

Bancos de Dados Relacionais

Uma introdução prática

PROGRAMA

- 1.Introdução
 - 1.1 Evolução do armazenamento de dados
 - 1.2 Arquitetura de um DBMS
- 2.Modelo Relacional
 - 2.1 Modelo E-R
 - 2.2 Modelo Relacional
 - 2.3 Operações

1.Introdução

1.1 Evolução do armazenamento de dados

1.2 Arquitetura de um DBMS

Evolução dos sistemas de armazenamento de massa

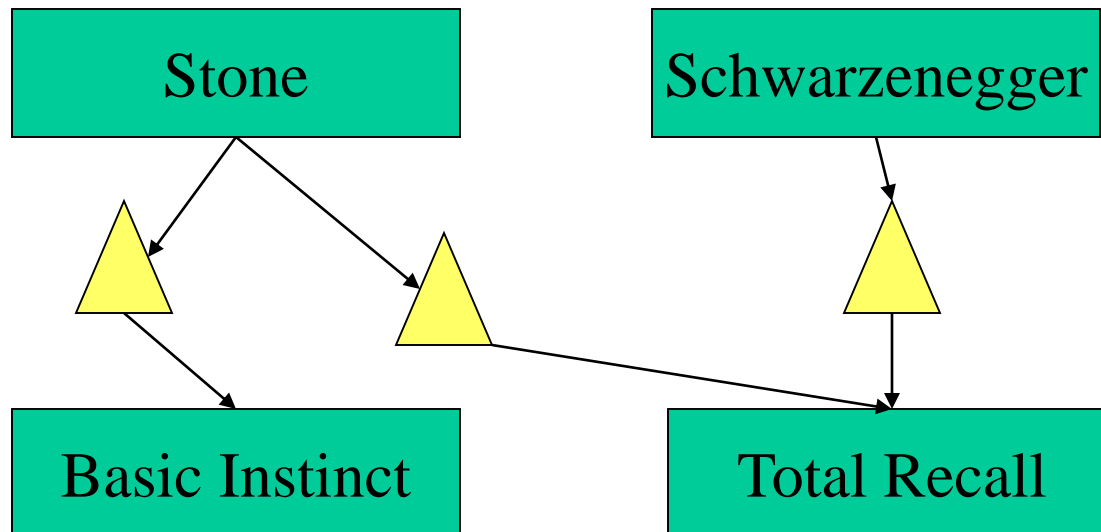
- Memória
- Sistema de arquivos
 - Acesso seqüencial
 - Acesso direto
 - Acesso indexado
- Bancos de dados

Características de um DBMS

- Usuários devem poder definir seu schema (estrutura lógica dos dados), usando uma linguagem apropriada (DDL, data definition language)
- Usuários devem poder realizar perguntas e acessos (“query”) sem conhecimento da estrutura de armazenamento dos dados, usando uma linguagem apropriada (DML, data manipulation language ou query-language)
- Usuários devem poder acessar grandes volumes de dados de forma eficiente
- Usuários devem poder acessar dados de forma independente uns dos outros, sob controle de acesso

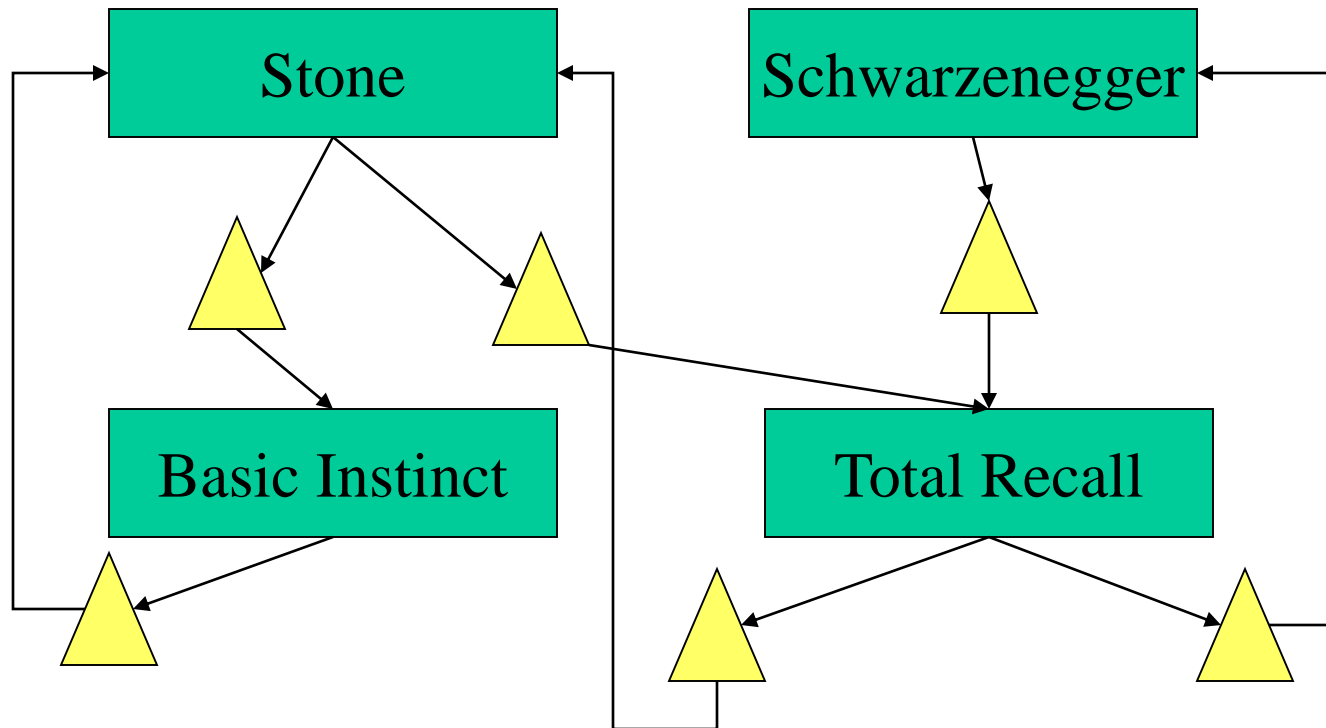
Evolução histórica dos DBMS

- Hierárquico (modelo em árvore, '60)



Evolução histórica dos DBMS

- Rede (CODASYL, fim '60)

