Graphs Task Force Report

Richard Cyganiak

1st RDF Working Group Face To Face Meeting, 13 April 2011, Amsterdam
Task before the WG

http://www.w3.org/2011/01/rdf-wg-charter

* (Required) “Standardize a model and semantics for multiple graphs and graphs stores”

* (Required) “Standardize the Turtle RDF Syntax. Either that syntax or a related syntax should also support multiple graphs and graph stores.”

* (Required) “Define and standardize a JSON Syntax for RDF … The goal is to [be] as complete as possible (including multiple graphs and graph stores)”

* (Time permitting) “Standardize additions to RDF/XML … Candidates
Inputs

- SPARQL’s RDF Dataset (and SPARQL Update’s Graph Store)
- Carroll et al.: Named Graphs
- Notation3: quoted graphs
- TriG
- N-Quads
- RDF/XML with rdf:graph (INRIA proposal)
- Reification
- (Typed graph literals)
Use Cases

http://www.w3.org/2011/rdf-wg/wiki/TF-Graphs-UC

- 5 Storage Use Cases
- 2 Query Use Cases
- 8 Provenance Use Cases
- 4 Use Cases for Providing a standard foundation for W3C specs
- 2 Advanced Annotations Use Cases
Proposals

- Lift SPARQL’s RDF Dataset into RDF Concepts and Abstract Syntax
  (a) Treat it only as an abstract syntax
  (b) Alex’s Semantics proposal
  (c) Antoine’s Semantics proposal

- Others???
  - Notation3-style quoted graphs
  - Only defining <IRI,graph> pairs might be sufficient
  - Reification or graph literals might have a role to play
Should we define Graph Literal datatypes?

- Datatypes ser:rdflxml, ser:turtle etc
- Lexical space is the format; value space the corresponding g-snap
- Turtle, RDF/XML could provide syntactic sugar
What is a named graph and what should we call it?

Name options:

(a) Named Graph

(b) named g-box

(c) g-pair

(d) IRI-graph-binding
ISSUE-15 (g-pair semantics)

What is the relationship between the IRI and the triples in a dataset/quad-syntax/etc?

- Leave it undefined (abstract syntax only)
- Define it
  - “IRI names this g-snap”
  - “IRI names a g-box containing this g-snap right now” (AWWW implications)
  - “IRI names a g-box that entails this g-snap right now”
- Let the user choose the relationship
ISSUE-17 (graph merge)

How are RDF datasets to be merged?

• What corresponds to RDF’s Graph merge?

• Merging an graph into a dataset; merging two datasets
ISSUE-21 (Shared Node-IDs)

Can Node-IDs be shared between parts of a quad/multigraph format?

- Corollary: Can blank nodes be shared between graphs?
- Most input proposals avoid this question
- Possibly related: blank node skolemization
Does multigraph syntax need to support empty graphs?

- TriG does
- N-Quads doesn’t
- SPARQL 1.1 distinguishes between stores that do and don’t
ISSUE-23 (multigraph-media-types)

Does going from single-graph to multi-graph require new format and new media types?

- Consumers would have to process quads instead of triples—implies major implementation and conceptual changes
- Hence: leave the existing formats triple-based: Turtle, RDF/XML, to a lesser degree N-Triples
- Distinct media types for multigraphs: N-Quads, TriG

Wednesday 13 April 2011
Issues: Discussion volume

<table>
<thead>
<tr>
<th>ISSUE</th>
<th>Nickname</th>
<th>Emails</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Graph Literals</td>
<td>57</td>
</tr>
<tr>
<td>14</td>
<td>named-g-boxes</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>g-pair semantics</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>graph merge</td>
<td>8</td>
</tr>
<tr>
<td>21</td>
<td>Shared Node-IDs</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td>empty graphs</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>multigraph-media-types</td>
<td>1</td>
</tr>
</tbody>
</table>
Candidate issues?

- Do we need nesting of graphs?
- Do we need a default graph?
- Do we need a concrete syntax?
- Reification—let this TF deal with it?
- Multigraph syntaxes might undermine AWWW’s notion of authority
Fulfilling the charter

- Look at a minimally invasive way of lifting SPARQL’s RDF Dataset into RDF Concepts and Abstract Syntax
- Evaluate additional possible features based on use cases
- Do not define a concrete syntax
- If we MUST have a concrete syntax: standardize N-Quads
- Avoid multigraphs in RDF/XML, JSON, Turtle