Comparing Topic Maps Constraint Specification Languages

Giovani Rubert Librelotto
giovani@unifra.br

José Carlos Ramalho
jcr@di.uminho.pt

Renato Preigschadt de Azevedo
rpa.renato@gmail.com

Pedro Rangel Henriques
prh@di.uminho.pt
Agenda

• Motivation
• Our work’s background
• What’s going on (models)?
• Constraints...
• XTche, OSL, Toma, and AsTMa!
• Case study
• Conclusions
Motivation

• Why do we need constraints?
  – Topic Maps are too abstract!
    ‣ Many ontologies have a rigid structure
    ‣ Many ontologies have semantic constraints
  – we do not control the process of creation
    ‣ uploads
    ‣ merges
  – manual editing of the Topic Map

• As a side effect we can have a Reporting Service
Ontology

Structure level

Catalog level
More questions

• Is a Schema a Constraint Language?

• Creating a new language ... Why?
  ‣ We are always creating new DSL arguing that we will hide complexity
  ‣ Aren´t we introducing complexity?

• How should we implement this new language?
Back in 2001 we had great expectations towards Topic Maps.

Returning to constraints

• What is really a TM?

• According to the XTM DTD
  ‣ A list of ... Topics, associations, occurrences
  ‣ Everything else is constructed by reference
  ‣ It is easy to get lost in a TM!

• That DTD is completely abstract
  ‣ There is room for “Specific Semantic Networks Schemas”
TMCL: Implementation wishes

• Do it simple

• Do it with existing technology

• Do it user friendly
  – If possible do not force the user to learn a new language
Schema constraints:
- Topic of type T must have a specified number of explicit names/occurrences/subject-indicators (cardinality);
- Topic of type T must have as name/occurrence/subject-indicators a value matching a particular pattern;
- Topic of type T must (not) have a name/occurrence with scope S;
- Topic of type T must have a name/occurrence, that is instance of topic type T, in scope S;

... 

Contextual constraints:
- Topic T can (only/not) be used for typing other topics;
- Topic T can (only/not) be used for typing subject indicator;
- Topic T can (only/not) be used for typing basenames;
- Topic T can (only/not) be used for typing occurrences;

...
XTche Specification Language

• XML Schema based
  – Constraint lang. = Schema lang.
  – We have some experience with XML Schemas
  – The tools we use have an interesting graphical support that we wanted to use
  – If we choose an existing language to host ours part of the work is already done

• XTche Spec. = XML Schema + semantic stamps
XTche skeleton

• Namespace: xmlns:xtche=http://www.di.uminho.pt/~gepl/xtche

• Imports: http://www.di.uminho.pt/~gepl/xtche/xtche-schema.xsd
xtche-schema.xsd

• A set of flags / semantic stamps

• Each flag represents a semantic validation

• The user will associate flags to concrete topics and these will map to semantic constraints that will be verified

• Remaining: **Did we cover the whole TMCL?**
Flags / Semantic Stamps

• **Schema constraints**
  – `<xs:attribute name="topicType"/>`
  – “this element represents a Topic type”

• **Contextual constraints**
  – `<xs:attribute name="associationTypeExclusive"/>`
  – “this topic can only be used to qualify associations”

• **Existence constraints**
  – To guarantee the existence of at least a specific topic or association
Astma language set

- Non-XML syntax

- Language for: structuring, adding, querying, constraining...

- An object algebra...

- It has its own model. The Topic Map must be uploaded to this model.
TOMA

• *Toma* is an "all-in-one" TM*L*:  
  – TMQL, TMML, and TMCL

• *Toma* offers statements like:  
  – SELECT, INSERT, UPDATE and DELETE

• *Toma* provides functions that allow to modify, convert and aggregate the data coming from the topic map.
Ontopia Schema Language (OSL)

- XML Schema based
- Allows the user to specify a schema for a specific Topic Map
- Since we are in the schema domain we are not allowed to use instance values
E-Sell case study (by ISO)

Customers, products, and orders

Diagram showing relationships between customers, products, and orders.
∀a ∈ Associations: if(Type(a) = “is-making-order” then MemberRoles((order, product, quantity) a)

**XTche approach:**
a) Restricting an association structure

**AsTMa! approach:**

forall [ * (is-making-order)  
    product: *  
    quantity: *  
    order: *   ]

