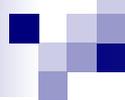


# Modelagem Semântica

Ademar Feil

Alessandro Quirolli

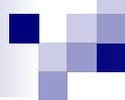


# Introdução

- Banco de Dados Tradicionais x Modelagem Semântica
- Uso da Modelagem Semântica
- Formalismo na Modelagem
- Complexidade

# Conceitos Fundamentais

<b>Conceito</b>	<b>Definição Informal</b>	<b>Exemplos</b>
ENTIDADE	Um objeto passível de distinção (de um tipo particular)	Fornecedor, Peça, Pessoa
PROPRIEDADE	Uma peça de informação que descreve uma entidade	Número de fornecedor, Altura da pessoa
ASSOCIAÇÃO	Um relacionamento de muitos-para-muitos (um-para-muitos etc.) entre entidades	Expedição (peça do fornecedor), Gravação (composição-orquestra regente)
SUBTIPO	O tipo de entidade Y é um subtipo da entidade de tipo X se, e apenas se, cada Y for necessariamente um X.	Funcionário é um subtipo de pessoa, Concerto é um subtipo de Composição.

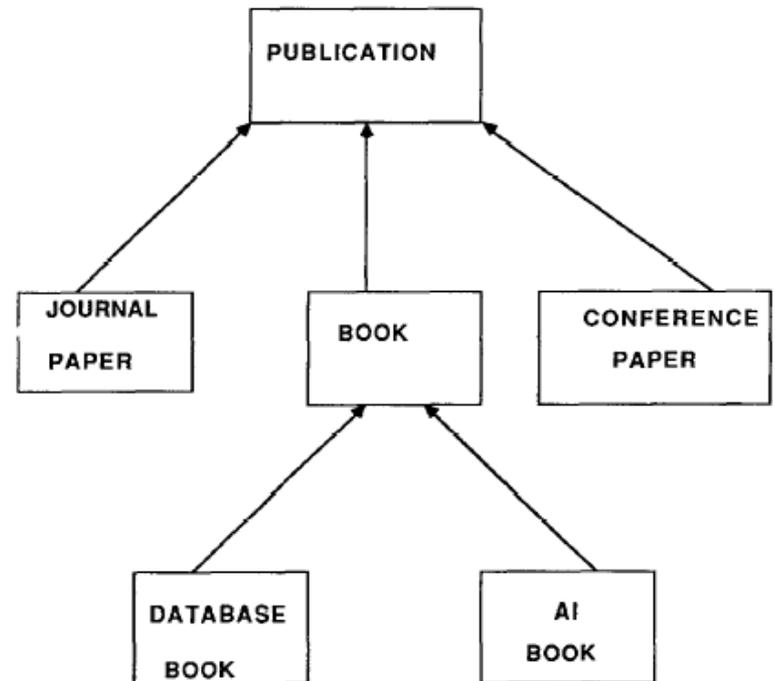


# MODELAGEM SEMÂNTICA E ORIENTAÇÃO A OBJETOS

- Grau de Abstração
- Independência dos Dados
- Captura Semântica dos Dados
- Generalização e Agregação
- Classificação e Associação

