

DFDL WG Call Minutes

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	DFDL 1.0
Meeting Date	24-May-16 (Tues)
Meeting Time	15:00 - 16:00

Created by Steve Hanson on 09-Mar-11

Last Modified by Steve Hanson on 25-May-16

OGF DFDL Working Group Call , 24 May 2016

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

1. Daffodil Open Source Project

Status update.

2. Deferred action 242

Mike has asked some questions about the behaviour of `dfdl:contentLength()`.

Steve thinks is covered by deferred action 242, so we should re-open it.

3. Clarify what empty string means for properties that are silent about it

Properties `dfdl:textBooleanTrueRep` and `dfdl:textBooleanFalseRep` do not state empty string behaviour.

Properties `dfdl:textStandardNanRep` and `dfdl:textStandardInfinityRep` do not state empty string behaviour.

4. US Gov Open-Data Initiative

See email from Mike.

5. DFDL needs additional function `dfdl:characterCode()` ?

Mike to elaborate.

6. AOB

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Attendees

Steve Hanson

Mike Beckerle

Apologies

Minutes

1. Daffodil Open Source Project

Not discussed.

2. Deferred action 242

Mike has asked some questions about the behaviour of `dfdl:contentLength()`.

Un-deferring action 242 to look at this again.

3. Clarify what empty string means for properties that are silent about it

Properties `dfdl:textBooleanTrueRep` and `dfdl:textBooleanFalseRep` should state empty string not allowed. **Issue <https://redmine.ogf.org/issues/312> created.**

Properties `dfdl:textStandardNaNRep` and `dfdl:textStandardInfinityRep` should state that empty string not allowed. **Issue <https://redmine.ogf.org/issues/313> created.**

4. US Gov Open-Data Initiative

https://fcw.com/articles/2016/04/14/open-data-bill.aspx?s=fcwdaily_150416

Consider doing a DFDL talk if any relevant meetings happen in Washington DC area.

5. DFDL needs additional function `dfdl:characterCode()` ?

Mike has encountered a format that encodes ASCII 0-9 and A-Z (minus 4 letters) in 5 bits.

One option is to treat the element as an integer then convert each 5 bits to a character using a new `dfdl:characterCode()` function. Will need `dfdl:inputValueCalc`, hidden groups.

Mike to consider whether a 5-bit encoding is a better way to solve this.

IPR Statement

"I acknowledge that participation in this meeting is subject to the OGF Intellectual Property Policy."

Meeting closed

15:45 UK

Next regular call

Tues 7th June 2016 @ 15:00 UK

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below. Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

Subject	Document Type	Created	Modified

Next action: **288**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
228	Review set of tutorial lessons (All) 17/9: Lesson 1 proposes a set of lessons, needs reviewing as over 2 years old. ... 22/10: No progress 31/10: Becoming a focus for Tresys. Steve to send his 'Modeling Data Formats using DFDL' powerpoint. ... 19/11: No further progress 26/11: Possibility of help from MITRE high-school student, and from Marisa at IBM. ... 11/3: No further progress 25/3: MITRE have produced a couple of new tutorials under the guidance of James Gariss . Jonathan to forward for review. Mike observed that an html tutorial could be generated from a tdml file using XSLT . 11/4: Not discussed 15/4: Jonathan will send 4 new mini-tutorials. Need to figure out best way to incorporate into the tutorial structure. 29/4: Tutorials received. Mark has taken a quick read. Mark & Steve to review and report back. 6/5: Still with Mark and Steve

