Standards-based Conformance Testing of IEC/ISO 29500 XML Resources

Seattle Edition
2009-09 (GB)

Menu

- From validation to conformance
- Ongoing work in this area
- A suggested methodology
- Technical Details
- Towards standardization

Conformity and Interoperability

- These concepts are closely related in JTC 1:

An assessment methodology for interoperability may include the specification of some or all of the following: terminology, basic concepts, requirements and guidance concerning test methods, the appropriate depth of testing, test specification and means of testing, and requirements and guidance concerning the operation of assessment services and the presentation of results. In technical areas where there is a conformity assessment methodology and an interoperability assessment methodology, the relationship between them must be specified.

(Directives, Annex I – Policy on Interoperability)

That Word, "Valid"

- In XML 1.x "valid" has a specific, technical meaning:
  - "An XML document is **valid** if it has an associated document type declaration and if the document complies with the constraints expressed in it"
- Today we often refer to "schema-valid" XML documents

A Taxonomy of Correctness

- Conformant
  - Obeys the provisions of a specification
- Valid
  - Of an XML document, that it obeys a schema
- Interoperable
  - A property of systems that interoperate
- Portable
  - A property of data that may be used interoperably

Conformance Testing – Why?

- Users !
  - Better correspondence between what the standard says, and what exists in reality …
  - … which benefits interoperability
  - … and straightforward "operability"
  - Generates feedback for the parent standard

The ODF Lesson

- ODF first published **May 2005**

- OASIS ODF Interoperability and Conformance (OIC) TC established **September 2008**

- And the lesson is for IS 29500 is:
  - Do not allow 3½ years to elapse before proactively addressing conformance issues

Work Underway

- ## SC 34/WG 5

  - "Develop principles of, and guidelines for, interoperability among documents represented using heterogeneous ISO/IEC document file formats… The initial work includes preparation of the Technical Report on ISO/IEC 26300 & ISO/IEC 29500 translation."

- ## Microsoft's DII Initiative

  - Microsoft partnership with Fraunhofer FOKUS to produce an "IS29500-Validator and Document-Library"

The Two Aspects

- Application Conformance
  - Does the application do what it must?
  - Does it handle/emit conformant documents
- Document Conformance
  - Are *documents* conformant to the standard
  - Valid
  - Obeying semantic provisions described in prose

Application Conformance

- Generally document format standards specify documents, not applications
- "Proving" applications correct is technically problematic
  - Testing is necessarily a scruffier art
- Scope is hard
  - Reading/Emitting valid documents?
  - Displaying italic text with an italic font?
  - Displaying Thai text correctly?

Use Case Based Testing

- Attempt to codify application conformance testing:
  - Assert pre-condition
  - Perform action
  - Assert post-condition
- Hard to automate

Document Conformance

- Package data integrity
  - A conformant ZIP
- Integrity of package parts
  - Graphics, text entities, etc.
- XML well-formedness
  - What claims to be XML, is XML
- XML validity
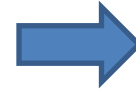 - XML valid to the IS 29500 schemas

Package Validation

50	4B	03	04	14
00	06	00	08	00
00	00	21	00	DD
FC	95	37	66	01
00	00	20	05	…

Package (bytes)

?

Validation Report

How do we standardize package validation?

Using XML Pipelines

- XProc – new W3C technology
- Declarative language for specifying multi-step operations to be performed on XML
- Strong on validation
  - Closely aligned with activities that have been taking place within SC 34
- Demo …

Going Further

- Some provisions of IS 29500 are not expressed in the schema, and yet are testable

- For example:

  - In CustomXML w:attr elements specify attribute names on a custom element; clearly no two w:attr elements can have the same name for one element

Semantic Validation

- Such "semantic" validation may be performed by applying a Schematron Schema

- Schematron is ISO/IEC 19757-3; invented by Rick Jelliffe

- A deliverable for conformance testing IS 29500 should include Schematron schema

- Gives deeper validation than allowed by the grammars alone

The double attribute issue

- "The w:attr element may be repeated to declare duplicate Namespace/Name pairs for attributes on a single element. Thus, the standard describes documents that can imply tagged text that cannot be processed using conformant XML processors"

Enhancing Application Validation

- Returning to application conformance …

- Both pre-conditions and post-conditions could be asserted using Schematron schemas

- Ideally the application action could be scripted too

  – One of many good arguments for standardizing office document scripting languages

The Case for Standardized Conformance Testing in SC 34

- Usual benefits of International Standardization, but especially:
  - Access to very wide constituency of users
  - Good mix of expertise in SC 34
  - Greater credibility than a MS-sponsored initiative

How Might it be Developed?

- Perhaps a TR (Technical Report)
  - Not a typical standard, can contain many kinds of content
  - Registry of rules? (use cases, schemas)
  - "Application Descriptions" (profiling)
- Developed across SC 34 Working Groups:
  - WG 1 (validation, conformance testing)
  - WG 4 (IS 29500 itself)
  - WG 5 (interoperability)

Next Steps

- The project needs to be formally proposed
  - "New Work Item Proposal" (NWIP)
  - Either by SC 34 itself or by a National Body
- 3 month JTC 1 ballot on the NWIP
  - Ideally, aiming to start work at the Stocholm meeting (2010)

Thank You

Any Questions?