

UFSC-CTC-INE

Curso de Sistemas de Informação

INE 5600 Bancos de Dados III

Gerência de Dados XML em Bancos de Dados

Gerência de Dados XML em BDs

SGBDs relacionais estendidos

- tipo de dado XML
- linguagem *SQL/XML* (extensão da SQL)
- falta de padronização na atualização de dados

SGBDs XML nativos

- modelo de armazenamento proprietário
- indexação de dados XML
- suporte à tecnologia XML da W3C
DTD, XSD, XPath, XQuery, DOM
- atualização de dados
linguagem *XUpdate* ou *XQuery* estendida
- limitações na gerência de dados
controle de RIs, visões, transações, ...

SGBDs Relacionais Estendidos

Tipo de dado XML

- associado a uma coluna de uma tabela
 - mantém documento XML ou conteúdo XML bem-formado
 - armazenamento do tipo CLOB
 - pode sofrer operações SQL/XML
- Exemplo

`xml`

Inserção de Dados do Tipo XML

Linguagem SQL/XML

Funções para manipulação de dados XML

- navegação e filtro sobre dados do tipo XML
suporte à expressões *XPath* e *XQuery*
- geração de conteúdo XML a partir de dados relacionais ou dados do tipo XML

Principais funções

- (gera um elemento XML)
- (gera um atributo XML)
- (gera uma sequência de elementos XML)
- (concatena conteúdo de uma coluna do tipo XML)
- (concatena 2 conteúdos XML)
- (predicado para expressões XPath)
- (executa consultas XPath) (SGBD Postgres)

SQL/XML - Exemplos

Pessoas

<u>ID</u>	CPF	nome	sexo
1	111111111111	Joao da Silva	M
2	222222222222	Ana Souza	F
3	333333333333	Pedro Santos	M
4	444444444444	Carlos Pereira	M
5	555555555555	Maria Oliveira	F

xmlelement

xmlattributes

Resultado: "<livro isbn="1"/>"

xmlelement

xmlattributes

xmlelement

xmlelement

Resultado: "<pessoa cpf="11111111111">
<nome>Joao da Silva</nome> <sexo>M </sexo></pessoa>"

SQL/XML - Exemplos

`xmlconcat`

`xmlforest`

Resultado: "<resultado/><cpf>2222222222</cpf><nome>Ana Souza</nome><sexo>F</sexo>"

`xmlagg`

Resultado: "<livro ISBN='1'><nome>Java</nome></livro>
<livro ISBN='2'><nome>XML</nome></livro>"

SQL/XML - Exemplos

`xpath`

Resultado: "<nome>Java</nome>
<nome>XML</nome>
<nome>SQL</nome>"

`xmlexists`

Resultado: 1

Atualização de Dados XML

Falta de padronização em termos de capacidades de atualização

- exemplos

Postgres: não possui operações de atualização

Oracle e DB2: funções proprietárias

Exemplo de atualização no Oracle

Livros ← tabela do tipo XML (XMLType)
← constante que indica conteúdo XML

updateXML

SGBDs XML Nativos

Armazenamento nativo de dados XML

Principais estratégias de armazenamento

Árvore DOM	BDOO
preserva ordem de elementos	não preserva ordem de elementos
qualquer esquema tem a mesma estrutura (<i>document</i> , <i>element</i> , ...)	esquema de classes gerado de acordo com os tipos de elementos complexos
clusterização por profundidade bom p/ buscas na ordem da hierarquia clusterização por largura bom p/ buscas por propriedades de um elemento	clusterização geralmente por instâncias da mesma classe bom para buscas por dados de determinados tipos de elementos

Tipos de Indexação

Indexação *por valor*

- indexa valores de elementos simples e atributos

exemplo

Indexação *por estrutura*

- indexa nomes de elementos ou de atributos

exemplo

Indexação *fulltext*

- indexa *tokens* presentes em nomes ou em valores de elementos ou de atributos

exemplo: buscar elementos que contenham o termo

Consulta a Dados XML

Uso das linguagens *XPath* e/ou *XQuery*

Alguns SGBDs XML com capacidades adicionais

- exemplo: SGBD Tamino (busca por padrão)