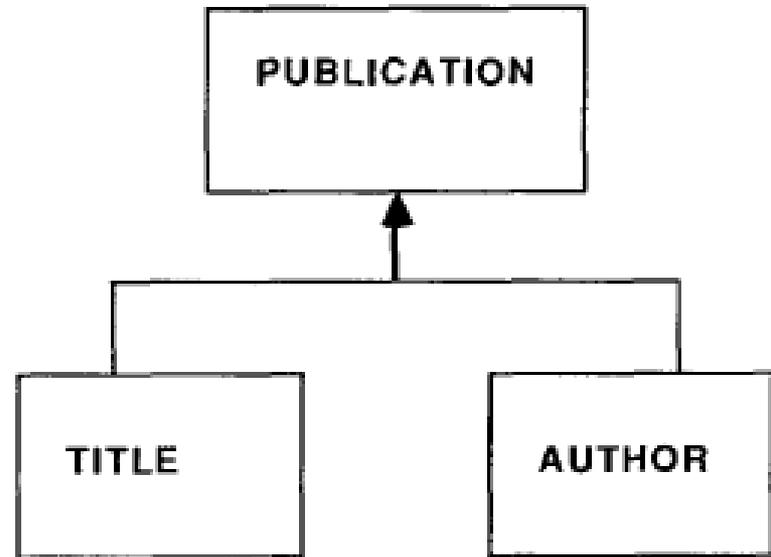
# Generalização

- Generalização é a forma que, diferenças entre objetos similares são ignoradas para formar um tipo em que a ênfase é dada às similaridades.



# Agregação

- A Agregação é o relacionamento entre tipos de baixo-nível que pode ser considerado um tipo de alto-nível.



# Classificação

- Classificação é uma forma de abstração em que uma coleção de objetos é considerada o nível alto da classe do objeto. Essencialmente, isto representa um relacionamento *is-instance-of* (é instância de).

# Associação

- Associação é uma forma de abstração em que um relacionamento entre objetos é considerado um conjunto de objetos de alto nível [Brodie 1984]. O conjunto DATABASE\_BOOKS é uma associação de objetos BOOK, assim como o conjunto AI\_BOOKS

# Separação das Estruturas Físicas e Lógicas

- Em modelos orientados a registro o acesso feito por usuários finais tendem a imitar a estrutura lógica do esquema de banco de dados.
- No modelo relacional um usuário deve simular ponteiros comparando identificadores de uma relação para outra (como em uma operação *join*).
- Os atributos dos modelos semânticos podem ser usados diretamente como ponteiros. Assim, usuários devem passar através de um nível extra imposto pelo modelo relacional, fazendo com que seja mais difícil formar objetos complexos de simples objetos.

# Mecanismos de Abstração

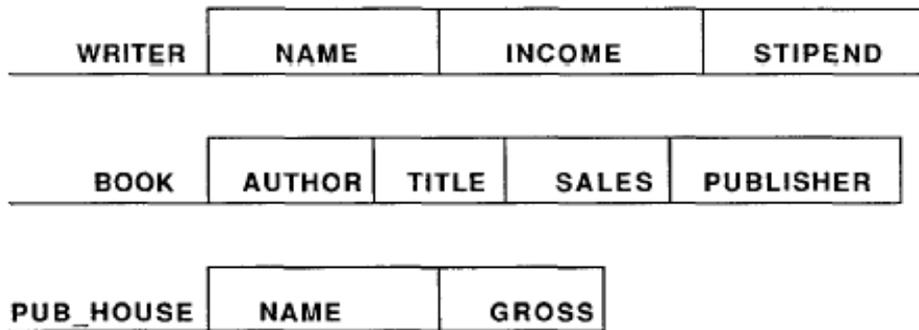
- A primeira dimensão de abstração consiste em um nível de detalhe em que porções do esquema podem ser vistos.
- A segunda dimensão de abstração nos modelos semânticos é o grau de modularidade fornecido.
- Uma terceira dimensão de abstração é conseguida derivando componentes do esquema que são suportados por alguns modelos semânticos e também por algumas implementações relacionais.

