a partir de vários outros tipos de registros menores.**

# Definições

- **Linha ou registro:** um conjunto fixo de atributos (campos) que descreve uma instância de uma entidade.
- **Chave:** um identificador único para uma linha de uma tabela, utilizado para selecionar linhas em consultas. Pode ser composto por vários campos.
- **Não-chave:** todos os outros campos da linha
- **Entidade:** objeto definido no modelo de sistema sobre quais dados são armazenados no banco de dados. Uma tabela em um banco de dados relacional.

# Primeira forma normal

- **Todas as linhas devem ter um comprimento fixo.**
  - Uma hipótese restritiva, não um princípio de projeto.
  - Não permite listas de tamanho variável.

# Segunda forma normal

Peça	Armazém	Quantidade	WarehouseAddress
42	Boston	2000	24 Main St
333	Boston	1000	24 Main St
390	Nova Iorque	3000	99 Broad St

- Todos os campos não-chave devem ser uma função da chave toda.
  - Exemplo de violação da segunda forma normal:
    - A chave é Peça + Armazém.
    - Alguém achou conveniente acrescentar o Endereço para tornar um relatório mais fácil.
    - WarehouseAddress (Endereço do Armazém) é um fato de Armazém e não de Peça.
  - Problemas:
    - O endereço do armazém é repetido em cada linha que se refere a uma peça armazenada em um armazém.
    - Se o endereço do armazém mudar, cada linha que se refere a uma peça armazenada nesse armazém deve ser atualizada.
    - Os dados podem se tornar inconsistentes, com diferentes registros mostrando diferentes endereços para o mesmo armazém.
    - Se em algum momento não houver peças armazenadas no armazém, não deve haver nenhum registro onde manter o endereço do armazém.

# Segunda forma normal

- **Solução**
  - **Dois tipos de entidade: Inventário e Armazém.**
  - **Vantagem: resolve problemas como o do slide anterior**
  - **Desvantagem: se a aplicação necessitar do endereço de cada armazém que estoca uma peça, deve acessar duas tabelas em vez de uma. Isso costumava a ser um problema, mas é raro agora.**

<b>Peça</b>	<b>Armazém</b>	<b>Quantidade</b>
<b>42</b>	<b>Boston</b>	<b>2000</b>
<b>333</b>	<b>Boston</b>	<b>1000</b>
<b>390</b>	<b>Nova Iorque</b>	<b>3000</b>

<b>Armazém</b>	<b>Endereço do Armazém</b>
<b>Boston</b>	<b>24 Main St</b>
<b>Nova Iorque</b>	<b>99 Broad St</b>

# Terceira forma normal

Funcionário	Departamento	Localização do Departamento
234	Financeiro	Boston
223	Financeiro	Boston
399	Operações	Washington

- **Campos não-chave não podem ser uma função de outros campos não-chave.**
  - **Exemplo de violação do terceira forma normal**
  - **A chave é funcionário**
  - **Alguém achou conveniente acrescentar a localização do departamento em um relatório**
    - **A localização do departamento é uma função do departamento, que não é uma chave**
  - **Problemas:**
    - **A localização do departamento é repetida em cada registro de funcionário**
    - **Se a localização do armazém mudar, cada registro que contenha esse dado deve ser alterado**
    - **Os dados podem ser tornar inconsistentes**
    - **Se um departamento não tiver funcionários, pode não haver lugar para armazenar sua localização**

# Terceira forma normal

- **Solução**
  - **Dois tipos de entidades: funcionário e departamento**

<b>Funcionário</b>	<b>Departamento</b>
<b>234</b>	<b>Financeiro</b>
<b>223</b>	<b>Financeiro</b>
<b>399</b>	<b>Operações</b>

<b>Departamento</b>	<b>Localização do Departamento</b>
<b>Financeiro</b>	<b>Boston</b>
<b>Operações</b>	<b>Washington</b>

