Conceptual Modeling for XML - A Survey

Martin Necasky
MFF UK
april, 2006

- requirements
- existing approaches
Synopsis

- introduction
- requirements for conceptual models for XML
- existing approaches
Why Conceptual Model for XML?

• recently, XML is frequently applied as a logical database model
• we have several XML schema languages
• however, these languages are unsuitable to use on the conceptual level:
  – too weak (DTD) or too complex syntax for designers (XML Schema)
  – better for use as schema languages on the logical level
Special Features of XML

- hierarchical structure
- irregular structure
- ordering
- mixed content
Requirements for conceptual models for XML
General Requirements (1)

- independence on XML schema languages
- formal foundations
- user-friendly graphical notation
- logical level mapping
General Requirements (2)

• support for hierarchical views
  – different users with different requirements accessing the modeled data
  – each of the users may require different hierarchical organization of the same data – hierarchical views
  – we need to model the hierarchical views on the conceptual level
General Requirements (3)

- integration with semantic web technologies
  - a translation from the conceptual level to the semantic web level where the structures from the conceptual level are described using OWL
  - automatic publication of internally represented data on the semantic web
  - automatic integration of data from the semantic web to the internal representation
Modeling Constructs Requirements

• classical features:
  – many-to-many rel. types
  – n-ary rel. types, attributes of rel. types

• special features:
  – hierarchical structure
  – ordering, mixed content
  – irregular structure
Approaches
Two main approaches (1)

- extension of the classical E-R model to be suitable for the modeling of special XML features
  - problem with the modeling of hierarchical structures and different hierarchical views of the same data
  - how to model ordering?
  - how to model data mixed with text values?
  - how to translate a non-hierarchical conceptual schema to a hierarchical logical schema?
Two main approaches (2)

- hierarchical approach
  - schemes are modeled as trees
  - allows to model hierarchical structures and different hierarchical views
  - problem with the modeling of many-to-many relationship types, n-ary relationship types, and attributes of relationship types
EReX -
An E-R based conceptual model for XML
EReX (1)

- allows to model irregular structure and ordering:
  - categorization of entity types
    - similar to IS-A hierarchies in E-R
  - total/exclusive coverage constraints
    - $E1 + E2 + E3 = E$
    - $E1 \mid E2 \mid E3$
EReX (2)

- ordering specified on an entity type \( E \) participating in a relationship type \( R \)
  - for each entity \( e \) from \( E \) the set of relationships from \( R \) having \( e \) as a participant is linear ordered
EReX (3)

Student + Professor = Person
Student | Professor
ORA-SS -
A hierarchical conceptual model for XML
ORA-SS (1)

- n-ary relationship types and attributes of relationship types
- cardinality constraints for both participants of relationship types
- ordering
- disjunction
ORA-SS (2)
ORA-SS (3)
ORA-SS (4)
Conclusion

- poor support of the specific XML features
- utilization of conceptual models for the data integration and for the integration with the semantic web has not been studied enough yet
- there is no conceptual model combining the advantages of the both approaches (E-R and hierarchical)
- open space for research
Thank you for your attention!