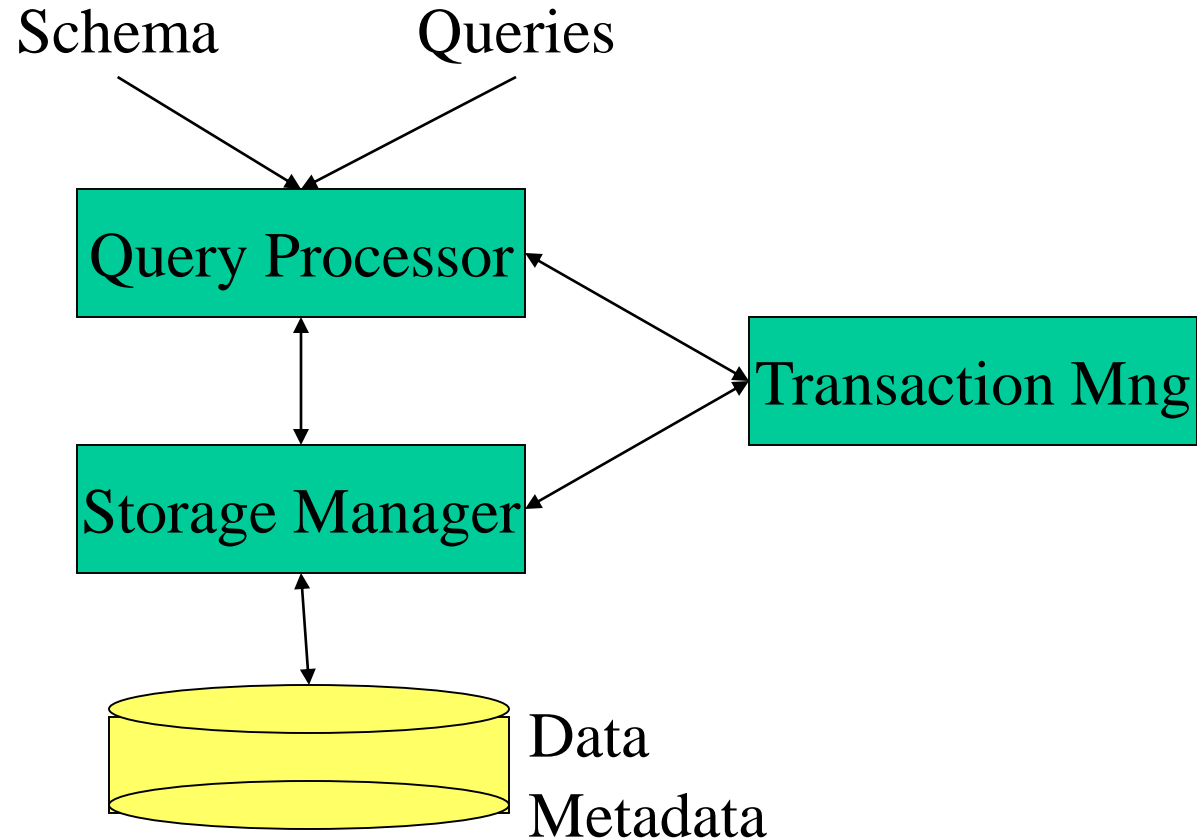


Evolução histórica dos DBMS

- Relacional (Codd, '70)

Stone	Basic Instinct
Schwarzenegger	Total Recall
Stone	Total Recall

Arquitetura de um DBMS



Storage Manager

Composto por

- File Manager, responsável pela manipulação do disco
- Buffer Manager, responsável pelos dados em memória (cache, paginação)

Query Manager

Composto por

- Query translator (de SQL para acessos a dados armazenados)
- Query optimizer

Transaction Manager

- Garante as propriedades ACID
 - Atomicidade (Atomicity): uma transação é indivisível; ou ela acontece ou não
 - Consistência (Consistency): Antes e após uma transação, o estado do banco de dados é consistente
 - Isolação (Isolation): Se duas ou mais transações acontecem simultaneamente, seus efeitos devem ser isolados
 - Durabilidade (Durability): Se uma transação se completa, seus efeitos não devem cessar mesmo que o sistema falhe imediatamente após

Transaction Manager

- Lock
 - Define o nível de isolamento
 - Página
 - Linha
- Logging
 - Armazenamento não volátil
- Commit
 - Durabilidade da transação

Transaction Manager

- Índice
 - Hash
 - Balanced tree
 - Binary tree

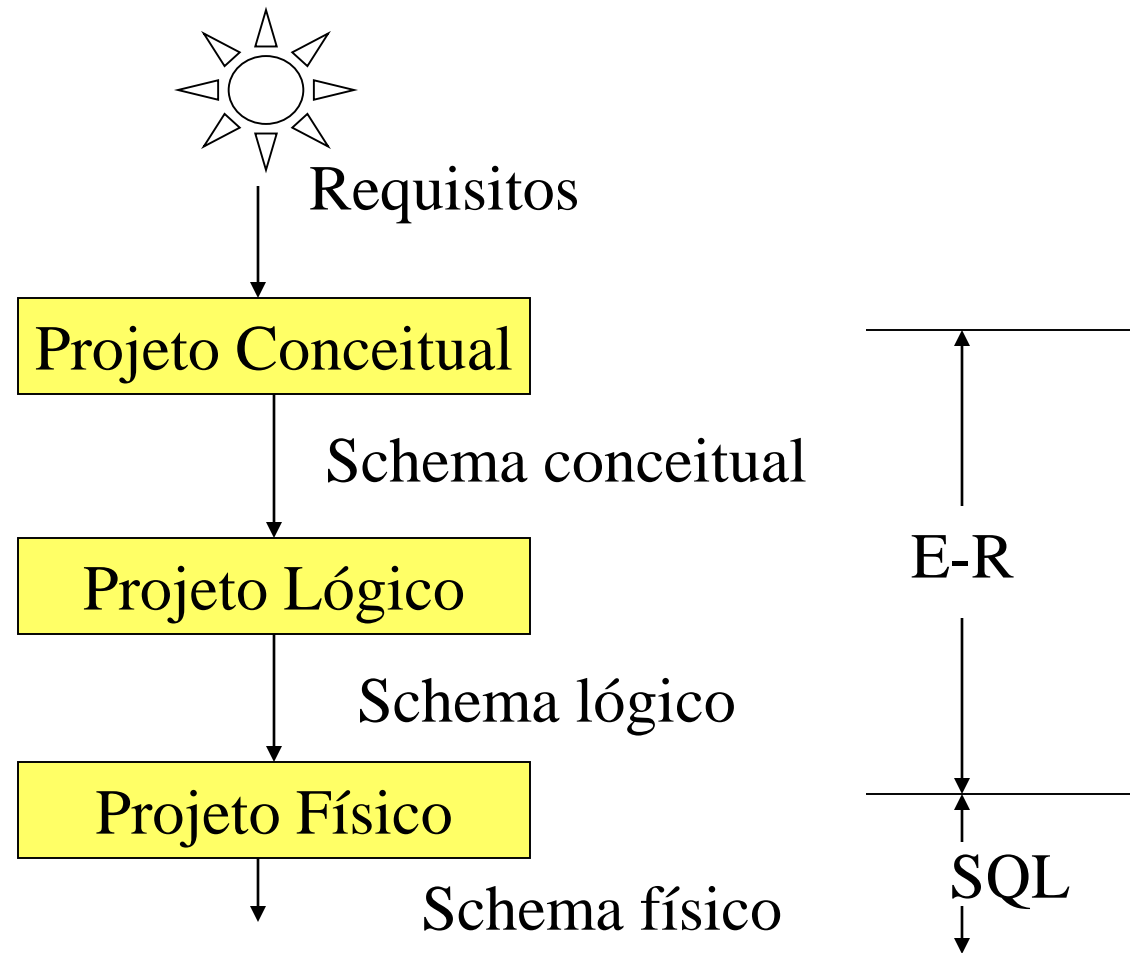
2. Modelo Relacional

2.1 Modelo E-R

2.2 Modelo Relacional

2.3 Operações

Ciclo de Projeto

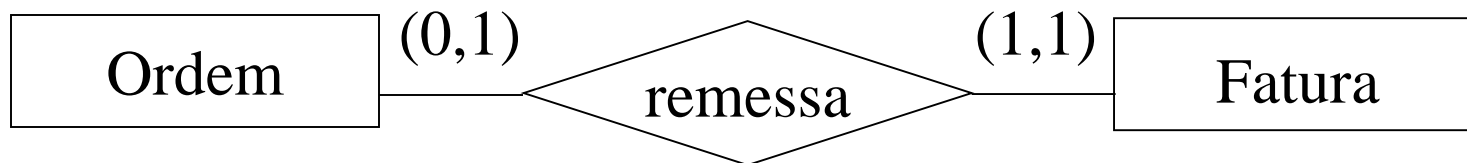


Conceitos básicos

- Entidade
 - Representam classes do mundo real
- Relacionamento
 - Representam agregações entre duas ou mais entidades
 - Os relacionamentos podem ser
 - Binários: entre duas entidades
 - N-ários: entre N entidades
 - Recursivos: de uma entidade para ela própria