## Quarta forma normal

Funcionário	Habilidade	Idioma
Brown	cozinhar	Inglês
Smith	digitar	Alemão

- **Uma linha não deve conter dois ou mais fatos de valores múltiplos independentes sobre uma entidade.**
  - **Exemplo de violação do quarta forma normal:**
    - Um funcionário pode ter várias habilidades e conhecer vários idiomas
  - **Problemas**
    - Incerteza em como manter as linhas. Várias abordagens são possíveis e diferentes programadores podem tomar diferentes caminhos, como mostrado no próximo slide

# Problemas da quarta forma normal

Funcionário	Habilidade	Idioma
Brown	cozinhar	
Brown	digitar	
Brown		Francês
Brown		Alemão
Brown		Grego

- **Formato desarticulado. Efetivamente o mesmo que 2 tipos de entidade.**
  - **Campos em branco ambíguos. A habilidade em branco poderia significar que:**
    - A pessoa não tem habilidades
    - O atributo não se aplica ao funcionário
    - O dado é desconhecido
    - O dado pode ser encontrado em outro registro (como nesse caso)
  - **Os programadores considerarão todas essas hipóteses o tempo todo, o mesmo será feito pelo pessoal responsável pelo preenchimento dos dados e os usuários**

# Problemas da quarta forma normal, continuação

Funcionário	Habilidade	Idioma
Brown	cozinhar	Francês
Brown	cozinhar	Alemão
Brown	cozinhar	Grego
Brown	digitar	Francês
Brown	digitar	Alemão
Brown	digitar	Grego

- **Formato de resultados cruzados. Problemas:**
  - **Repetições:** as atualizações devem ser feitas em muitos registros, e pode haver inconsistências
  - **A inserção de uma nova habilidade** pode envolver a procura de um registro de habilidade em branco, a inserção de um novo registro com, possivelmente, um idioma ou uma habilidade em branco, ou a inserção de um novo registro emparelhando a habilidade com alguns ou todos os idiomas.
  - **A anulação é pior:** significa apagar uma habilidade em um ou mais registros, e em seguida ter que verificar se não ficaram 2 registros com o mesmo idioma e nenhuma habilidade, ou pode significar o apagamento de um ou mais registros, certificando-se de não apagar a última referência de um idioma que não deve ser apagado.

# Solução para a quarta forma normal

- Solução: dois tipos de entidades
  - Funcionário – habilidade e funcionário – idioma

Funcionário	Habilidade
Brown	cozinhar
Brown	digitar

Funcionário	Idioma
Smith	Francês
Smith	Alemão
Smith	Grego

- Observe que habilidades e idiomas podem estar relacionados, caso em que o exemplo inicial estava correto:
  - Se Smith só souber cozinhar comida francesa, e puder digitar em francês e grego, então habilidade e idioma não são fatos independentes múltiplos sobre o funcionário, logo não se violou a quarta forma normal.

# Quinta forma normal

- Um registro não pode ser reconstruído a partir de vários tipos de registros menores.
- Exemplo:
  - Os vendedores representam as empresas
  - As empresas fazem produtos
  - Os vendedores vendem os produtos
- Caso mais geral (permite qualquer combinação):

Vendedor	Empresa	Produto
Smith	Ford	carro
Smith	GM	caminhão

- O Smith não vende nem caminhões da Ford, nem carros da GM

# Quinta forma normal

- É em uma alteração que ocorre o problema
  - Se um vendedor trabalha com um determinado produto, e ele representa uma empresa, então ele vende esses produtos pela empresa.