	<p>20/5: Mark has reviewed. Will ask IBM information development to recommend a way to portray the existing and new lessons, preferably web-based. Find somewhere to host them. OGF? GitHub? developerWorks? NCSA?</p> <p>3/6: Steve has also reviewed.</p> <p>...</p> <p>17/6: No further progress on tutorials. Tim is looking into the creation of some DFDL how-to videos using the IBM Integration Studio.</p> <p>...</p> <p>31/3: No further progress</p> <p>14/4: Agreed that the need for better tutorials has become pressing for Daffodil users who aren't using IBM's tools and material. Discussed creating tutorials based on a tdml file with comments that is processed to produce html. Mike to investigate.</p> <p>28/4: Mike has sent an example tdml file which embeds instances of a new 'tutorial' element in various places. These elements contain html which can be extracted and formatted in a browser. Suggest future DFDL tutorials are created using this technology.</p> <p>12/5: Not discussed</p> <p>...</p> <p>22/9: No further progress</p> <p>3/11: Daffodil team has someone working on the new 'tutorial' element in tdml files. In time this should result in some new tutorials and re-working of existing tutorials.</p> <p>5/1/16: Mike has started a bitOrder tutorial using the tdml file approach (uses stylesheets to render html).</p> <p>16/2: The bitOrder tutorial is available on the web @ https://opensource.ncsa.illinois.edu/bamboo/artifact/DFDL-MASTER21/JOB1/build-132/Tutorial s/bitorder.tutorial.tdml.xml</p> <p>1/3: Awaiting review. Web-based try out facility under development at Tresys.</p> <p>...</p> <p>24/5: No further progress</p>
242	<p>Public comment : dfdl:valueLength and dfdl :contentLength descriptions (Mike)</p> <p>19/11: http://redmine.ogf.org/boards/15/topics/63. Agreed that the function names were ok as per errata 3.18, and that the spec is clear that they refer to the grammar regions. However the grammar regions mentioned do not fully include literal nil values. Discussed what happens when parsing - remember the length or re-parse? What about lengthUnits 'characters' when the data is binary? Also the 'Notes' that follow the table need to be reworked.</p> <p>26/11: Needs wording to handle all the issues found, assigned to Mike.</p> <p>...</p> <p>11/3: Still with Mike</p> <p>25/3: Mike has sent out revised wording, reviewed by Mark and Steve. Noted that the words need to explain the concept of building a complex element from the bottom up, and these words are equally applicable to several places in section 12.3. Mike to revise accordingly.</p> <p>11/4: More revised wording sent by Mike. Started to review but realised it needed some off-line preparation and thought. Review for next call.</p> <p>15/4: Review comments from Steve and Tim. The functions need to be clear that they work off the infoSet value. The detailed wording is needed but should be removed to a new sub-section of 12.3, probably at end. Most sub-sections of 12.3, and the functions in 23.5.3 will refer to this new sub-section. 23.5.3 should limit itself to behaviour specific to the functions, such as not potentially represented, the effect of the \$lengthUnits argument. Also discussed what happens if \$path argument returns a nodeSet > 1; should be a processing error, can always use a predicate to select one node of an array.</p> <p>29/4: See various email discussions. Several things noted by Mike, and he recommends a rewrite of some of section 12.3. Then the description of the two functions becomes much simpler. Deferring for now, and will resurrect after current spec revision is finalised.</p> <p>6/5: Mike is working on a mind map for the length section. Deferring until needed.</p> <p>....</p> <p>23/9: Rewrite should be postponed to future 1.1. Still need to answer the original questions about the functions though...</p>

