

Semantic Data Modeling using XML Schemas

Murali Mani, Dongwon Lee,
Prof. Richard R Muntz
UCLA
Nov 27, 2001

Outline of the presentation

- Data modeling using XML Schemas – why do we need to study this and how difficult is it?
- XGrammar characteristics and notations
- XGrammar ↔ Entity relationship model
- Some interesting issues...

Motivation

- Data modeling is very crucial in successful database design.
- Significantly new features provided by XML Schemas
 - Ordered relationships

Ordered relationships



Book (0, *) => Person (1, *)

Motivation (contd...)

- Data modeling is very crucial in successful database design.
- Significantly new features provided by XML Schemas
 - Ordered relationships
 - Closure under union

XGrammar

- Defined by a 7-tuple $(N_T, N_H, T, S, E, H, A)$
 - N_T is the set of tree types
 - N_H is the set of hedge types
 - T is the set of terminal symbols
 - S is the set of start symbols, $S \subseteq N_T$
 - E is the set of element production rules
 - H is the set of hedge production rules
 - A is the set of attribute production rules

- $N_T = \{\text{Library, Book, Magazine, Person}\}$
- $N_H = \{\text{LibraryContent}\}$
- $S = \{\text{Library}\}$
- E : Library \rightarrow library (LibraryContent)
 Book \rightarrow book (ϵ)
 Magazine \rightarrow magazine (ϵ)
- H : LibraryContent \rightarrow (Book*, Magazine*)
- A : Book \rightarrow book (@authors::IDREFS \rightsquigarrow Person+)
 Magazine \rightarrow magazine (@editor::IDREF \rightsquigarrow Person)

XGrammars are closed under union

- Union of any two tree types is a tree type
 Book \rightarrow book (Title, Author+)
 Magazine \rightarrow magazine (Name, Editor)
 ReadingMaterial \rightarrow (Book | Magazine)
- Schema language proposals such as RELAX, TREX, RELAX-NG, XDoc are closed under union.

Mapping from EER \leftrightarrow XGrammar

1. Using Element Production rules – ordered 1:n binary relationships



Book (1, 1) \Rightarrow Person (1, *)

E: Book \rightarrow book (Person+)

2. Using IDREF – unordered 1:n binary relationships



Book (1, 1) \rightarrow Person (0, *)

A: Person \rightarrow person (@authorOf::IDREF \rightsquigarrow Book)

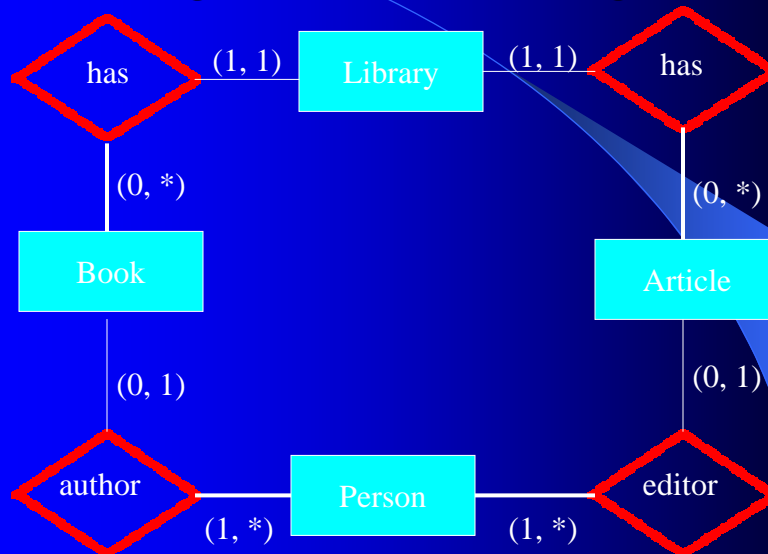
3. Using IDREFS – ordered n:m binary relationships



Book (0, *) => Person (1, *)

A: Book → book (@author::IDREFS ↪ Person+)

Significant challenges



Best XGrammar possible at present

E: Library \rightarrow library (Book*, Article*, Person*)
A: Book \rightarrow book (@authors::IDREFS \rightsquigarrow Person+)
Article \rightarrow article (@authors::IDREFS \rightsquigarrow Person+)

Desired XGrammar

E: Library \rightarrow library (Book*, Article*)
Book \rightarrow book (Person+)
Article \rightarrow article (Person+)

We need more information for this

- $\text{Person}_{\text{Book}} \cap \text{Person}_{\text{Article}} = \phi$
- $\text{Person}_{\text{Book}} \cup \text{Person}_{\text{Article}} = \text{Person}$