`~= '*XML*'`

- exemplo: SGBD eXists (busca por palavra-chave)

`&= 'banco XML'`

Atualização de Dados XML

Principais abordagens

- linguagem *XQuery* estendida
- linguagem *XUpdate*

XQuery estendida

- exemplo: SGBD Tamino
update

replace

- exemplo: SGBD BaseX
insert node
into

Atualização de Dados XML

XUpdate

- Sintaxe XML

inclusão/elementos de elementos, atributos e texto
alteração do conteúdo de elementos e atributos

- Exemplos

```
xupdate:append
```

```
/xupdate:append
```

```
xupdate:remove
```

Projeto de BDs XML

- Não há uma metodologia consolidada
- Etapas do projeto tradicional de um BD
 - (i) especificação de requisitos; (ii) modelagem conceitual; (iii) modelagem lógica e (iv) modelagem física ou implementação
 - podem ser aplicadas a um BD XML
 - no

Guia para Projeto de BDs XML

1. Especificação de requisitos

levantamento das necessidades de dados

2. Modelagem conceitual

uso de algum modelo clássico (exemplo: ER)

3. Modelagem lógica

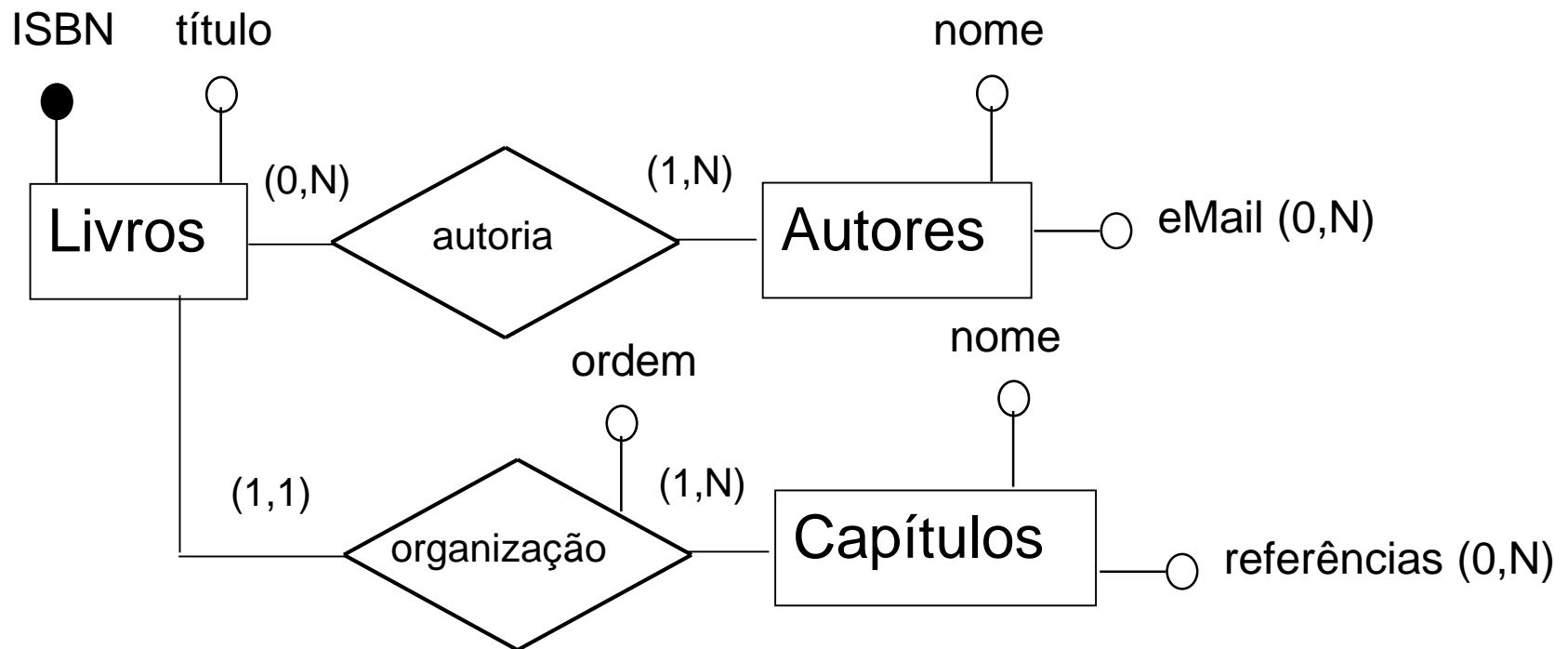
uso de algum modelo baseado em grafo direcionado