# VANTAGENS DA MODELAGEM SEMÂNTICA

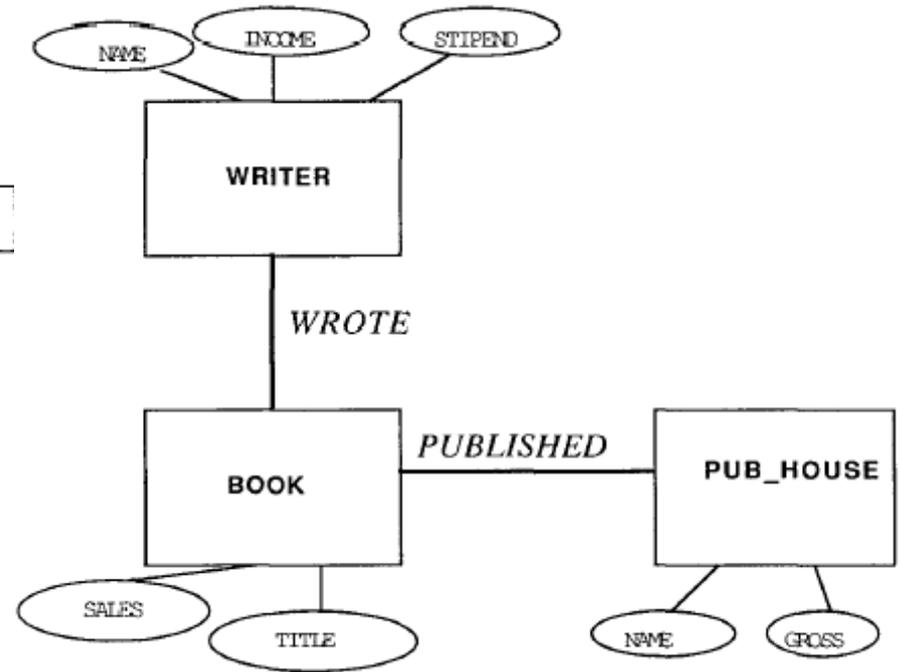
## ■ Economia de Expressão

- Os modelos de dados semânticos são geralmente completos no sentido de que o usuário é capaz de extrair qualquer informação da base de dados facilmente.
- No modelo semântico, operações são definidas explicitamente sobre os relacionamentos.

# Economia de Expressão



Esquema relacional da base de dados livraria



Esquema semântico da base de dados livraria

# Economia de Expressão

## Notação No Esquema relacional

```
SELECT NAME FROM WRITER
WHERE WRITER.INCOME < 2000
AND WRITER.NAME = BOOK.AUTHOR
AND BOOK.SALES > 50000
AND BOOK.PUBLISHER = PUBLISHER.NAME
AND PUB_HOUSE.GROSS > 1000000
```

## Notação de Tsichritzis e Lochovsky [1982]

```
SELECT NAME FROM WRITER
WHERE WRITER.INCOME < 2000
AND BOOK.SALES > 50000/WROTE
AND PUB_HOUSE.GROSS > 1000000/PUBLISHED
```

# VANTAGENS DA MODELAGEM SEMÂNTICA

- **Manutenção da Integridade:** Os modelos tradicionais forçam o usuário a manter conexões entre os objetos da base de dados ou a manter a consistência entre os objetos.
- **Flexibilidade de Modelagem:** Os modelos de dados mais tradicionais fornecem somente uma maneira de representar os dados. Modelos de dados semânticos, através de uso de abstrações, permitem que o usuário modele e visualize os dados de vários níveis diferentes.

# VANTAGENS DA MODELAGEM SEMÂNTICA

- **Eficiência na Modelagem:** Em um projeto de banco de dados, o projetista não precisa implementar instruções de baixo nível. A maioria dos modelos semânticos contém operações elementares já explícitas.

# **BASES DE COMPARAÇÃO**

- Tsichritzis e Lochovsky [1982] classificou os modelos como tradicional, entidade-relacionamento, binário e rede semântica.
- Hull e King [1987] classificou de forma diferente: um modelo construído com os componentes fundamentais da modelagem semântica é usado para comparação com outro modelo.

# Características

- **Representação de Objetos Não-Estruturados**
- **Representação de Relacionamentos**
- **Existência de Abstração Comum**
- **Rede ou Hierarquia de Relacionamentos**
- **Derivação / Herança**
- **Inserção / Exclusão / Modificação**

