Ontological engineering for rhetorical figures and beyond

Cliff O'Reilly

Computational Rhetoric Workshop
Waterloo, 13th August 2016

Outline

Expressivity

- There are infinite ways to model the world
- There are restrictions for a given framework, e.g. types and properties
- It's a creative exercise
- There are no wrong answers (until the model breaks)

Architecture

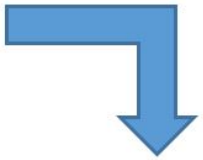
- OWL/RDFS due to its interoperability and expressivity
- Key distinction between the ontology (e.g. owl file) and the knowledge base (e.g. RDF triples)
  - Deliver the knowledge base and ontology together
  - Use an RDF Store to separate ontology from Knowledge Base
- Thinking about usage is important

Architecture - example

- Anaphora is a Rhetorical Device
- Expressed as an RDF triple in XML, e.g.

RhetoricalDevices.owl

```
<owl:Class rdf:ID="RhetoricalDevice"/>
<owl:ObjectProperty rdf:ID="hasRhetoricalDevice">
  <rdfs:range rdf:resource="#RhetoricalDevice"/>
  <rdfs:domain rdf:resource="http://repositori.com/sw/onto/DocStruct.owl#Doc"/>
</owl:ObjectProperty>
.
.
```



MyExampleInstances.rdf

```
<rdf:Description rdf:about="http://repositori.com/sw/onto/RhetoricalDevices.owl#RhetoricalDevice"/>
  <rdf:type rdf:resource="Anaphora"/>
<rdf:Description/>
<rdf:Description rdf:about="http://repositori.com/sw/onto/RhetoricalDevices.owl#RhetoricalDevice"/>
  <rdf:type rdf:resource="Epizeuxis"/>
<rdf:Description/>
.
.
```

Why?

- Easy to do
- Functional separation
- Powerful, e.g. increased scope for inference
- All elements can be accessed from the linked ontology
- Semantic Web principles - opportunity for targeted re-use
- Can become unwieldy

How?

Linking ontologies

W3C standard gives the owl:imports element which is added to the header in the containing OWL ontology and pointed to the resource to import, e.g. <owl:imports rdf:resource="http://repositori.com/sw/onto/RhetoricalDevices.owl"/>

Background

- 2010 MSc dissertation project
- Automatic rhetorical figure detection using Semantic Web technology
- OWL ontologies and rules in SWRL (Semantic Web Rule Language)
- A personal interest in rhetorical patterns and history
- Unpublished paper: "Lassoing Rhetoric with OWL and SWRL"

The model

- Java program, calling on GATE[1], Protege[2] and JESS[3] reasoner APIs
- Suite of OWL ontologies
- Extension into Semantic Web - Linked Data
- Designed for inference - discovering new knowledge via logic rules
- Ontologies designed with the goal in mind, i.e. reverse engineered from logical rules

[1] https://gate.ac.uk/
[2] http://protegewiki.stanford.edu/wiki/ProtegeOWL_API_Programmers_Guide
[3] http://www.jessrules.com/jess/index.shtml

In detail - ontologies

- Controller ontology
  - Lassoing Rhetoric (`http://repositori.com/sw/onto/LassoingRhetoric.owl`)
- Imported ontologies
  - Document Structure (`http://repositori.com/sw/onto/DocStruct.owl`)
  - Rhetorical Devices (`http://repositori.com/sw/onto/RhetoricalDevices.owl`)
  - Rhetorical Structures (`http://repositori.com/sw/onto/RhetoricalStruct.owl`)
  - Language Tags (`http://repositori.com/sw/onto/langtag.owl`)
  - GATE (`http://repositori.com/sw/onto/gate.owl`)
  - Verb Noun Combinations (`http://repositori.com/sw/onto/VerbNounCombo.owl`)

In detail - example 1

Anaphora (Epanaphora)

A figure characterised by a repeating word at the beginning of successive clauses, e.g. Churchill's speech to Parliament in 1940: "We shall go on to the end. We shall fight in France. We shall fight on the seas and oceans..".
The pattern chosen to be modelled in this project is a two word phrase repeating in successive sentence beginnings.

In detail - example 1

Anaphora (Epanaphora) - example 1 - simplified syllogism

IF Document HAS Paragraph
AND IF Paragraph HAS Sentence
AND IF Sentence HAS Clause A
AND IF Sentence HAS Clause B
AND IF Clause A HAS Word X
AND IF Clause A HAS Word Y
AND IF Clause B HAS Word F
AND IF Clause B HAS Word G
AND IF Word X IS THE SAME AS Word F
AND IF Word Y IS THE SAME AS Word G
THEN Document HAS Anaphora

In detail - example 1

Anaphora (Epanaphora) - Horn Clause

DocStruct:Doc(?h) ∧ DocStruct:hasParagraph(?h,?i)
∧ gate:Sentence(?z) ∧ gate:Sentence(?c)
∧ gate:word(?a) ∧ gate:word(?x)
∧ DocStruct:hasSentence(?i,?z) ∧ DocStruct:hasNextSentence(?z,?c)
∧ DocStruct:hasFirstWord(?z,?x) ∧ DocStruct:hasFirstWord(?c,?a)
∧ DocStruct:hasNextWord(?x,?y) ∧ DocStruct:hasNextWord(?a,?b)
∧ gate:hasString(?x,?d) ∧ gate:hasString(?y,?e) ∧ gate:hasString(?a,?f) ∧
gate:hasString(?b,?g)
∧ swrlb:equal(?d,?f) ∧ swrlb:equal(?e,?g)
∧ gate:hasStartNode(?x,?j) ∧ gate:hasEndNode(?b,?k)