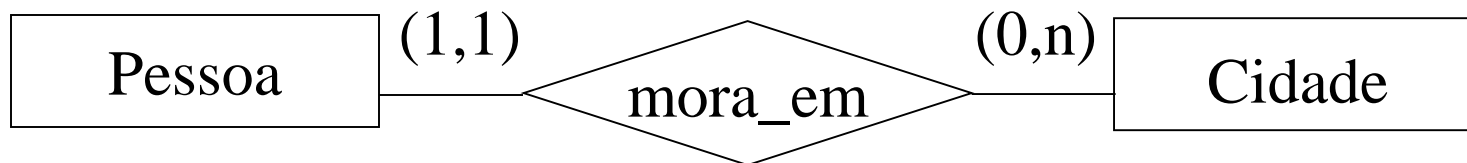
Cardinalidade dos Relacionamentos

- 1 para 1
 - $\text{min_card}(\text{Ordem}, \text{remessa}) = 0$
 - $\text{max_card}(\text{Ordem}, \text{remessa}) = 1$
 - $\text{min_card}(\text{Fatura}, \text{remessa}) = 1$
 - $\text{max_card}(\text{Fatura}, \text{remessa}) = 1$



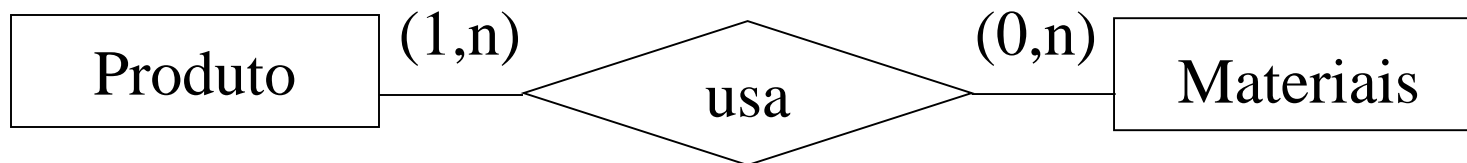
Cardinalidade dos Relacionamentos

- 1 para n
 - $\text{min_card}(\text{Pessoa}, \text{mora_em}) = 0$
 - $\text{max_card}(\text{Pessoa}, \text{mora_em}) = 1$
 - $\text{min_card}(\text{Cidade}, \text{mora_em}) = 0$
 - $\text{max_card}(\text{Cidade}, \text{mora_em}) = n$



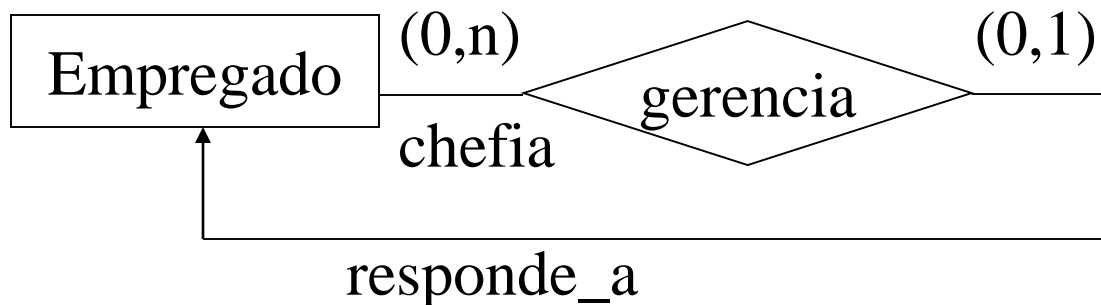
Cardinalidade dos Relacionamentos

- n para n
 - $\text{min_card}(\text{Produto}, \text{usa}) = 1$
 - $\text{max_card}(\text{Produto}, \text{usa}) = n$
 - $\text{min_card}(\text{Materiais}, \text{usa}) = 0$
 - $\text{max_card}(\text{Materiais}, \text{usa}) = n$



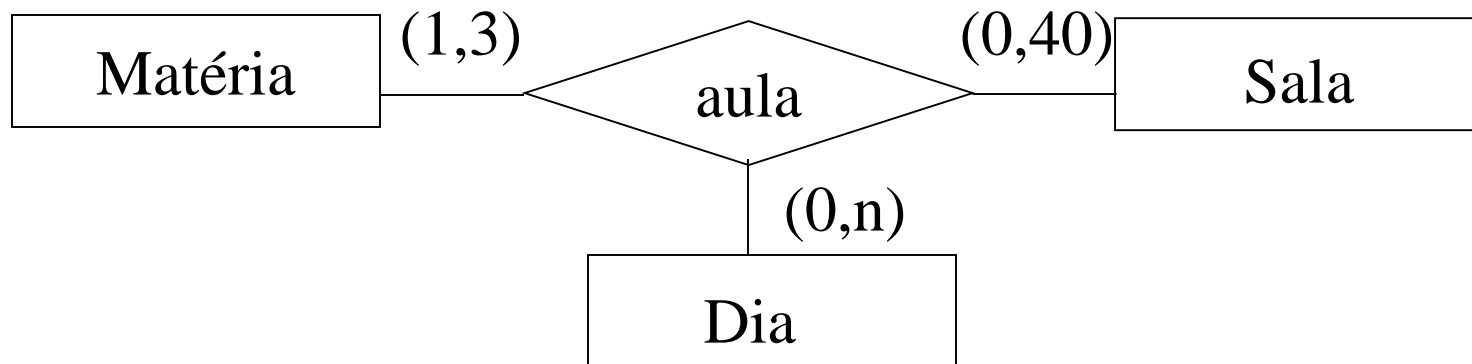
Cardinalidade dos Relacionamentos

- Recursivo (1-n)
 - $\text{min_card}(\text{Empregado}, \text{chefia}) = 0$
 - $\text{max_card}(\text{Empregado}, \text{chefia}) = n$
 - $\text{min_card}(\text{Empregado}, \text{responde_a}) = 0$
 - $\text{max_card}(\text{Empregado}, \text{responde_a}) = 1$