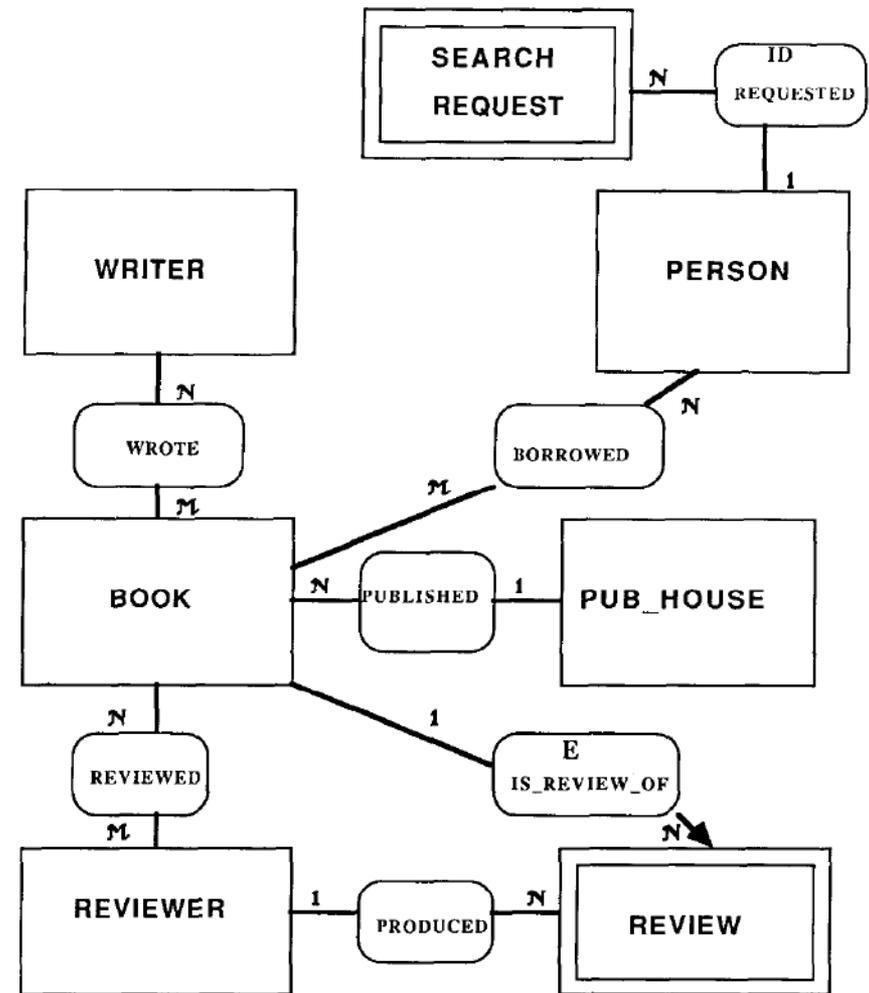
# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

## ■ Modelo de Entidade-Relacionamento

- Chen[1976]
- Entidade é “uma coisa que pode ser identificada distintamente”
- Relacionamento como “uma associação entre entidades” (uma-para-uma ou muitas-para-uma ou muitas-para-muitas).
- Um relacionamento é algo que pode ser “identificado”.

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

Na figura 6, temos os elementos 1, N e M sendo a cardinalidade dos relacionamentos, onde os elementos M e N denotam muitos. Por exemplo, a entidade Review “Revisões”, depende da entidade book “livro” para existir, desta forma se um elemento de book for excluído, todas as review relativas serão excluídas. A dependência de existência é demarcada pelo quadro duplo na entidade review e colocando um comentário “E” entre as entidades.



# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

## ■ TAXIS

- TAXIS [Borgida et al 1984, Mylopoulos et al 1980, Nixon et al 1987 e O'Brien 1983] é uma linguagem para modelagem de sistema de banco de dados interativos que enfatiza na classificação e hierarquia de generalização/especialização.
- Este modelo combina a idéia de linguagem de programação e teoria de banco de dados.

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

- TAXIS é um sistema onde a hierarquia de classificação e generalização são enfatizadas e estendidas ao banco de dados.
- Para descrever as classes, são utilizados os seguintes grupos de categorias:
  - Chaves – identificar uma instância
  - Características – agrupam as propriedades que não se alteram
  - Atributos – agrupam as propriedades que se alteram
- Através da pré-definição de algumas classes é possível gerar novas classes especializadas.