- adequado à representação de uma hierarquia XML

4. Modelagem física

especificação do esquema XML (DTD ou XSD)

Modelagem Conceitual - Exemplo



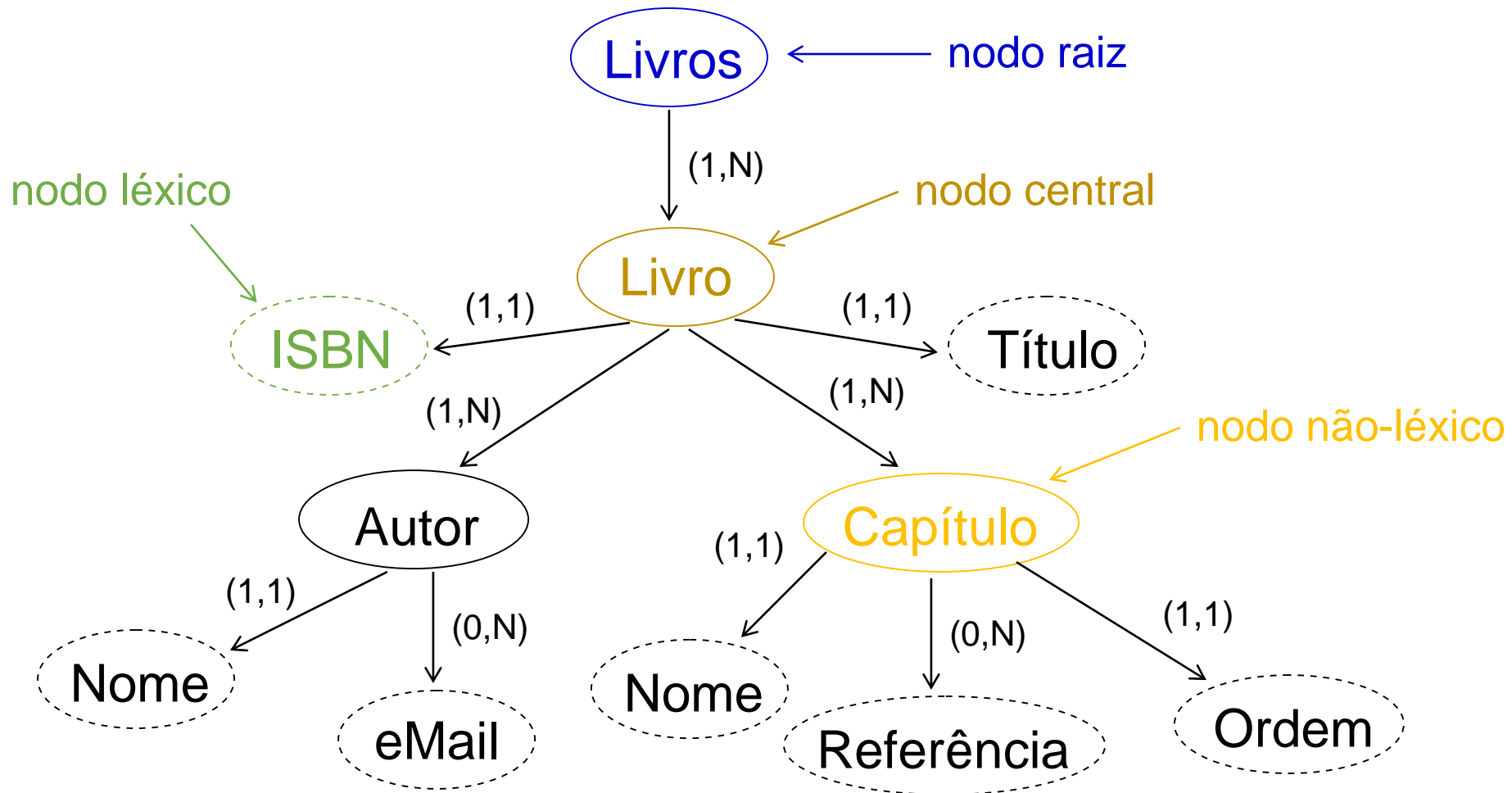
Modelagem Lógica

- Grafo direcionado com 2 tipos de nodos
- **Nodos não-terminais (ou não-léxicos)**
 - mapeamento de entidades do ER
 - modelam elementos compostos do XML
- **Nodos terminais (ou léxicos)**
 - mapeamento de atributos do ER
 - modelam conteúdo de elemento ou de atributo no XML
- **Arestas rotuladas com restrições de cardinalidade**
 - mapeamento de relacionamentos ou ligações entidade-atributo do ER
 - modelam relacionamentos hierárquicos ou ligações elemento-atributo no XML
 - 2 ou mais arestas podem ter uma **restrição de disjunção** (útil para modelar especializações exclusivas)