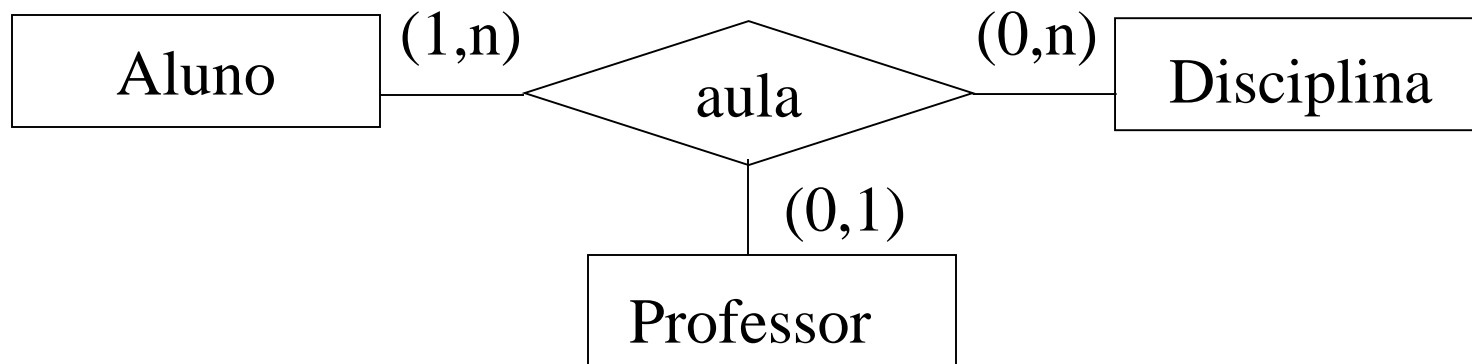
Cardinalidade dos Relacionamentos

- N-ário



Cardinalidade dos Relacionamentos

- N-ário



Cardinalidade dos Relacionamentos

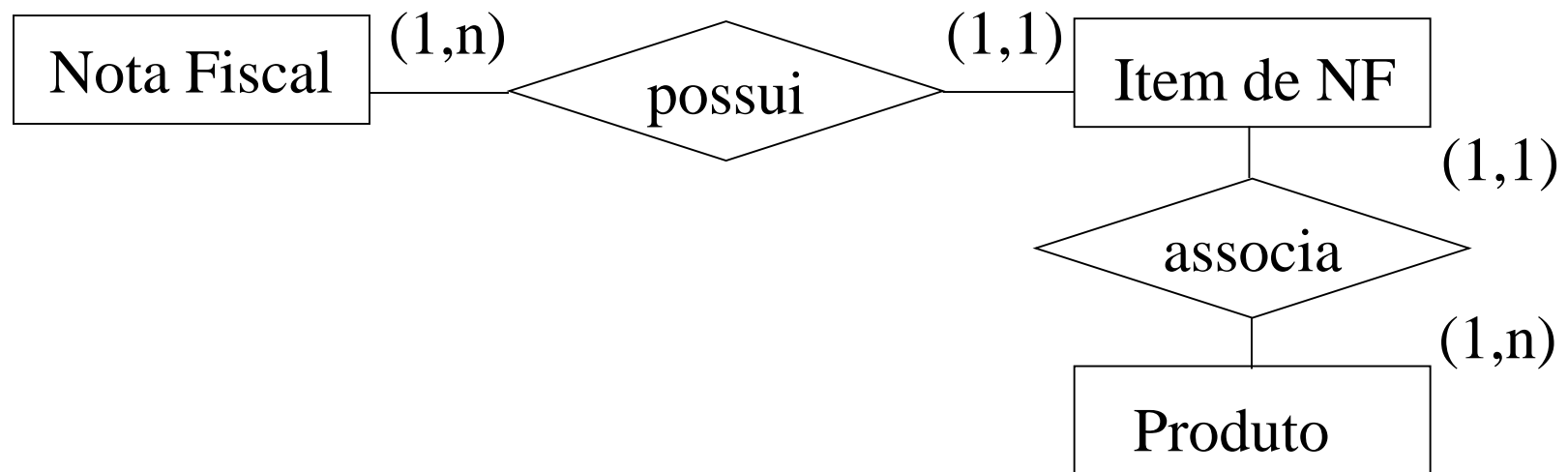
- N-ário
 - Separar a entidade *Aluno* e analisar o par *Professor, Disciplina*. Para cada par *Professor/Disciplina* podemos ter de 1 até N *Alunos* relacionados;
 - Separar a entidade *Professor* e analisar o par *Aluno, Disciplina*. Para cada par *Aluno/Disciplina* podemos ter 1 e somente 1 *Professor* relacionado;
 - Separar a entidade *Disciplina* e analisar *Professor, Aluno*. Para cada par *Professor/Aluno* podemos ter de 1 até N *Disciplinas* relacionadas.

Cardinalidade dos Relacionamentos

- N-ário
 - Quando um aluno está matriculado em uma disciplina, este tem sempre um professor;
 - Um aluno pode estar matriculado em várias disciplinas;
 - Uma disciplina tem vários alunos, e somente um professor;
 - Um professor leciona uma disciplina para vários alunos.

Decidir a Cardinalidade

- Modelagem de Dados é uma Arte
 - Nota Fiscal, Item de Nota Fiscal e Produto

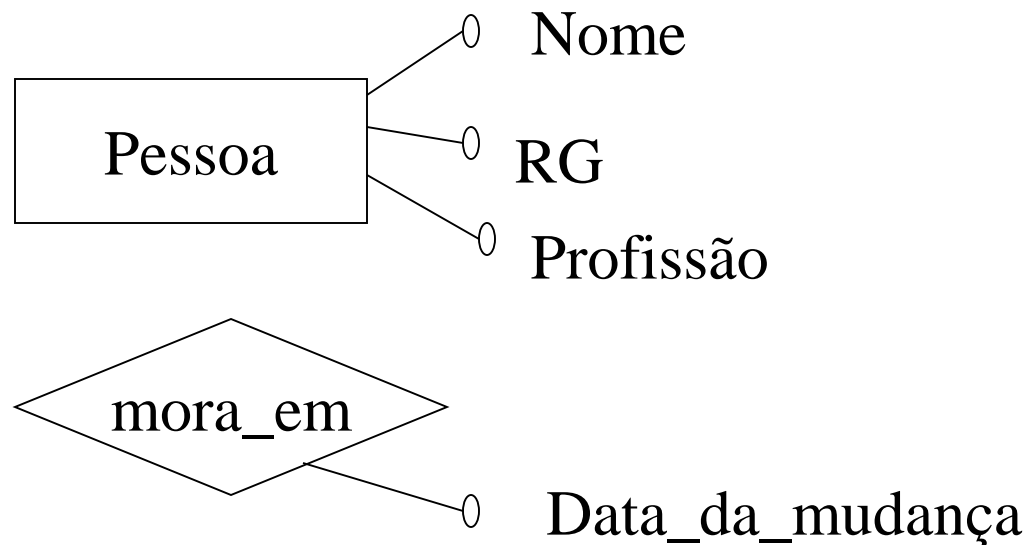


Decidir a Cardinalidade

- Modelagem de Dados é uma Arte
 - Todas as notas fiscais têm, no mínimo, um item de nota fiscal relacionado;
 - Todo item de nota fiscal está relacionado a uma nota fiscal;
 - Todo item de nota fiscal está relacionado a um produto.
- A entidade item de NF está relacionando NF e produto. Ela também é conhecida como Entidade Associativa.

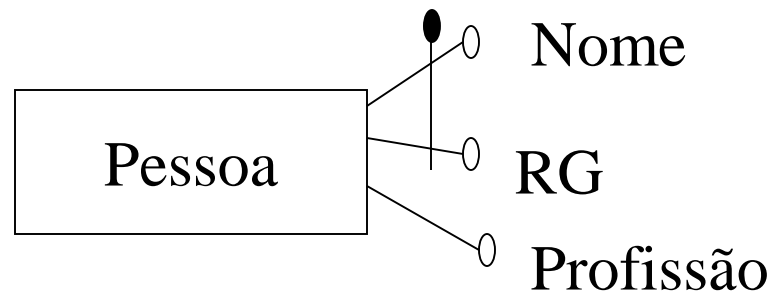
Atributos

- Atributos são propriedades elementares de entidades e relacionamentos



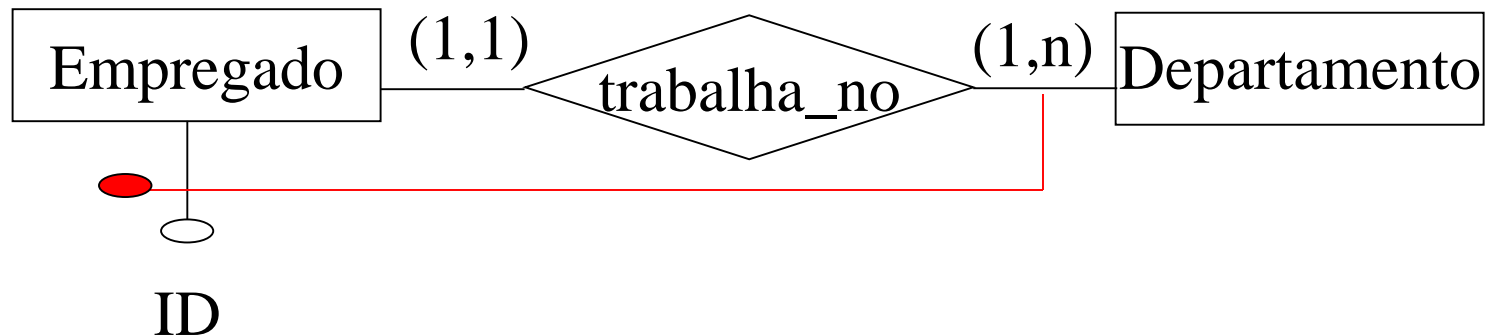
Identificadores (chaves)

