

UFSC-CTC-INE

INE 5600 – Bancos de Dados III

Modelagem de Dados XML

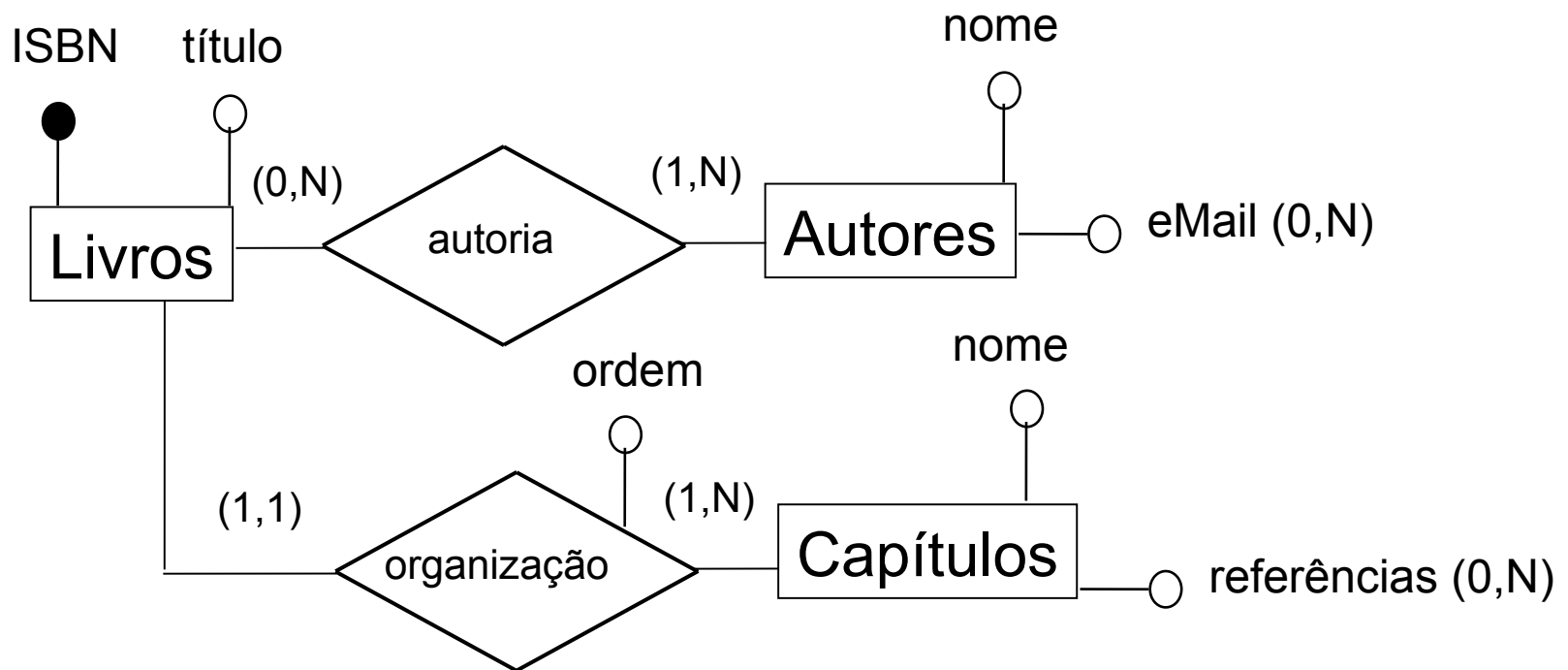
Projeto de BDs XML

- Não há uma metodologia consolidada
- Etapas do projeto tradicional de um BD
 - (i) levantamento de requisitos de dados; (ii) modelagem conceitual; (iii) modelagem lógica e (iv) modelagem física ou implementação
 - podem ser aplicadas a um BD XML
 - no caso de dados XML **fortemente semiestruturados**
 - revisão da modelagem física: considerar a existência de informação textual não-estruturada dentro do conteúdo de elementos

Guia para Projeto de BDs XML

- **Modelagem conceitual**
 - uso de algum modelo clássico (exemplo: ER)
- **Modelagem lógica**
 - uso de algum modelo baseado em grafo direcionado
 - adequado à representação de uma hierarquia XML
- **Modelagem física**
 - especificação do esquema XML (DTD ou XSD)