**OSL approach:**

```xml
<association>
    <instanceOf>
        <internalTopicRef href="#is-making-order"/>
    </instanceOf>
    <role min="1" max="1">
        <instanceOf>
            <internalTopicRef href="#product"/>
        </instanceOf>
        <player>
            <internalTopicRef href="#product"/>
        </player>
    </role>
    <role min="1" max="1">
        <instanceOf>
            <internalTopicRef href="#order"/>
        </instanceOf>
        <player>
            <internalTopicRef href="#order"/>
        </player>
    </role>
    <role min="1" max="1">
        <instanceOf>
            <internalTopicRef href="#quantity"/>
        </instanceOf>
        <player><any/></player>
    </role>
</association>
```

**TOMA approach:**

define constraint is_making_order_constraint  
each association $a$(is-making-order)  
satisfies exists $a$(is-making-order)$\rightarrow$product = $$  
and $a$(is-making-order)$\rightarrow$quantity = $$  
and $a$(is-making-order)$\rightarrow$order = $$;
b) Each person must have a phone or fax field

\[ \forall t \in \text{Topics}, \text{TopicType}(t) = "customer" \Rightarrow \exists r \in \text{Occurrences} : r \subset t \wedge (\text{OccType}(r) = "phone" \lor \text{OccType}(r) = "fax") \]
b) Each person must have a phone or fax field

AsTMa!

\[\forall c \in \text{customer} \implies \exists c \in \text{phone} \lor \exists c \in \text{fax}\]

TOMA

define constraint customer_must_have_a_contact_number_constraint
    each topic $t$
        where $t$.type = 'customer'
        satisfies exists $t$.oc.id = 'phone'
        or $t$.oc.id = 'fax';
b) Each person must have a phone or fax field

```
<topic>
  <instanceOf>
    <internalTopicRef href="#customer"/>
  </instanceOf>
  <occurrence min="0" max="1">
    <instanceOf>
      <internalTopicRef href="#phone"/>
    </instanceOf>
  </occurrence>
  <occurrence min="0" max="1">
    <instanceOf>
      <internalTopicRef href="#fax"/>
    </instanceOf>
  </occurrence>
</topic>
```

The code above defines a topic instance of customer that has zero or one phone occurrence and zero or one fax occurrence.

But, according this OSL specification, there is no way to impede that a topic instance of customer has both occurrences.
c) Contextual constraint

Restricting the use of a specific topic: topic “wine” can only be used as a Topic Type

**XTche**

forall \[ \text{wine} \] => exists \[ w \ (\text{wine}) \]

**AsTMa!**

forall [ wine  
bn: Wine]  
=> exists [ $w \ (\text{wine}) \ ]
c) Contextual constraint

Restricting the use of a specific topic: topic “wine” can only be used as a Topic Type

**TOMA**

```plaintext
define constraint wine_is_type
  each topic $t
  satisfies exists $t.instance
  where $t.id = 'wine'
```

**OSL**

There are no constructors for that.
## Comparing the 4 Languages

<table>
<thead>
<tr>
<th></th>
<th>XTche</th>
<th>AsTMa!</th>
<th>TOMA</th>
<th>OSL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Background</strong></td>
<td>XML Schema + Stamps</td>
<td>Specific Syntax</td>
<td>SQL, Tolog, OO, AsTMa*</td>
<td>XML Schema</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td>Abstract Doc. Tree (XTM)</td>
<td>Specific (should be open)</td>
<td>TMCL, TMML, and TMQL</td>
<td>Abstract Doc. Tree (XTM)</td>
</tr>
<tr>
<td><strong>Range of App.</strong></td>
<td>Till now... it looks promising</td>
<td>Some doubts...</td>
<td>TMCL, TMML, and TMQL</td>
<td>Only structural constraints</td>
</tr>
<tr>
<td><strong>Scalability</strong></td>
<td>Does not scale</td>
<td>???</td>
<td>Does not scale</td>
<td>Does not scale</td>
</tr>
</tbody>
</table>
## Comparing the 4 Languages

<table>
<thead>
<tr>
<th></th>
<th>XTche</th>
<th>AsTMa!</th>
<th>TOMA</th>
<th>OSL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TM syntax</strong></td>
<td>XTM</td>
<td>AsTMa=</td>
<td>Database ?</td>
<td>XTM, HyTM, LTM, RDF</td>
</tr>
<tr>
<td><strong>Existence, boolean, and conditional constraints</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Data types</strong></td>
<td>Based on XML Schema data types</td>
<td>Strings</td>
<td>Basic data type: string, int, float, ...</td>
<td>No</td>
</tr>
</tbody>
</table>

26
Conclusions

• The main conclusion is that XTche, Toma and AsTMa! comply with all requirements stated for TMCL whereas OSL just includes topic maps structure validation.

• There is work to be done
  – To evaluate the processors performance;
Questions...?

Giovani Rubert Librelotto
giovani@unifra.br

José Carlos Ramalho
jcr@di.uminho.pt