- Identificadores (chaves) são conjuntos de atributos que podem determinar unicamente uma entidade



Identificadores (chaves)

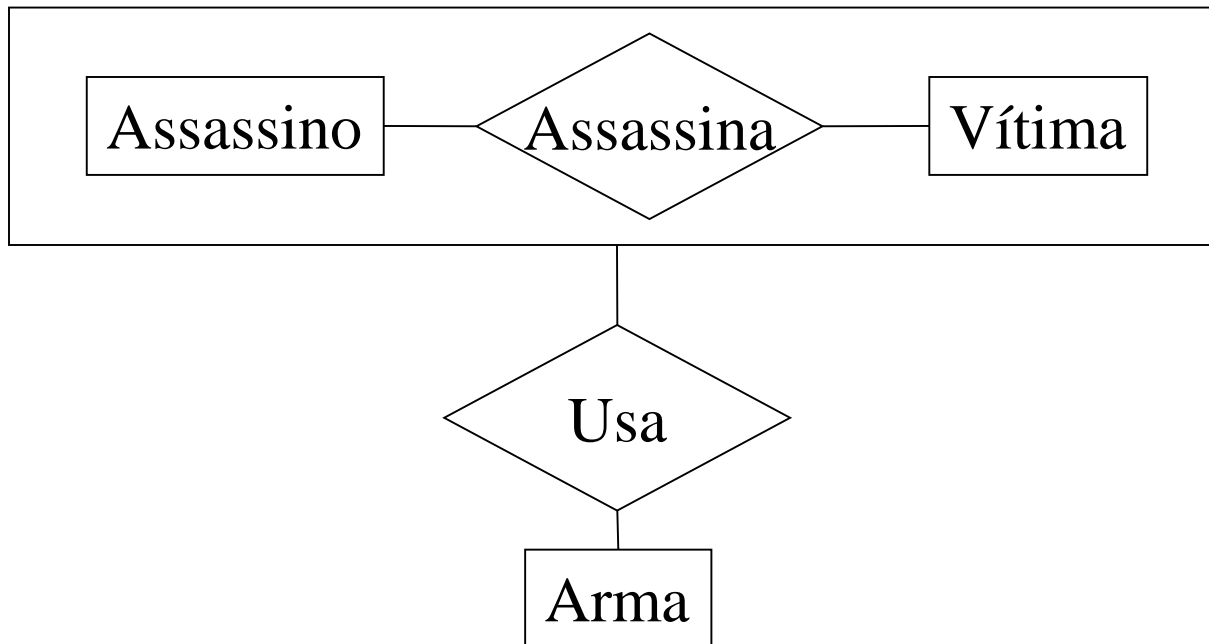
- Simples ou compostos
- Internos ou externos



Nota: Empregado é uma *entidade fraca*

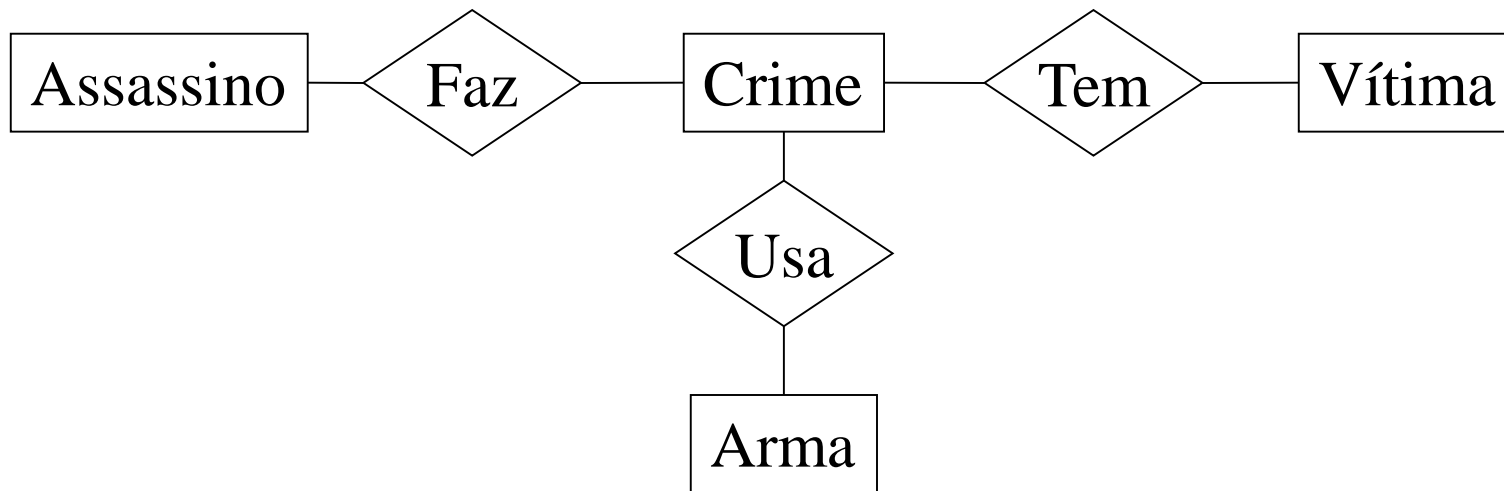
Agregação

- Um assassino faz vítimas. Quando ele assassina, utiliza armas.



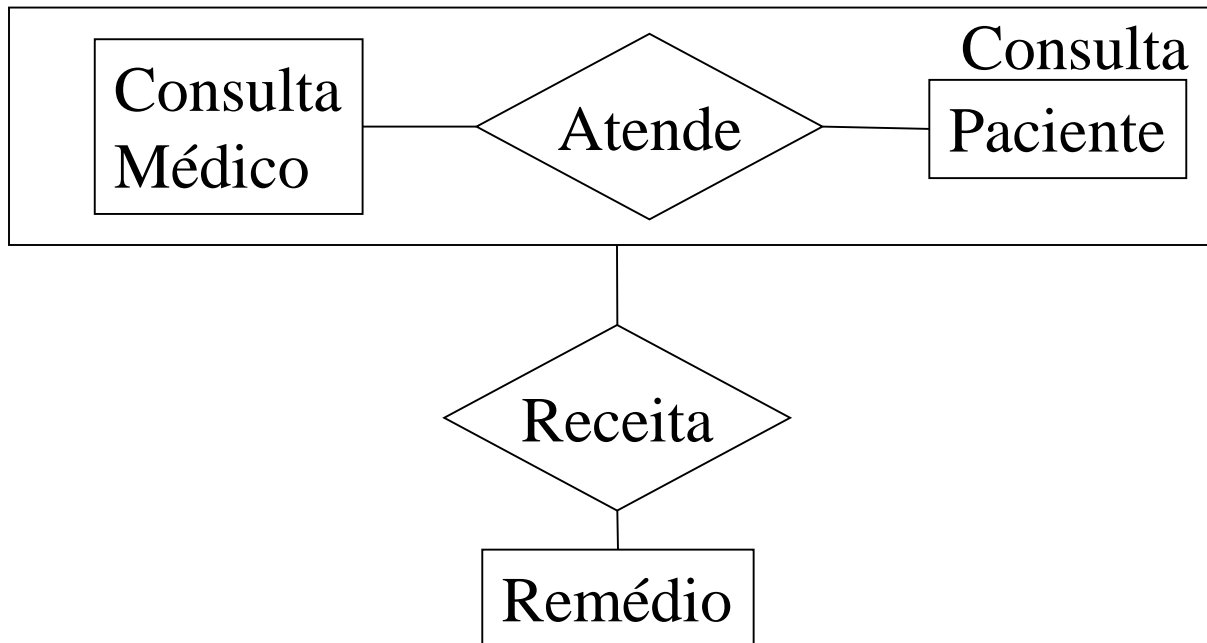
Agregação

- Existem na realidade dois relacionamentos para retratar um fato por completo.
- O que desejamos é relacionar uma ocorrência de Arma, com uma ocorrência do fato, do relacionamento Assassina.