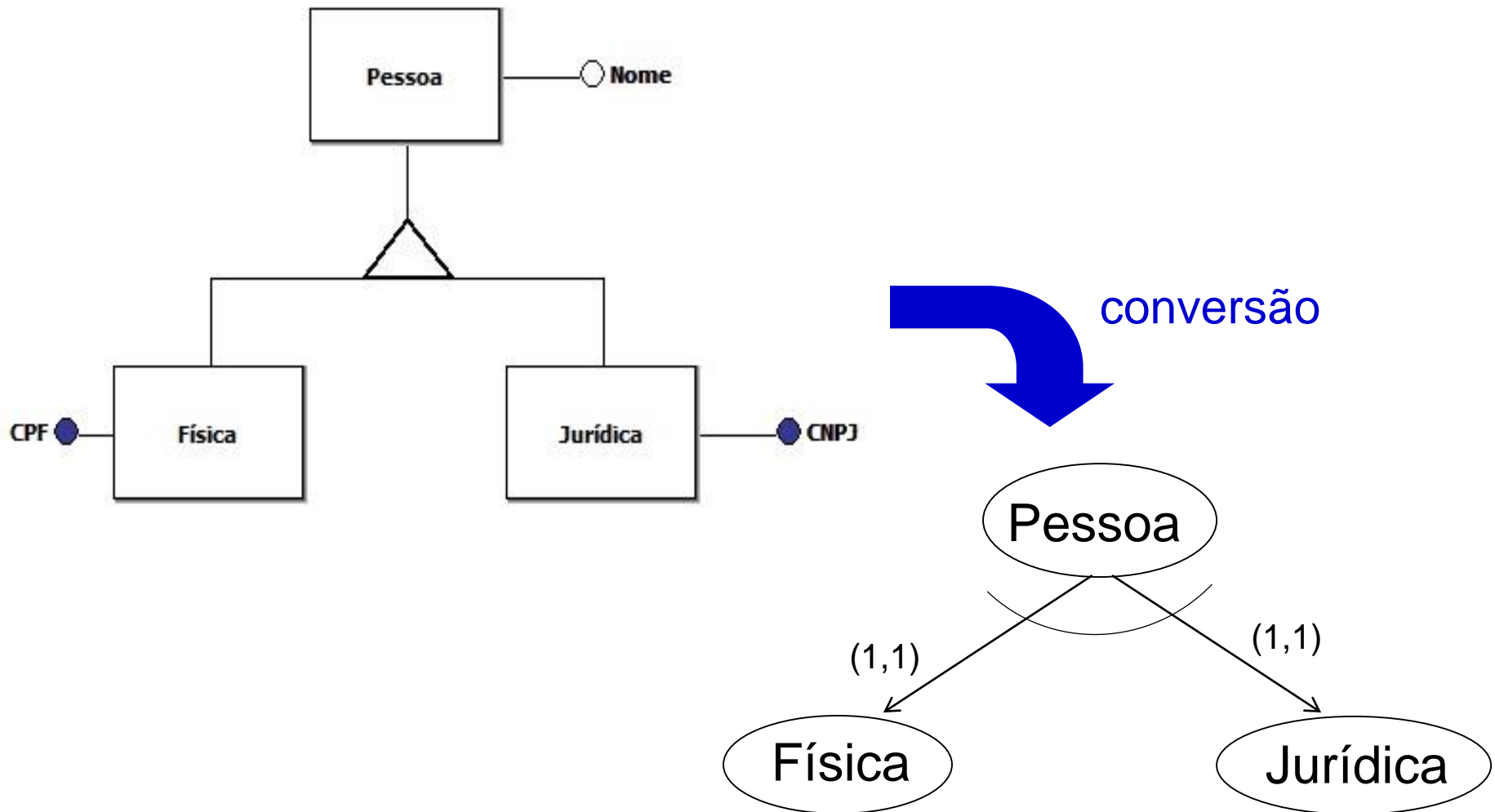
Modelagem Lógica

- Eleição do **nodo não-léxico central**
 - **entidade central na modelagem conceitual**
 - a partir dela uma hierarquia de nodos é definida com base nos seus relacionamentos no ER (preferência por caminhos com cardinalidade 1:1 ou 1:N)
 - exemplo: **Livro**
 - **um nodo raiz deve ser definido como t]TJ4este nodo (com cardinalidade 1:N)**
 - sugestões de nomenclatura