Modelagem Conceitual - Exemplo



Modelagem Lógica: Grafo Direcionado

- **Nodos não-terminais**

- mapeamento de entidades do ER
- representam elementos compostos do XML

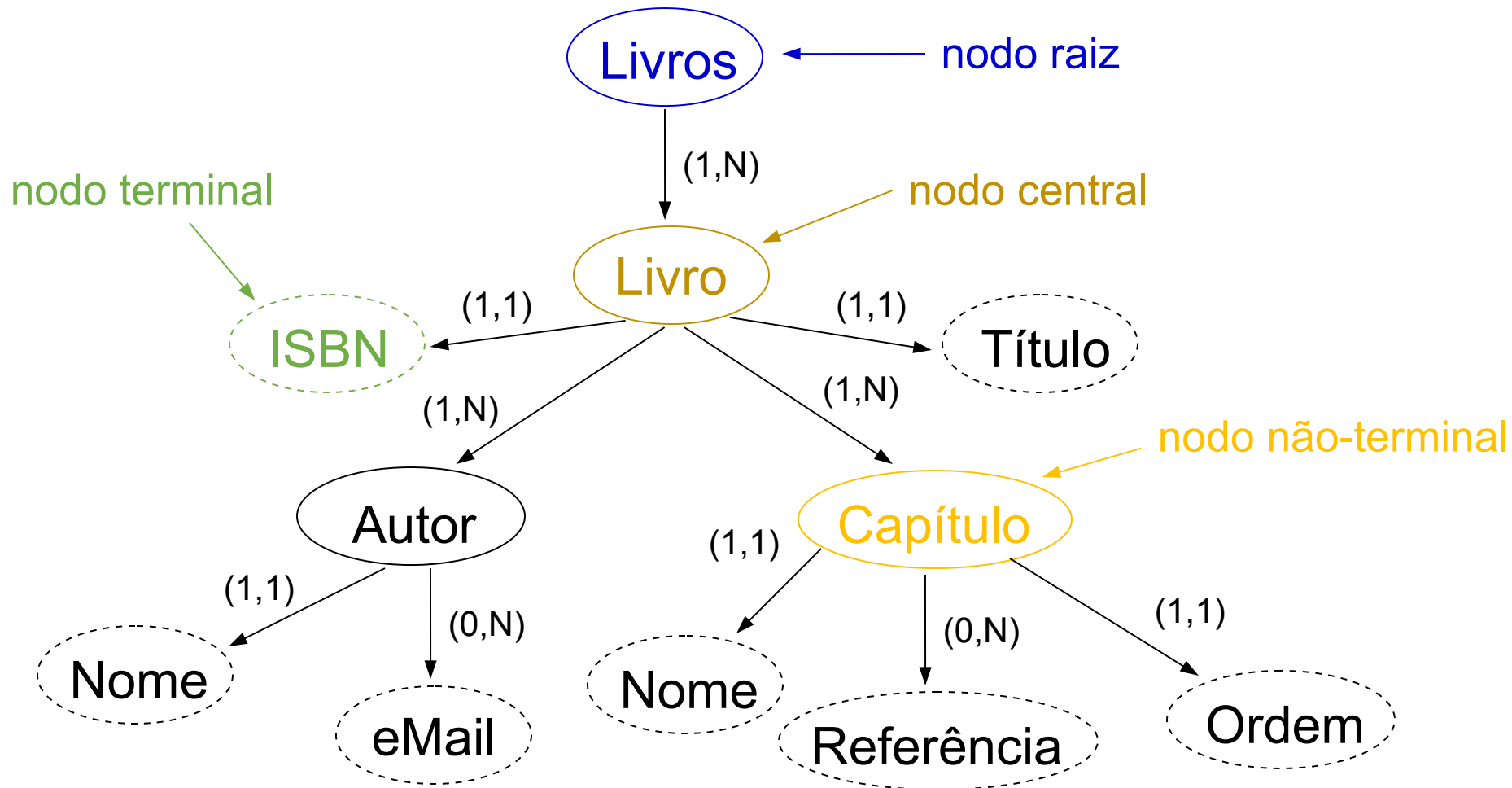
- **Nodos terminais**

- mapeamento de atributos do ER
- representam elementos simples ou atributos do XML