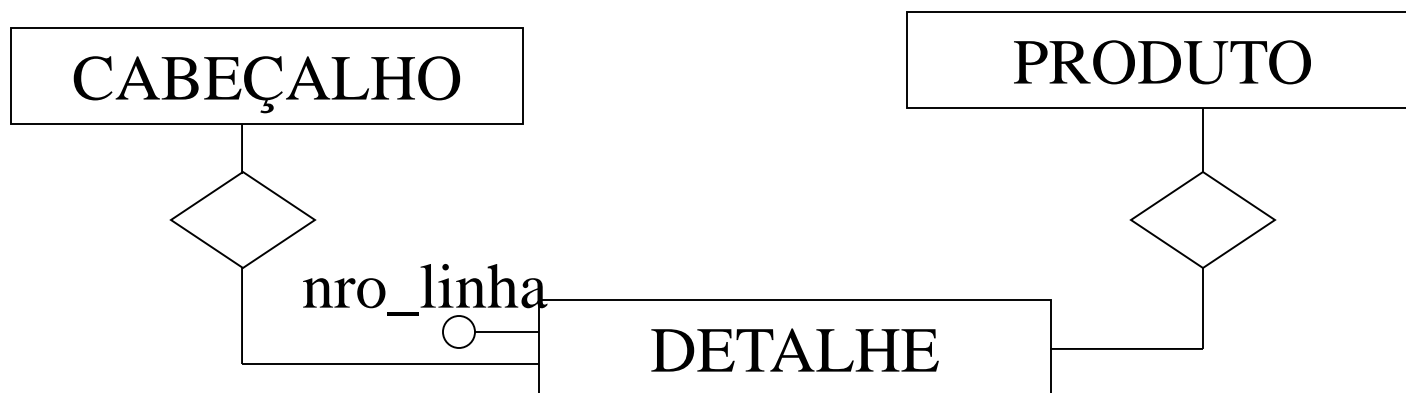
Agregação

- Um médico atende a muitos pacientes, que o consultam, e um paciente pode realizar consultas com muitos médicos. Sempre que um paciente consulta um médico, este fornece uma receita, que pode ter um, ou vários remédios.



Problemas 2.1

- Prepare um diagrama E-R para um sistema de controle de pedidos
- Mostre dois identificadores para a entidade DETALHE



Conceitos básicos

- **Relação:** tabela bi-dimensional
- **Atributos:** nomes das colunas da tabela
- **Schema:** nome da relação e atributos
 - Filme(Título, Ano, Duração, Classificação)
- **Tupla:** linhas de uma relação
 - ("Star Wars", 1977, 124, Ficção)
- **Domínio:** valores que um atributo pode assumir

Formas normais

- O conceito de normalização foi introduzido por E.F.Codd em 1970 (primeira forma normal). Esta técnica é um processo matemático que tem seus fundamentos na teoria dos conjuntos.
- Através deste processo pode-se, gradativamente, substituir um conjunto de entidades e relacionamentos por um outro, o qual se apresenta *purificado* em relação às anomalias de atualização (inclusão, alteração e exclusão) as quais podem causar certos problemas, tais como: grupos repetitivos (atributos multivalorados) de dados, dependências parciais em relação a uma chave concatenada, redundância desnecessária de dados, dificuldade na representação de fatos da realidade observada e dependências transitivas entre atributos.

Fomulário de Pedido

Num. Ped.	Cliente	Endereço	CGC	IE	Cod. Prod.	Unid.	Quant.	Descrição	Val. Unit.	Tot. Prod.	Tot. do Ped.
3445	TCA	R. Meira	1111111	111111	45	L	20	álcool	5,00	100,00	1799,00
3445	TCA	R. Meira	1111111	111111	130	M	2	tecido	20,00	40,00	1799,00
3445	TCA	R. Meira	1111111	111111	335	Kg	30	farinha	1,00	30,00	1799,00
3445	TCA	R. Meira	1111111	111111	78	Kg	50	cimento	30,00	1500,00	1799,00
3445	TCA	R. Meira	1111111	111111	90	L	40	cola	3,00	120,00	1799,00
3445	TCA	R. Meira	1111111	111111	39	Kg	3	chumbo	3,00	9,00	1799,00
2610	Lopes	R. 127	23232323	34343434	45	L	50	álcool	5,00	250,00	2650,00
2610	Lopes	R. 127	23232323	34343434	78	Kg	47	cimento	30,00	1410,00	2650,00
2610	Lopes	R. 127	23232323	34343434	21	Kg	20	pregos	5,00	100,00	2650,00
2610	Lopes	R. 127	23232323	34343434	98	L	15	tinta azul	25,00	375,00	2650,00
2610	Lopes	R. 127	23232323	34343434	90	L	15	cola	3,00	45,00	2650,00
2610	Lopes	R. 127	23232323	34343434	43	M	10	arame	3,00	30,00	2650,00
2610	Lopes	R. 127	23232323	34343434	25	F	10	algodão	2,00	20,00	2650,00
2610	Lopes	R. 127	23232323	34343434	65	L	5	querosene	8,00	40,00	2650,00
2610	Lopes	R. 127	23232323	34343434	51	M	20	fio elétrico	13,00	260,00	2650,00
2610	Lopes	R. 127	23232323	34343434	74	M	30	linha 10	4,00	120,00	2650,00

Fomulário de Pedido (Atributos)

- Número do pedido
- Prazo de entrega
- Cliente
- Endereço
- Cidade
- UF
- CGC
- Inscrição Estadual
- Código do produto
- Unidade do produto
- Quantidade do produto
- Descrição do produto
- Valor unitário do produto
- Valor total do produto
- Valor total do pedido
- Código do vendedor
- Nome do vendedor

Formas normais

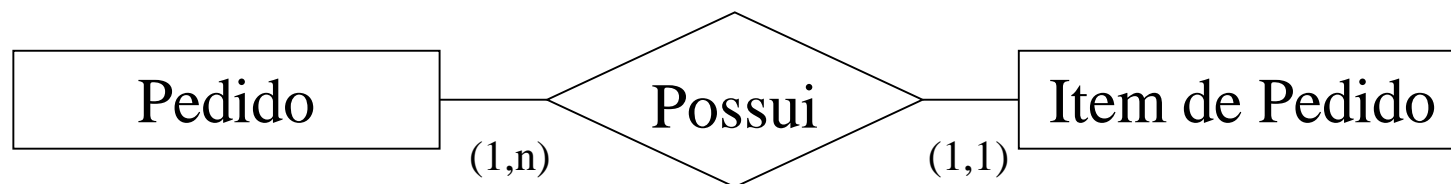
- Caso esta entidade fosse implementada como uma tabela em um banco de dados, as seguintes anomalias iriam aparecer:
 - Anomalia de inclusão: ao ser incluído um novo cliente, o mesmo tem que estar relacionado a uma venda;
 - Anomalia de exclusão: ao ser excluído um cliente, os dados referentes as suas compras serão perdidos;
 - Anomalia de alteração: caso algum fabricante de produto altere a faixa de preço de uma determinada classe de produtos, será preciso percorrer toda a entidade para se realizar múltiplas alterações.

