HCU

XML and TEI

OUCS

TEI and Schema

Sebastian Rahtz
November 2001

Summary

1. Schemas are a useful replacement for DTDs
2. The TEI can benefit from schema in source maintenance
3. Or we can generate schema from DTDs
4. There is little to lose from working with schemas

What are schemas for?

Doing the same job as DTDs, but

☛ Expressing the syntactic constraints in an XML-based language, allowing for extensions to DTD features

☛ An opportunity to add extensive data-typing

☛ Providing a resource which standard XML tools can access

☛ Selective validation, using namespaces

In fact, so good that we have several Schema languages! We have looked at Relax NG.

XML and TEI

HCU

OUCS

Examples

- ☞ We can model a content rule for `<frontmatter>` which says "`<title>`, `<author>` and `<date>` must be present, but the order does not matter

- ☞ We can constrain the value of the attribute "age" to be a positive integer between 1 and 256

- ☞ We can specify rules for the PCDATA *content* of an element, not only attributes

- ☞ We can say that the content of the TEI `<formula>` element is in MathML and validation that with a different schema

XML and TEI

The TEI in Schema

We can work in two ways:

1. Leave the source as is, and add a new feature to the Pizza Chef to generate a Schema from the compiled DTD

2. Convert the source ODD files to schema, and generate DTD files on demand.

The attractions of the latter route are:

1. We can encapsulate some of the current prose into formal schema data-typing

2. We can validate the ODD files right to the lowest level

ODD to pseudo-Relax NG

Odd:

```
<elemDecl> %om.RR; (#PCDATA | %m.placePart;
    | %m.phrase; | %m.Incl;)*</elemDecl>
```

New Odd:

```
<schemaDecl omit="RR"><schemaFragment>
  <rng:zeroOrMore
    xmlns:rng="http://relaxng.org/ns/structure/0.9">
    <rng:choice>
      <rng:text/>
      <rng:ref name="m.placePart"/>
      <rng:ref name="m.phrase"/>
      <rng:ref name="m.Incl"/>
    </rng:choice>
  </rng:zeroOrMore>
</schemaFragment></schemaDecl>
```

XML and TEI

Relax NG sample

```
<define name="set">
      <ref name="TEI.drama"/>
      <element name="set">
          <ref name="al.set"/>
          <group>
              <zeroOrMore>
                  <ref name="m.Incl"/>
              </zeroOrMore>
              <optional>
                  <ref name="head"/>
              </optional>
              <ref name="component.seq"/>
          </group>
      </element>
   </define>
```

A real Relax NG schema

```xml
<?xml version="1.0" encoding="iso-8859-1"?>
<grammar xmlns="http://relaxng.org/ns/structure/0.9"
  xmlns:a="http://relaxng.org/ns/compatibility/annotations/0.9

  datatypeLibrary="http://www.w3.org/2001/XMLSchema-
datatypes">

<include href="tei2.dtd.rng">
<define name="TEI.linking"><ref name="INCLUDE"/></define>

<define name="TEI.figures"><ref name="INCLUDE"/></define>

<define name="ab"><notAllowed/></define>
<define name="al.xref" combine="interleave">
   <ref name="a.global"/>
   <ref name="a.xPointer"/>
   <optional><attribute name="url"><text/></attribute></option

   <optional>
     <attribute name="TEIform" a:defaultValue="xref">
        <text/>
     </attribute>
   </optional>
 </define>
</include>
</grammar>
```

XML and TEI

Relax NG Schema resources (1)

☞ `http://www.relaxng.org/` - Relax NG home site, with specifications and papers.

☞ `http://sourceforge.net/projects/relaxng` - RELAX NG project at Sourceforge.

☞ `http://www.sun.com/software/xml/developers/multischema/` - Sun Multi-Schema XML Validator: a validator capable of using several schema languages, including RELAX NG.

☞ `http://www.thaiopensource.com/relaxng/jing.html` - Jing: A RELAX NG Validator in Java.

Relax NG Schema resources (2)

☛ `http://www.sun.com/software/xml/` `developers/instancegenerator/` - XML instance generator of XML documents from a schema.

☛ `http://www.sun.com/software/xml/` `developers/relaxngconverter/` - converts DTD, RELAX Core/Namespace, TREX and W3C XML Schema to the corresponding RELAX NG.

☛ `http://www.sun.com/software/xml/` `developers/schematronaddon/` - experimental validator: processes the combination of RELAX NG and Schematron.

☛ `http:` `//www.thaiopensource.com/dtdinst/` - DTDinst: a DTD to RELAX NG converter.

HCU

XML and TEI

OUCS