Vendedor	Empresa	Produto
Smith	Ford	carro
Smith	Ford	caminhão
Smith	GM	carro
Smith	GM	caminhão
Jones	Ford	carro

(Repetição de fatos)

- Podemos reconstruir todos os fatos verdadeiros em 3 tabelas em vez de 1:

Vendedor	Empresa
Smith	Ford
Smith	GM
Jones	Ford

Vendedor	Produto
Smith	carro
Smith	caminhão
Jones	carro

Empresa	Produto
Ford	carro
Ford	caminhão
GM	carro
GM	caminhão

(Sem repetição de fatos)

# Quinta forma normal

- **Problemas com a forma da tabela 1**
  - Os fatos são registrados várias vezes. Por exemplo, o fato de que o Smith vende carros é registrado duas vezes. Se Smith parar de vender carros, 2 linhas devem ser atualizadas e faltará uma.
  - O tamanho desta tabela aumenta em progressão geométrica, enquanto as tabelas normalizadas crescem em progressão aritmética. Para grandes operações, isso faz uma grande diferença.
    - $100.000 \times 100.000$  é bem maior do que  $100.000 + 100.000$
- **É mais fácil escrever as regras comerciais a partir da quinta normal**
  - As regras são mais explícitas
  - (Cadeias de suprimentos têm todos os tipos de problemas da 5ª. normal)

# Quinta forma normal, conclusão

- Um exemplo de um conjunto sutil de condições

Não-normal

Vendedor	Empresa	Produto
Smith	Ford	carro
Smith	Ford	caminhão
Smith	GM	carro
Smith	GM	caminhão
Jones	Ford	carro
Jones	Ford	caminhão
Brown	Ford	carro
Brown	GM	carro
Brown	Toyota	carro
Brown	Toyota	ônibus

Quinta normal

Vendedor	Empresa
Smith	Ford
Smith	GM
Jones	Ford
Brown	Ford
Brown	GM
Brown	Toyota

Empresa	Produto
Ford	carro
Ford	caminhão
GM	carro
GM	caminhão
Toyota	carro
Toyota	ônibus

Vendedor	Produto
Smith	carro
Smith	caminhão
Jones	carro
Jones	caminhão
Brown	carro
Brown	ônibus

- Jones vende carros e a GM fabrica carros, mas Jones não é representante da GM
- Brown é representante da Ford e a Ford fabrica caminhões, mas Brown não vende caminhões
- Brown é representante da Ford e Brown vende ônibus, mas a Ford não fabrica ônibus

# Moral da história

- **Sistemas são efêmeros**
- **Dados são permanentes**
- **Se você desordenar um sistema, é só reescrevê-lo e o problema estará resolvido**
- **Se você desordenar os dados, eles geralmente ficam irrecuperáveis**
- **Negócios reais têm regras comerciais sutis**
  - **É necessário dar a devida atenção à modelagem dos dados e às regras comerciais para se conseguir uma boa qualidade de dados**
  - **É necessário normalizar os dados para preservar sua qualidade**

# Ferramentas

- **Modelagem de dados**
  - Visio: barato, amplamente utilizado para modelagem inicial
    - Útil para requisitos iniciais ao longo de empresas múltiplas
  - ERWin, PowerDesigner, Oracle Designer, outros... são mais poderosos
    - Podem gerar banco de dados físicos no Oracle, e outros.

# Empresa Transportadora MIT

- A transportadora emprega mecânicos para manter os veículos. Para cada mecânico se atribui pelo menos um veículo, e freqüentemente muitos. Um veículo está sempre sob a responsabilidade de 1 mecânico.
- A transportadora tem varias garagens. Uma garagem pode conter muitas vagas. Entretanto, nem todas as garagens têm vagas. Uma vaga deve estar em uma garagem.
- A garagem mantém uma lista de pessoal especializado em reparo, que é utilizado em caso de necessidade. Um especialista pode trabalhar em muitos veículos. Um veiculo pode ou não ser reparado por um especialista.
- A garagem emprega gerentes para monitorar os custos de disponibilidade e reparo. A cada veículo é atribuído pelo menos 1, e, freqüentemente, muitos veículos. Um veículo pode ou não ter um gerente responsável.