Primeira Forma Normal (1FN)

- Em uma determinada realidade, às vezes encontramos algumas informações que se repetem (atributos multivalorados), retratando ocorrências de um mesmo fato dentro de uma única linha e vinculadas a sua chave primária;
- Ao observarmos a entidade PEDIDO, apresentada anteriormente, visualizamos que um certo grupo de atributos (produtos solicitados) se repete (número de ocorrências não definidas) ao longo do processo de entrada de dados na entidade;
- A 1FN diz que: cada ocorrência da chave primária deve corresponder a uma e somente uma informação de cada atributo, ou seja, a entidade não deve conter grupos repetitivos (multivalorados);
- Para se obter entidades na 1FN, é necessário decompor cada entidade não normalizada em tantas entidades quanto for o número de conjuntos de atributos repetitivos. Nas novas entidades criadas, a chave primária é a concatenação da chave primária original mais o(s) atributo(s) do grupo repetitivo visualizado(s) como chave primária deste grupo.

Primeira Forma Normal (1FN)

- Ao aplicarmos a 1FN sobre a entidade PEDIDO, obtemos mais uma entidade chamada ITEM-DE-PEDIDO, que herdará os atributos repetitivos e destacados da entidade PEDIDO;
- Um PEDIDO possui no mínimo 1 e no máximo N elementos em ITEM-DE-PEDIDO e um ITEM-DE-PEDIDO pertence a 1 e somente 1 PEDIDO.



Primeira Forma Normal (1FN)

- **PEDIDO**

- Número do pedido
- Prazo de entrega
- Cliente
- Endereço
- Cidade
- UF
- CGC
- Inscrição Estadual
- Valor total do pedido
- Valor total do pedido
- Código do vendedor
- Nome do vendedor

- **ITEM DE PEDIDO**

- Número do pedido
- Código do produto
- Quantidade do produto
- Descrição do produto
- Valor unitário do produto
- Valor total do produto
- Unidade do produto

Dependência Funcional

- Para descrevermos as próximas formas normais, se faz necessária a introdução do conceito de dependência funcional, sobre o qual a maior parte da teoria de normalização foi baseada;
- Dada uma entidade qualquer, dizemos que um atributo ou conjunto de atributos A é dependente funcional de um outro atributo B contido na mesma entidade, se a cada valor B existir nas linhas da entidade em que aparece, um único valor de A. Em outras palavras, A depende funcionalmente de B;
- Na entidade PEDIDO, o atributo PRAZO-DE-ENTREGA depende funcionalmente de NUMERO-DO-PEDIDO;
- O exame das relações existentes entre os atributos de uma entidade deve ser feito a partir do conhecimento (conceitual) que tem sobre a realidade a ser modelada.

Dependência Funcional Total e Parcial

- Na ocorrência de uma chave primária concatenada, dizemos que um atributo ou conjunto de atributos depende de forma completa ou total desta chave primária concatenada, se e somente se, a cada valor da chave (e não parte dela), está associado um valor para cada atributo, ou seja, um atributo não se apresenta com dependência completa ou total quando só depende de parte da chave primária concatenada e não dela como um todo;
- Dependência total – na entidade ITEM-DO-PEDIDO, o atributo QUANTIDADE-DO-PRODUTO depende de forma total ou completa da chave primária concatenada (NÚMERO-DO-PEDIDO + CÓDIGO-DO-PRODUTO);
- A dependência total ou completa só ocorre quando a chave primária for composta por vários atributos, ou seja, em uma entidade de chave primária composta de um único atributo não ocorre este tipo de dependência.

Dependência Funcional Transitiva

- Quando um atributo ou conjunto de atributos A depende de outro atributo B que não pertence à chave primária, mas é dependente funcional desta, dizemos que A é dependente transitivo de B.
- Dependência transitiva – na entidade PEDIDO, os atributos ENDEREÇO, CIDADE, UF, CGC e INSCRIÇÃO ESTADUAL são dependentes transitivos do atributo CLIENTE. Nesta mesma entidade, o atributo NOME-DO-VENDEDOR é dependente transitivo do atributo CÓDIGO-DO-VENDEDOR;
- Com base na teoria sobre as dependências funcionais entre atributos de uma entidade, podemos continuar com a apresentação das outras formas normais.

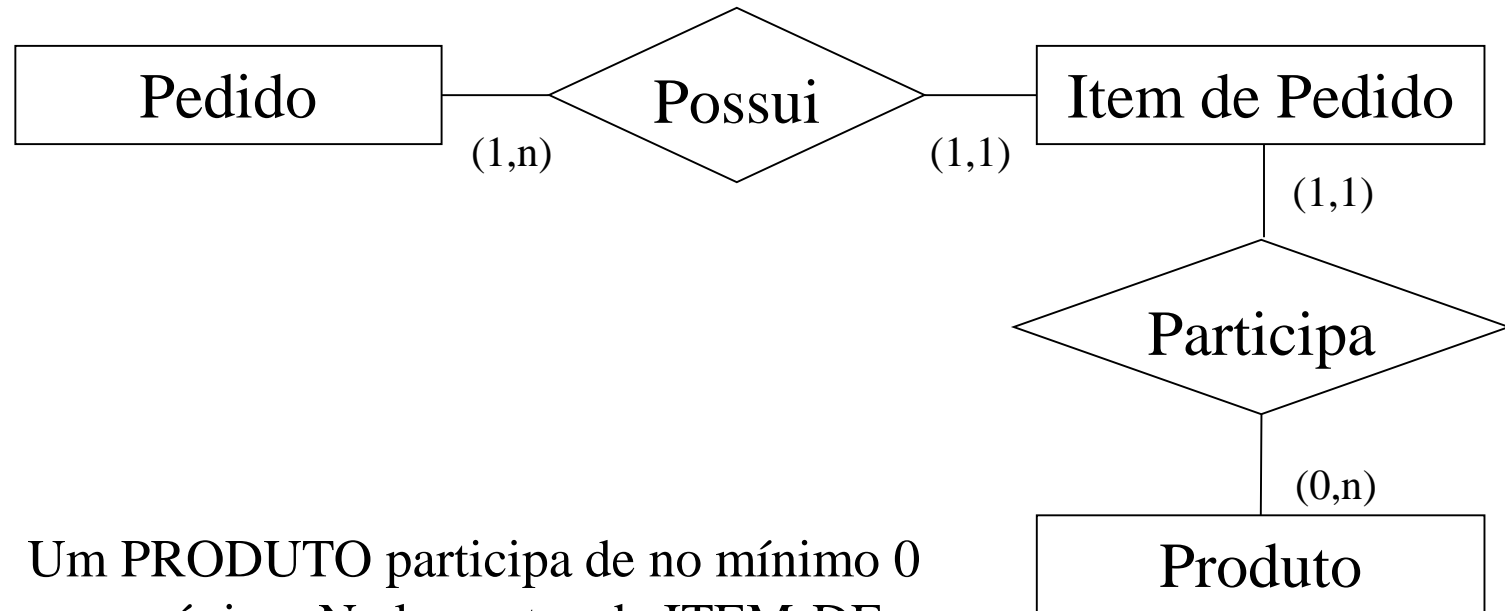
Segunda Forma Normal (2FN)

- Devemos observar se alguma entidade possui chave primária concatenada, e para aquelas que satisfizerem esta condição, analisar se existe algum atributo ou conjunto de atributos com dependência parcial em relação a algum elemento da chave primária concatenada;
- Com a finalidade de tornar ainda mais estável o modelo de dados, a aplicação da 2FN sobre as entidades em observação geram novas entidades, que herdarão a chave parcial e todos os atributos que dependem desta chave parcial, ou seja, uma entidade para estar na 2FN não pode ter atributos com dependência parcial em relação à chave primária.