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

- TAXIS é um modelo fortemente baseado na abstração de herança (IS-A) utilizando-a para estruturar dados e procedimentos de uma aplicação, incluindo as expressões, transações e exceções.

```
DATACLASS BOOK WITH
  ATTRIBUTES
    AUTHOR: SET OF WRITER,
    SALES: 0...999999999;
    TITLE: STRING;
    PUBLISHER: PUB-HOUSE;
    BORROWER: SET OF PERSON;
    REVIEWERS: SET OF REVIEWER;
    REVIEWS: SET OF REVIEW;
end BOOK;
```

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

## ■ SDM

- SDM [Hammer and McLeod 1981] traz uma aproximação diferenciando do modelo E/R e do TAXIS, incorporando um grande número de construtores de modelagem em uma abstração simples, **a classe**. Busca permitir ao designer de banco de dados expressar o banco de dados claramente com mecanismos para mapear diretamente o conceito de design do programador.
- Classificação e associação possuem uma maior ênfase no SDM do que agregação e generalização.

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

- Um esquema SDM é uma coleção de entidades organizadas em classes e conexões entre as classes e atributos derivados.
- Nas classes são especificados os atributos dos membros e das classes.
- Existem dois tipos de conexões entre as classes, uma que representa o mecanismo de agrupamento e outro que representa o mecanismo de generalização/ especialização.

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

DESCRIPTION: ALL BOOKS WITHIN THE  
LIBRARY MEMBER  
ATTRIBUTES:  
TITLE  
VALUE CLASS: STRINGS  
AUTHOR  
VALUE CLASS: PERSON  
PUBLISHER  
VALUE CLASS: PUB-HOUSE  
PRICE  
DESCRIPTION: PRICE OF BOOK AT  
LIBRARY RATE  
VALUE CLASS: DOLLARS  
TOPIC  
VALUE CLASS: TOPICS  
INTEREST-INDEX  
VALUE CLASS: INTEGERS  
ACQUISITION-PRIORITY  
VALUE CLASS: INTEGERS  
DERIVATION: ORDER BY DECREASING  
INTEREST-  
INDEX WITHIN BOOK.  
IDENTIFIERS  
TITLE

**DATABASE-BOOK**  
DESCRIPTION: ALL DATABASE BOOKS  
WITHIN THE LIBRARY INTERCLASS  
CONNECTION: SUBCLASS OF BOOK  
WHERE  
BOOK.TOPIC = 'DATABASE'  
MEMBER ATTRIBUTES:  
SUBTOPIC  
VALUE CLASS: DB-SUB-TOPICS  
CLASS ATTRIBUTES:  
TOTAL-COST  
VALUE CLASS: DOLLARS  
DERIVATION: SUM OF PRICE  
MEMBERS OF THIS CLASS

OVER

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

## ■ FUNCTIONAL DATA MODEL

- O Function Data Model [Shipman 1981] foi construído em conjunto com a linguagem de definição de dados DAPLEX. O objetivo era prover um modelo e uma linguagem de definição/manipulação que fosse capaz de representar aplicativos com naturalidade e simplicidade.

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

- Neste modelo é possível definir funções para especificar agregações de atributos utilizados para formar a entidade, como no exemplo abaixo, onde o símbolo ==>> indica elementos multivalorados e a declaração Author(Book) representa a relação entre book e person:

```
DECLARE BOOK( )==>> ENTITY
DECLARE TITLE (BOOK) ==> STRING
DECLARE PUBLISHER (BOOK) ==> PUB-HOUSE
DECLARE AUTHOR (BOOK) ==>> PERSON
```

- O modelo não traz um meio explícito para generalização e classificação, sendo necessário o usuário definir funções para obtê-las.