 - conjunto de ocorrências da entidade central (ex.: **Livros**)
 - nome do domínio (ex.: **Livraria**, **Biblioteca**, ...)
- Mais de um nodo central pode existir...
 - entidades (ex.: livros e funcionários de uma biblioteca)
 - todos serão filhos do nodo raiz (ex.: biblioteca)

Modelagem Lógica - Exemplo



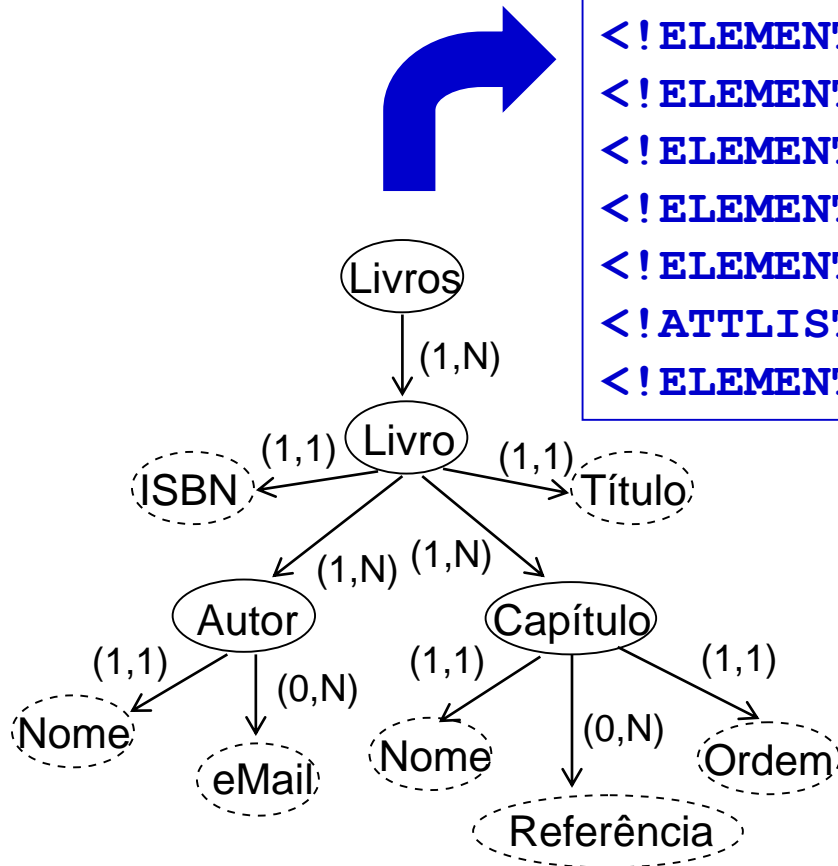
Modelagem Lógica Exemplo de Restrição de Disjunção



