

Translating Relational Schemas to XML Schemas

Dongwon Lee, Murali Mani,
Frank Chiu, Wesley. W. Chu
UCLA

Our goals .

- Map from relational to XML model automatically.
- Maintain semantic constraints during the mapping.
- Existing work such as XML Extender (from IBM), db2XML, XML-DBMS, SilkRoute, EXPERANTO rely on user-specified mapping, and do not maintain semantic constraints.

CoT: Constraint-based Translation: Step 1

- Consider IND $s[\alpha] \subseteq t[\beta]$, where $\alpha \subseteq X$, $\beta \subseteq Y$, β is primary key, and α is non-nullable.
 - If α is unique, $M(t) = (Y, s?)$, else $M(t) = (Y, s^*)$
 - $M(s) = (X - \alpha)$
 - Key for s is $(K_s - \alpha)$

Sname	Advisor	Course
John	Muntz	DB
John	Muntz	N/W

Pname	Age
Muntz	60
Chu	55


```

<professor>
  <Pname>Muntz</Pname>
  <Age>60</Age>
</student>
  <Sname>John</Sname>
  <Course>DB</Course>
</student>
</student>
  <Sname>John</Sname>
  <Course>N/W</Course>
</student>
</professor>
    
```

CoT: Step 2

- Consider tables s , t_1 , t_2 with column set X , Y_1 , Y_2 , and INDs $s[\alpha] \subseteq t_1[\beta_1]$, and $s[\gamma] \subseteq t_2[\beta_2]$, where β_1 , β_2 are primary keys and α, γ are non-nullable
- Translate one IND as in Step 1, and translate the other to IDREF as:

$$M(t_1) = (Y_1, s^*), M(t_2) = (Y_2), M(s) = (X - \alpha - \gamma), A(t_2) = \{ID_t2::ID\}, A(s) = \{Ref_t2::IDREF\}$$

Cname	Room
DB	4760
N/W	4549


```

<course
  ID_course="DB"/>
<course
  ID_course="N/W"/>
</professor>
  <Pname>Chu</Pname>
  <Age>55</Age>
</professor>
  <Pname>Muntz</Pname>
  <Age>60</Age>
  <student Ref_course="DB">
    <Sname>John</Sname>
  </student>
  <student Ref_course="N/W">
    <Sname>John</Sname>
  </student>
</professor>
    
```

CoT: Step 3

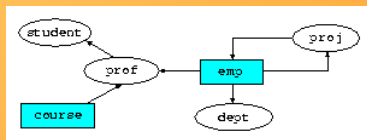
- Consider a relational schema with tables $\{t_1, t_2, \dots, t_n\}$ and INDs $t_i[\alpha_i] \subseteq t_j[\beta_j]$
- Construct an IND-graph
- Identify top-nodes as:
 - Nodes that do not have any IND are top-nodes
 - In a strongly connected component formed from table-set, S , if there is no IND from a node in S to a node outside S , then one of the nodes in S must be a top-node.
- Perform BFS and translate the IND as in Step 1 or Step 2.

student (<u>Sid</u> , Name, Advisor) emp [Eid, Name, ProjName] prof [Eid, Name, Teach] course [<u>Cid</u> , Title, Room] dept (<u>Dno</u> , Mgr) proj [Pname, Pmgr]
student [Advisor] \subseteq prof [Eid] emp [projName] \subseteq proj [Pname] prof [Teach] \subseteq course [Cid] prof [Eid, Name] \subseteq emp [Eid, Name] dept [Mgr] \subseteq emp [Eid] proj [Pmgr] \subseteq emp [Eid]

$M(\text{course}) = (\text{Cid}, \text{Title}, \text{Room}, \text{prof}^*)$
 $M(\text{prof}) = (\text{student}^*)$
 $M(\text{student}) = (\text{Sid}, \text{Name})$
 $M(\text{emp}) = (\text{Eid}, \text{Name}, \text{dept}^*, \text{proj}^*)$
 $M(\text{dept}) = (\text{Dno})$
 $M(\text{proj}) = (\text{Pname})$

$A(\text{emp}) = \{\text{ID_emp::ID}, \text{Ref_proj::IDREF}\}$
 $A(\text{prof}) = \{\text{Ref_emp::IDREF}\}$
 $A(\text{proj}) = \{\text{ID_proj::ID}\}$

IND-Graph



Implementation Grammar syntax

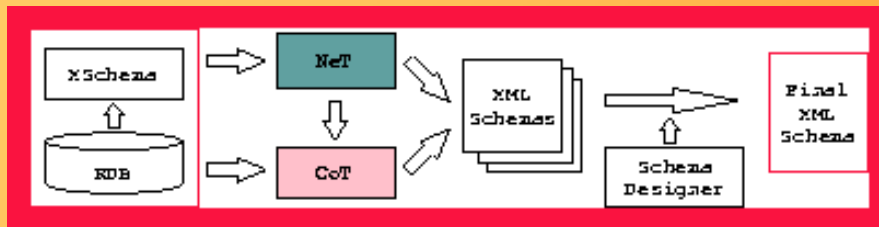
```
root → (course *, emp *)
course → (Cid, Title, Room, prof *)
emp → (Eid, Name, dept *, proj *)
@emp → (@ID_emp, @Ref_proj]
prof → (student *)
@prof → (@Ref_emp)
student → [Sid, Name)
dept → (Dno)
proj → (Pname)
@proj → (@ID_proj)
```

```
primKeys (course) = Cid
primKeys (emp) = Eid
primKeys (prof) = @Ref_emp
primKeys (student) = Sid
primKeys (dept) = Dno
primKeys [proj] = Pname
```

CoT Implementation

- Available at <http://www.cs.ucla.edu/~mani/net/vldb/code/cot.jar>
- Converts a set of relations into a single XML document, and a schema for it.
- Document presently constructed using DOM.
- Resulting schema has following features
 - Local tree grammar
 - No recursion
 - An element type corresponding to a relation occurs only once in the grammar

Conclusions



- **Automatically map from relational to a good XML model.**
- **Maintain semantic constraints.**
- **Remove some of the redundancies that could have been present in the relational model.**