- **Arestas rotuladas com restrições de cardinalidade**

- mapeamento de relacionamentos ou ligações entidade-atributo do ER
- modelam relacionamentos hierárquicos ou ligações elemento-atributo do XML
- 2 ou mais arestas podem ter uma **restrição de disjunção** (útil para modelar especializações exclusivas)

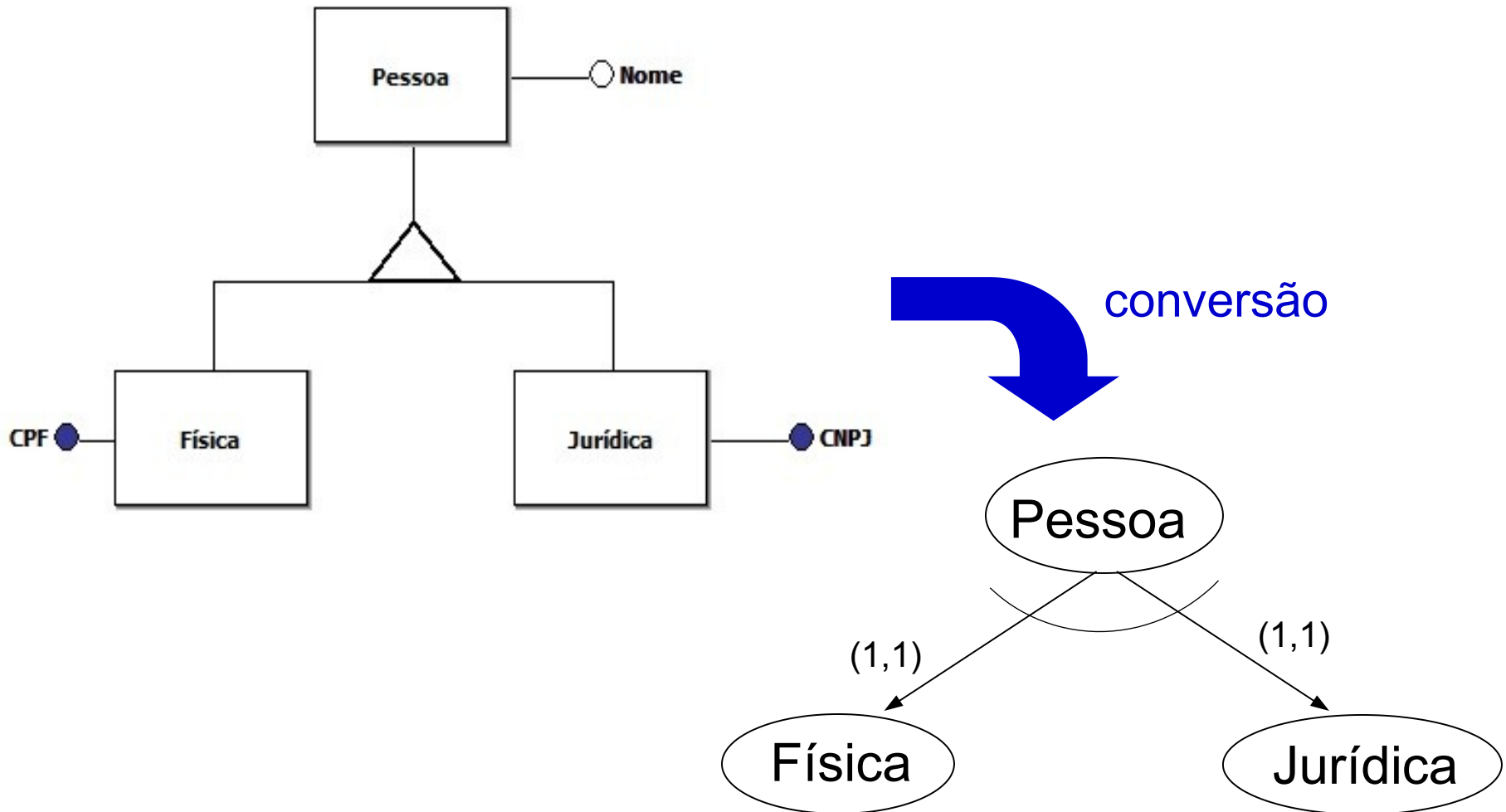
Modelagem Lógica - Exemplo



Modelagem Lógica

- Eleição do **nodo não-terminal 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 pai deste 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 “independentes” (ex.: livros e funcionários de uma biblioteca)
 - todos serão filhos do nodo raiz (ex.: Biblioteca)