Modelagem Física

- Definição de elementos e atributos do esquema
 - nodos não-léxicos \Rightarrow elementos compostos
 - nodos léxicos \Rightarrow elementos simples ou atributos
- Definição da **ordem** de subelementos
 - ordenar as arestas que partem do nodo não-léxico
- Modelagem física de um nodo léxico
 - **como atributo**
 - economia de espaço no documento XML
 - pode-se definir algumas RIs específicas
 - **como elemento**
 - conteúdos extensos (uso de *tags* para representa-lo)
 - necessário quando se deseja subelementos com cardinalidade > 1
 - melhor legibilidade do documento XML

Modelagem Física - Exemplo



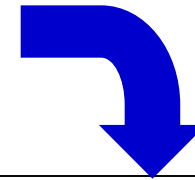
```
<!ELEMENT Livros (Livro+)>
<!ELEMENT Livro (Título, Autor+, Capítulo+)>
<!ATTLIST Livro ISBN CDATA>
<!ELEMENT Título (#PCDATA)>
<!ELEMENT Autor (Nome, eMail*)>
<!ELEMENT Nome (#PCDATA)>
<!ELEMENT eMail (#PCDATA)>
<!ELEMENT Capítulo (Nome, Referência*)>
<!ATTLIST Capítulo ordem CDATA>
<!ELEMENT Referência (#PCDATA)>
```

DTD

Modelagem Física - Revisão

```
<!ELEMENT Livros (Livro+)>
<!ELEMENT Livro (Título, Autor+, Capítulo+)>
<!ATTLIST Livro ISBN CDATA>
<!ELEMENT Título (#PCDATA)>
<!ELEMENT Autor (Nome, eMail*)>
<!ELEMENT Nome (#PCDATA)>
<!ELEMENT eMail (#PCDATA)>
<!ELEMENT Capítulo (Nome, Referência*)>
<!ATTLIST Capítulo ordem CDATA>
<!ELEMENT Referência (#PCDATA)>
```

capítulos possuem
conteúdo
semiestruturado



```
<!ELEMENT Livros (Livro+)>
<!ELEMENT Livro (Título, Autor+, Capítulo+)>
<!ATTLIST Livro ISBN CDATA>
<!ELEMENT Título (#PCDATA)>
<!ELEMENT Autor (Nome, eMail*)>
<!ELEMENT Nome (#PCDATA)>
<!ELEMENT eMail (#PCDATA)>
<!ELEMENT Capítulo (Nome, Texto)>
<!ELEMENT Texto (#PCDATA | Referência)*>
<!ATTLIST Capítulo ordem CDATA>
<!ELEMENT Referência (#PCDATA)>
```


Atividade 4

Apresente uma modelagem lógica para um BD XML e uma modelagem física em DTD a partir desta modelagem ER abaixo no domínio de um Museu