Segunda Forma Normal (2FN)

- A entidade ITEM-DO-PEDIDO apresenta uma chave primária concatenada e por observação, notamos que os atributos UNIDADE-DO-PRODUTO, DESCRIÇÃO-DO-PRODUTO e VALOR-UNITÁRIO dependem de forma parcial do atributo CÓDIGO-DO-PRODUTO, que faz parte da chave primária. Logo devemos aplicar a 2FN sobre esta entidade. Quando aplicarmos a 2FN sobre ITEM-DO-PRODUTO, será criada a entidade PRODUTO que herdará os atributos UNIDADE-DO-PRODUTO, DESCRIÇÃO-DO-PRODUTO e VALOR-UNITÁRIO e terá como chave primária o CÓDIGO-DO-PRODUTO.

Segunda Forma Normal (2FN)



- Um PRODUTO participa de no mínimo 0 e no máximo N elementos de ITEM-DE-PEDIDO e um ITEM-DE-PEDIDO só pode conter 1 somente 1 PRODUTO.

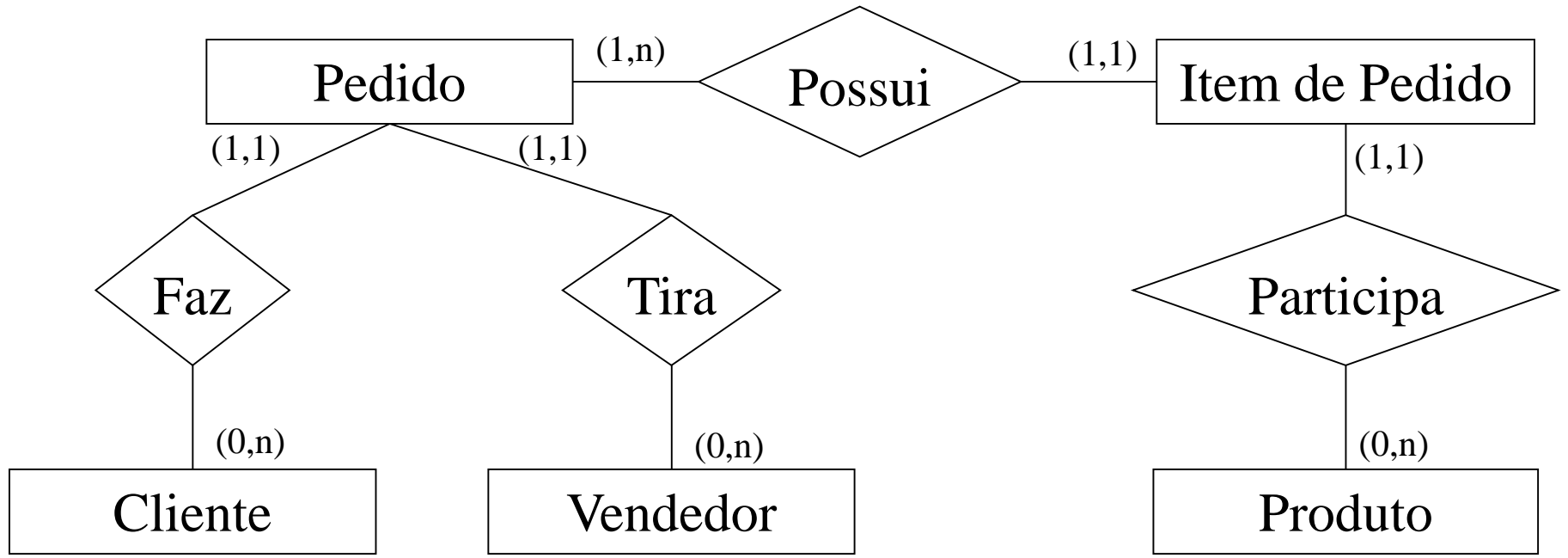
Terceira Forma Normal (3FN)

- Uma entidade está na 3FN se nenhum de seus atributos possui dependência transitiva em relação a outro atributo da entidade que não participe da chave primária, ou seja, não exista nenhum atributo intermediário entre a chave primária e o próprio atributo observado;
- Ao retirarmos a dependência transitiva, devemos criar uma nova entidade que contenha os atributos que dependem transitivamente de outro e a sua chave primária é o atributo que causou esta dependência;
- Além de não conter atributos com dependência transitiva, entidades na 3FN não devem conter atributos que sejam o resultado de algum cálculo sobre outro atributo, que de certa forma pode ser encarada como uma dependência funcional.

Terceira Forma Normal (3FN)

- Na entidade PEDIDO, podemos observar que o atributo NOME-DO-VENDEDOR depende transitivamente do atributo CÓDIGO-DO-VENDEDOR que não pertence à chave primária. Para eliminarmos esta anomalia devemos criar a entidade VENDEDOR, com o atributo NOME-DO-VENDEDOR e tendo como chave primária o atributo CÓDIGO-DO-VENDEDOR;
- Encontramos ainda o conjunto de atributos formados por ENDEREÇO, CIDADE, UF, CGC e INSCRIÇÃO ESTADUAL que dependem transitivamente do atributo CLIENTE. Neste caso, devemos criar a entidade CLIENTE que conterà os atributos ENDEREÇO, CIDADE, UF, CGC e INSCRIÇÃO ESTADUAL. Para chave primária desta entidade vamos criar um atributo chamado CÓDIGO-DO-CLIENTE que funcionará melhor como chave primária do que NOME-DO-CLIENTE, deixando este último como simples atributo da entidade CLIENTE.

Terceira Forma Normal (3FN)



Terceira Forma Normal (3FN)

- Um PEDIDO só é feito por um e somente um CLIENTE e um CLIENTE pode fazer de 0 até N elementos de PEDIDO. Um PEDIDO só é tirado por um e somente um VENDEDOR e VENDEDOR pode tirar de 0 a N elementos de PEDIDO.

Seleção

Produz uma nova relação R' com um subconjunto das tuplas originais

$\sigma_{\text{Título,ano}=\text{“Star Wars”}}$ (Filmes)

selecionaria os atributos Título e Ano da relação Filmes cujos títulos são “Star Wars”.

Operações de conjunto

R, S sendo relações, tem-se

- $R \cup S$: união
- $R \cap S$: intersecção
- $R - S$: diferença

aplicam-se sobre os elementos de R e S

Projeção

Seleção de um ou mais atributos de uma relação

$\Pi_{\text{Título, ano}}(\text{Filmes})$

selecionaria os atributos Título e Ano da relação
Filmes.

Produto Cartesiano

O produto $R \times S$ é o conjunto de pares formados tomando-se o primeiro elemento do par como sendo um elemento qualquer de R e o segundo, um elemento de S

A	B		B	C	D
1	2		2	5	6
3	4		4	7	8
			9	10	11
A	R.B	S.B	C	D	
1	2	2	5	6	
1	2	4	7	8	
1	2	9	10	11	
3	4	2	5	6	
3	4	4	7	8	
3	4	9	10	11	

Join

O join natural $R \bowtie S$ é o produto das duas relações, relacionando somente as tuplas que correspondem em um dado sentido

A	B	C		B	C	D
1	2	3		2	3	4
6	7	8		2	3	5
9	7	8		7	8	10
A	B	C	D			
1	2	3	4			
1	2	3	5			
6	7	8	10			
9	7	8	10			

Problemas 2.3

- a. Para o modelo relacional do sistema de pedidos, calcule as operações para alguma(s) das relações.

Produto Cartesiano

O produto $R \times S$ é o conjunto de pares formados tomando-se o primeiro elemento do par como sendo um elemento qualquer de R e o segundo, um elemento de S

Problemas 2.2

- a. Converta o modelo E-R para o relacional
- b. Analise o modelo relacional obtido em a. em relação às formas normais