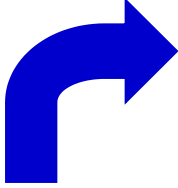
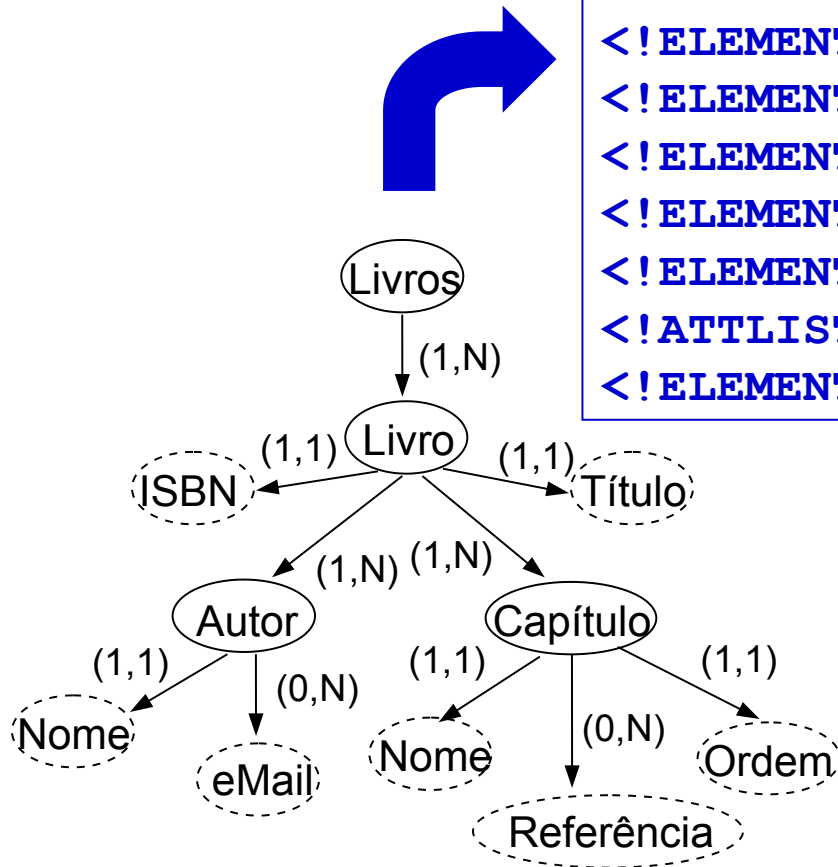
Modelagem Lógica – Exemplo de Restrição de Disjunção



Modelagem Física

- Definição de elementos e atributos do esquema XML
 - nodos não-terminais → elementos compostos
 - nodos terminais → elementos simples ou atributos
- Ordem de subelementos é desejada?
 - SIM: ordenar as arestas que partem do nodo não-terminal
- Modelagem física de um nodo terminal
 - como atributo
 - economia de espaço no documento XML
 - pode-se definir algumas RIs
 - como elemento
 - conteúdo mais extenso (uso de *tags* para representá-lo)
 - necessário quando se deseja subelementos com cardinalidade > 1
 - melhor legibilidade do documento XML

Modelagem Física - Exemplo



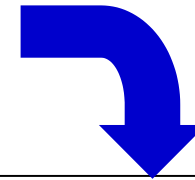
```
<!ELEMENT Livros (Livro+)>
<!ELEMENT Livro (Título, Autor+, Capítulo+)>
<!ATTLIST Livro ISBN CDATA>
<!ELEMENT Título (#PCDATA)>
<!ELEMENT Autor (Nome, eMail*)>
<!ELEMENT Nome (#PCDATA)>
<!ELEMENT eMail (#PCDATA)>
<!ELEMENT Capítulo (Nome, Referência*)>
<!ATTLIST Capítulo ordem CDATA>
<!ELEMENT Referência (#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 3

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