$\Longrightarrow$ $RhetDev : hasRhetoricalDevice(?h, RhetDev : Anaphora)$∧
gate:hasStartNode(RhetDev:Anaphora,?j) ∧ gate:hasEndNode(RhetDev:Anaphora,?k)

In detail - example 2

Epizeuxis

A figure characterised by an emphatic repetition of a word with no other words in between, e.g. Poe's poem The Bells:
"How they tinkle, tinkle, tinkle ... From the bells, bells, bells, bells ...".
The pattern chosen to be modelled in this project is the same word repeating three times in succession.

In detail - example 2

Epizeuxis - Horn Clause

DocStruct:Doc(?d) $\wedge$ DocStruct:hasNextWord(?x, ?y) $\wedge$ DocStruct:hasNextWord(?y, ?z)
$\wedge$ gate:hasString(?x, ?a) $\wedge$ gate:hasString(?y, ?b) $\wedge$ gate:hasString(?z, ?c)
$\wedge$ swrlb:equal(?a, ?b) $\wedge$ swrlb:equal(?a, ?c)
$\wedge$ gate:hasStartNode(?x, ?e) $\wedge$ gate:hasEndNode(?z, ?f)
$\Longrightarrow$ $RhetDev : hasRhetoricalDevice(?d, RhetDev : Epizeuxis)\wedge$
gate:hasStartNode(RhetDev:Epizeuxis, ?e) $\wedge$ gate:hasEndNode(RhetDev:Epizeuxis, ?f)

In detail - extension

- A key idea from the project was to take advantage of Linked Data, e.g.
  - DBPedia -utilising data from the web to inform the ontology/knowledge base, such as world knowledge
  - Wordnet - using data to extend the knowledge base, e.g. by including synonyms and resultant inferences
- Facilitated via the Java program, but dynamic and parameterised

In detail - example 3

Historic Present

A figure characterised by a reference to a thing in the past as if it were still referable in the present, such as a person no longer alive having thoughts in the present day. For example,
"What Socrates is trying to do here is...".
The pattern chosen to be modelled in this project is based on the instantiation of a class called DeadPerson and this being referenced by a verb in the present tense.

In detail - example 3

Historic Present - Horn Clause

DocStruct:Doc(?x) $\wedge$ gate:paragraph(?y) $\wedge$ DocStruct:hasParagraph(?x, ?y)
$\wedge$ gate:Sentence(?z) $\wedge$ DocStruct:hasSentence(?y, ?z)
$\wedge$ gate:word(?a) $\wedge$ DocStruct:hasWord(?z, ?a)
$\wedge$ gate:word(?b) $\wedge$ DocStruct:hasWord(?z, ?b)
$\wedge$ DeadPerson(?a) $\wedge$ langtag:ThirdPersonSingularPresentVerb(?b) $\wedge$
gate:hasDependency(?b, ?a)
$\wedge$ gate:hasStartNode(?a, ?c) $\wedge$ gate:hasEndNode(?a, ?d)

$\Longrightarrow$ $RhetDev : hasRhetoricalDevice(?x, RhetDev : Historic\_Present)\wedge$
gate:hasStartNode(RhetDev:Historic_Present, ?c) $\wedge$
gate:hasEndNode(RhetDev:Historic_Present, ?d)

Ontology of cognition and intelligence

- A new ontology to capture primitive cognitive patterns
- Early stages of development
- PhD proposal
- Domain: Intelligent agents in 3D gaming
- ... however, it could apply to any application especially analysis of Rhetoric

Ontology of cognition and intelligence

- Influence from Cognitive Science, Cognitive Linguistics and Embodiment theory
- Currently the most primitive aspects of cognition are under development, e.g.
  - Identity / Categorisation
  - Perception of primitive shapes, e.g. lines and boundaries etc
  - Material behaviour
  - Quantity / number
  - Persistence / temporality
  - Size, Scale etc
  - Proximity
  - Difference / similarity
- The scope includes developing more complex functions involving intentionality, goals, decisions, beliefs, frames etc

Cognitive Linguistics and Embodiment

- Image Schema; Conceptual Metaphor - Lakoff and Johnson
  - Containment, location, motion, lines, adjacency, contact, orientation etc
  - Metaphors - More-Is-Up, Less-Is-Down, Love-As-Journey etc
- Embodiment; Situated/Embodied Cognition
  - Cognition is grounded in physical aspects and implications of being in a body in the world
- Natural language, and therefore Rhetorical forms, is strongly influenced by this underlying cognitive framework, e.g. via metaphors and words such as Up/Down etc

How does this relate to rhetoric?

- Descriptive, e.g. adding more knowledge to our models
- New ways of describing patterns
- Discovering new knowledge via inference
- "Ask not what your country can do for you; ask what you can do for your country" (Antimetable) could be described in terms of:
  - Relative proximity (physical and/or other) of like-terms
  - "Actions are self-propelled movements" - Event Structure conceptual metaphor
  - Size comparison of what a country can do versus what you can do

Discussion

- Questions?
- Suggestions?
- Comments?