Finite Models in RDF(S), with datatypes

Peter F. Patel-Schneider[1] Pat Hayes[2]
[1]Nuance Communications [2]IHMC

RDF(S) Models

- Models and interpretations are the foundation of RDF(S) [1, 2].
- An interpretation is a way that the world might be, containing a universe of entities (including individuals, classes, etc.) and relationships between them.
- Interpretations map IRIs, blank nodes, and literals (*names*) to elements of the universe (entities).
- A model of an RDF graph is an interpretation that makes the statements in the graph true.
- The meaning of RDF(S), including entailment, is defined in terms of interpretations and models.
- Examining the properties of models and interpretations uncovers the properties of RDF(S).

How Big must RDF(S) Universes be?

- It appears that infinite universes are needed!
- There are an infinite number of container membership properties that all belong to `rdfs:ContainerMemebershipProperty`.
- All IRIs are interpreted. (New for 2013.)
- There are an infinite number of decimals (including integers).
- There are lots of floats.
- Looks like RDF(S) universes are infinite, even without data values.
- Are infinite universes really needed?

How Small can RDF(S) Universes be?

- If there are no datatypes, is it possible to consider only small (or finite) RDF(S) universes, without changing the meaning of RDFS?
- Ter Horst [3] showed how to ignore unused container membership properties in reasoning.
  - In model-theoretic terms, all unused container membership properties can be interpreted as a single entity.
  - I(`rdf:_n`) =I(`rdf:_n+1`) = I(`rdf:_n+2`)
  - Even though container membership properties are non-trivial, they all look the same.
- In the same way, unused IRIs and blank nodes can be interpreted as a single entity.
  - I(`ex:unused`) =I(`ex:notused`) = I(`ex:wasntused`)
  - Unused IRIs and blank nodes are all trivial (i.e., they all have no properties except belonging to `rdfs:Resource`).
- Removes two supports for the need for infinite universes.
  - Also shows up an issue with container membership properties.
- If there are no recognized datatypes, the RDF(S) universe can be finite (linear) without changing meaning.
- Constrast with OWL, where simple ontologies can require infinite universes.

How Small can RDF(S) Universes be, with datatypes?

- Is it possible to consider only small (or finite) RDF(S) universes, without changing the meaning of RDFS?
- Techniques similar to those above can be used to show that unused recognized literals can also be ignored.
  - Technically, they still make the universe infinite because they are distinct, but unused data values can't have any extra properties associated with them.
  - Can define pre-interpretations, where literals not appearing in an RDF graph are not interpreted.
  - Pre-interpretations don't change the meaning of RDF(S).
  - Removes the other support for the need for infinite universes.
  - Shows how weak RDF(S) datatypes are.
- The RDF(S) universe can be finite (linear) without changing meaning.

Theorem

To reason in RDF(S) it suffices to consider (pre-)interpretations whose universe is the same size as the number of names in an RDF graph (plus the unused important RDF(S) vocabulary, plus one).

Sub-Linear Universes with Disjointness or Disjunction

- Do universes smaller than the number of names in an RDF graph suffice?
- In RDF plus `owl:differentFrom`, smaller universes are not adequate.
- Just ask whether all the names are different from each other.
- If RDF plus disjunction, again smaller universes are not adequate.
- Consider the RDF graph

  $S_i\ S_1\ S_i.$ for $1 <= i <= n$.
- In an interpretation with less than n domain elements some particular different two of the S_i and S_j have the same denotation.
- So in any interpretation with less than n domain elements $S_i\ S_1\ S_j$. for different i, j.
- So the disjunction of all these triples is true in all such interpretations, but this is not a valid entailment.
- Thus interpretations with at least n elements must be considered.

Sub-Linear Universes with Blank Nodes

- Do universes smaller than the number of names in an RDF graph suffice?
- Consider the RDF graph

  $S_i\ S_1\ S_j$. for $1 \leq i \neq j \leq n$.
- In any model with less than n domain elements, some particular different two of the S_i and S_j have the same denotation.
- This entity is then related to the denotation of each of the S_i by S_1.
- So the RDF graph

  _:x $S_1\ S_j$. for $1 <= j <= n$

  is true in each of these models, but this graph is not entailed.
- Thus interpretations with at least n elements must be considered.
- If blank nodes are allowed in entailments, smaller universes are not adequate.

Sub-Linear Universes without Blank Nodes

- Do universes smaller than the number of names in an RDF graph suffice, if blank nodes are not permitted in entailments?
- Consider an interpretation I containing two domain elements e_1 and e_2 that are neither properties nor classes nor data values (call these domain elements *ordinary*).
  - Form I' from I by simply replacing e_1 and e_2 with a single domain element e throughout.
  - For N_1 and N_2 IRIs whose denotations in I are neither e_1 nor e_2, I' supports any triple of the form $N_1\ P\ N_2$. if and only if I supports the triple.
- For any particular B_1 and B_2 this process can be repeated until only three ordinary domain elements remain, producing an interpretation that doesn't add any entailments for triples between B_1 and B_2.
- Without blank nodes, three ordinary domain elements are adequate.

Theorem

To reason in RDF(S) without blank nodes it suffices to consider interpretations with only three ordinary entities (plus entities for classes, properties, and literals).

Conclusions

- RDF is very weak:
  - Can't require existence of unmentioned resources.
- RDF without blank nodes is extremely weak:
  - Can't require existence of more than three ordinary resources.
- Adding something like `owl:differentFrom` would strengthen RDF.

References

[1] Patrick Hayes. RDF semantics. W3C Recommendation, `http://www.w3.org/TR/rdf-mt/`, 2004.

[2] Patrick Hayes and Peter F. Patel-Schneider. RDF 1.1 semantics. W3C Working Draft, `http://www.w3.org/TR/rdf11-mt/`, July 2013.

[3] Herman J. ter Horst. Completeness, decidability and complexity of entailment for RDF Schema and a semantic extension involving the OWL vocabulary. *Journal of Web Semantics*, 3(2-3):79–115, 2005.

The Fine Print

- Lots of shortcuts have been taken in this presentation.
- No shortcut invalidates the results; no shortcut is misleading.

pfpschneider@gmail.com