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

## ■ RM/T

- Foi primeiramente definido por Codd.
- Uma diferença imediata entre o RM/T e o modelo E/R é que o RM/T não faz distinções desnecessárias entre entidades e relacionamentos, um relacionamento é meramente considerado um tipo especial de entidade.
- A segunda diferença é que os aspectos estruturais e de integridade do modelo são mais ampliados e definidos de forma muito mais precisa no RM/T.
- A terceira diferença é que o RM/T inclui os seus próprios operadores especiais, além dos operadores do modelo relacional básico.

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

- **Relações E** - registram a existência das entidades
- **Relações P** - registram certas propriedades dessas entidades
- O RM/T inclui uma estrutura de catálogo formal através do qual o sistema toma conhecimento dos relacionamentos (restrições de integridade) e existem vários operadores de alto nível para facilitarem a manipulação dos vários objetos RM/T (relações E, relações P, relações de catálogo, etc.).
- O modelo RM/T divide as entidades em:
  - **Entidades-semente** (possuem existência independente),
  - **Entidades-característica** (uma entidade cuja única função é de descrever uma outra entidade)
  - **Entidade-associativa** (uma entidade cuja função é representar o relacionamento de entidades).

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

- Exemplo de E-Relation e P-Relation de Book onde a E-Relation “Relações E” representa a existência de uma entidade e as P-Relation “Relações P” que representa as propriedades da entidade Book.

## E-RELATION

BOOK	BOOK_ID
------	---------

## P-RELATIONS

AUTH	BOOK_ID	AUTHOR
------	---------	--------

TITL	BOOK_ID	TITLE
------	---------	-------

PUBL	BOOK_ID	PUBLISHER
------	---------	-----------

# MODELOS DE REPRESENTAÇÃO SEMÂNTICA

- Um modelo como o RM/T pode ser útil como auxílio ao projeto de banco de dados, contudo sua complexidade ultrapassa em muito a complexidade do modelo relacional. Parte das extensões semânticas do RM/T são feitas no dicionário de dados do modelo relacional, através de relações que descrevem os inter-relacionamentos existentes junto a novos operadores.

# Comparação entre Modelos

	<b>Representação de Objetos Desestruturado</b>	<b>Representação de Relacionamentos</b>	<b>Abstração Standard</b>	<b>Rede vs Hierarquia</b>	<b>Derivação/herança</b>	<b>Inserção/exclusão</b>	<b>Semântica de Relacionamentos</b>
<b>E/R</b>	Limitado	Independente e tabelas	Agregação	Forte Rede	Não	Específico do Usuário	Selecionada pelo usuário
<b>TAXIS</b>	Limitado	Entidade (Classes)	Generalização Agregação Classificação	Forte Hierarquia	Herança	Específica do Usuário exeto na hierarquia IS-A	Predefinida
<b>SDM</b>	Limitado	Independente e entidades (classes)	Generalização Agregação Classificação	Presença de uma hierarquia geral	Elaborada e Variada	Automática	Definida pelo Usuário
<b>Funcional</b>	Limitado	Funções	Associação Agregação Classificação	Sem suporte para ambos	Funcional	Específica do usuário	Definida pelo Usuário
<b>RM/T</b>	Limitado	Independente	Generalização Agregação Classificação	Presença de uma hierarquia geral	Herança	Automática e específica do usuário	Predefinida

# Conclusão

- Recursos e Facilidades
- Esquema de Banco de Dados mais legível
- Abstração
- Generalização, Especialização, Agregação e Herança

# BIBLIOGRAFIA

- PECKHAM, JOAN; MARYANSKI, FRED. ***Semantic Data Models***. ACM Computing Surveys (CSUR). Volume 20 , Issue 3 (September 1988).
- HULL, RICHARD; KING, ROGER. ***Semantic Database Modeling: Survey, Applications, and Research Issues***. ACM Computing Surveys (CSUR). Volume 19 , Issue 3 (September 1987).
- DATE, C. J.. ***Introdução a Sistemas de Bancos de Dados***. Tradução de Daniel Vieira. Rio de Janeiro: Elsevier, 2003.
- DATE, C. J.. ***Introdução a Sistemas de Bancos de Dados***. Tradução da 4ª Edição. Rio de Janeiro: Campus, 1990.