	<p>...</p> <p>25/4/2016: Undeferring action as some of these issues are now impacting Daffodil team as they write their unparser. Steve has sent the email threads on this action to Mike. Mike will combine with his issues and distill into a single thread.</p>
250	<p>Standardise on a single tdml format for DFDL tests (All)</p> <p>5/2: Steve has requested permission for IBM to view / use the Daffodil tdml files, as a precursor to trying to standardise on a common tdml format. Was formerly part of action 066.</p> <p>...</p> <p>18/2: No further progress</p> <p>11/3: Mike and Steve discussing the best way to share and cooperate on tdml format.</p> <p>25/3: Discussed the creation of an OGF document that will own and define a standardised tdml format.</p> <p>11/4: Proposal is for the OGF document to define a tdml format without Tresys or IBM copyright statement.</p> <p>15/4: Draft document on Redmine</p> <p>...</p> <p>6/5: No further progress</p> <p>20/5: Mark has read through the document. Particularly concerned with how namespaces are handled in the info set.</p> <p>...</p> <p>17/6: No further progress</p> <p>25/6: Mike has added bit order capability as per action 233.</p> <p>...</p> <p>9/12: No further progress</p> <p>6/1/15: Mike to resurrect this as Tresys would like to run their tdml suite against both Daffodil and IBM DFDL.</p> <p>...</p> <p>10/2: No further progress</p> <p>24/2: Mike updating the Daffodil TDML test runner to handle unparser (ie, serializer) tests</p> <p>...</p> <p>14/4: No further progress</p> <p>28/4: Tresys have enhanced their tdml runner to allow unparser tests and round-trip tests (parser->unparser->parser) as well as the new tutorial tag (see action 228)</p> <p>12/5: Not discussed</p> <p>...</p> <p>3/11: No progress</p> <p>5/1/16: No progress. Needs more interoperability between implementations to be really useful.</p> <p>...</p> <p>24/5: No progress</p>
279	<p>Improve defaulting description to explicitly cover local groups (Mike)</p> <p>28/4: Only talks about elements, should mention local sequence and choice.</p> <p>12/5: Not discussed</p> <p>23/6: Section 15.1.3 needs to say what happens when a choice branch does not contain any elements; such a choice branch is selected (but see action 280 below as minOccurs '0' might change this). Section 9.4 also needs updating to say what happens when local groups are found within a complex type.</p> <p>11/8: Steve did some tests with IBM DFDL. Just need some words as above. Action assigned to Mike.</p> <p>25/8: In progress</p> <p>...</p> <p>5/1/16: No progress</p> <p>...</p> <p>24/5: No progress</p>
286	<p>Create table that compares Daffodil features to IBM DFDL features (Steve/Mike)</p> <p>16/2: Take table from section 22 of DFDL 1.0 spec and add columns showing implementations.</p> <p>1/3: No progress</p>

	10/5: No progress. Work on this for next call. 24/5: No further progress
287	Find a way to handle a variable path step in DFDL expression (All) 1/3: DFDL4S currently using a hack that embeds a regex in a path step. 10/5: No progress 24/5: Need example from DFDL4S

Closed actions

No	Action
282	<p>Does XPath have operators for checking a value is in a range ? (Steve)</p> <p>12/5: Investigate whether equivalent to DFDL4S 'in' operator exists.</p> <p>23/6: Mike has found an XPath 'intersect' operator. Handles the enumeration case well, but not as convenient for ranges as DFDL4S's 'in' operator.</p> <p>11/8: Looked back at the motivating example from DFDL4S. Agreed that DFDL functions to do the equivalent of 'in' and 'inrange' would be useful if nothing can be re-used from XPath. Steve to write up a proposal taking into account different data types.</p> <p>...</p> <p>22/9: Proposal sent by Steve for new functions <code>dfdl:checkValues()</code>, <code>dfdl:checkRangeInclusive()</code>, <code>dfdl:checkRangeExclusive()</code>. Discussed whether both range functions needed, and whether they should be allowed for float and double. Mike noted that <code>dfdl:checkConstraints()</code> and simple types could be used instead of all three functions if values were static. Follow-up with DFDL4S team to see if they had thought of that.</p> <p>3/11: DFDL4S not happy with usability of <code>dfdl:checkConstraints</code> type, but ok with the intersection operator. Next step is to see what the DFDL4S schema would look like if rewritten to use <code>dfdl:checkConstraints</code> and intersection.</p> <p>5/1/16: Agreed that <code>dfdl:checkConstraints()</code> is not ideal as it requires creation of union simple types. Steve to go back to his proposal from 22/9/15 and rework to include use of the intersect operator.</p> <p>16/2: No progress</p> <p>1/3: Discussed with ESA team. Steve to rework his proposal from 22/9/15.</p> <p>10/5: Reworked proposal mailed to WG. DFDL4S to review for next call please.</p> <p>24/5: Closed. DFDL4S have reviewed the proposal and are happy with it. Issue https://redmine.ogf.org/issues/314 created.</p>

Deferred actions

No	Action
241	<p>Public comment : Bi-di properties placement in precedence section (All)</p> <p>7/11: This looks deliberate but the asymmetry between parsing and unparsing is unclear. Really needs Daffodil or IBM DFDL to implement these properties, which has not happened yet. Deferring this action.</p> <p>...</p> <p>23/9: Candidate to be moved out to 1.1 ?</p>

Work items:

No	Item	Owner	Target	Status

