A Short Survey of Discourse Representation Models

Tudor Groza, Siegfried Handschuh, Tim Clark, Simon Buckingham Shum and Anita de Waard

Semantic Web Applications in Scientific Discourse Workshop @ ISWC 2009

26 October 2009
Chantilly, Virginia, US
Outline

- Introduction
- Analysis features
- Models
- Analysis overview
- Conclusion
Outline

- Introduction
- Analysis features
- Models
- Analysis overview
- Conclusion
Introduction

- **Dissemination – communication process**
- **Externalization**
  - Implicit – materialized as publications
  - Explicit – required for machine interpretation
- **Discourse structuring and analysis**
  - Discourse representation models
  - Computational linguistic approaches – automatic extraction of epistemic items
Introduction (cont.)

- Our focus: Discourse representation models
  - Succinct overview
  - Brief comparative analysis

- Tentative goal: an unified discourse representation model
Analysis features

- Coarse-grained rhetorical structure
- Fine-grained rhetorical structure
- Relations – types of relations used
- Polarity – explicit positive vs. negative
- Weights – explicit numeric weight of relations
- Provenance – localization in text
- Shallow metadata support
- Domain knowledge
- Purpose – intended use of the model
- Evaluation and uptake
Outline

- Introduction
- Analysis features
- Models
- Analysis overview
- Conclusion
Models

- Harmsze’s model
- The Scholarly Ontologies project
- De Waard’s model
- The SWAN Ontology
- The SALT Framework
Harmsze’s model

- **Purpose**: presentation of scientific information in electronic articles

- **Coarse-grained structure**
  - Modules and elementary modules
  - E.g.: Meta-information, Positioning, Methods, Results, Interpretation, Outcome

- **Relations**
  - Organizational links
  - Scientific discourse links
    - Communicative function – *elucidation, argumentation, clarification*
    - Content relations – *elaboration, aggregation*
The Scholarly Ontologies Project

- **Purpose**: General structuring of coherence and argumentation

- **Fine-grained structure**
  - Atomic nodes – short pieces of text
  - Claims – connected nodes

- **Relations**
  - Cognitive Coherent Relations – Sanders et al.
  - Explicit polarity and weights – *proves* (+1) vs. *refutes* (-1)
  - Types: causal, problem related, similarity, general, supports/challenges, taxonomic
De Waard’s model

- **Purpose**: Modularization of scientific publications

- **Coarse-grained structure**
  - Annotation – Background – Contribution – Discussion – Entities

- **Relations**
  - Argumentative
  - Explicit polarity
  - E.g.: *proves vs. refutes; agrees vs. disagrees*
The SWAN Ontology

- **Purpose**: Creation of knowledge bases
  - Initially in the context of the Alzheimer Disease Research

- **General structure**
  - 6 main elements: people, bibliographic records, life science entities, tags, versions, discourse elements

- **Fine-grained structure**
  - Discourse Element, Research Statement, Research Question, Structure Comment

- **Relations**
  - Argumentative
  - E.g.: consistentWith, inconsistentWith, discusses
The SALT Framework

- **Purpose**: Structuring of rhetoric and argumentation in scientific publications

- **General structure**
  - 3 layers

- **Coarse-grained structure**
  - Rhetorical blocks: Introduction, Conclusion, ...

- **Fine-grained structure**
  - Rhetorical elements: Claims, Supports, ...

- **Relations**
  - Rhetorical relations (Rhetorical Structure of Text – Mann et al.): *Antithesis, Consequence, ...*
  - Argumentative relations
Outline

- Introduction
- Analysis features
- Models
- Analysis overview
- Conclusion
<table>
<thead>
<tr>
<th>Feature/Approach</th>
<th>Coarse-grained rhetorical structure</th>
<th>Fine-grained rhetorical structure</th>
<th>Relations</th>
<th>Polarity</th>
<th>Weights</th>
<th>Provenance</th>
<th>Shallow Metadata Support</th>
<th>Domain Knowledge</th>
<th>Purpose</th>
<th>Evaluation and uptake</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harmsze</td>
<td>Modules</td>
<td>Elementary module</td>
<td>Structuring, organisational and discourse</td>
<td>Implicit (within relations)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Open</td>
<td>Presentation of scientific information in electronic articles</td>
<td>Preliminary evaluation</td>
</tr>
<tr>
<td>ScholOnto</td>
<td>No</td>
<td>Node, Claim</td>
<td>Cognitive Coherent</td>
<td>Explicit (+/-)</td>
<td>Explicit (1, 2)</td>
<td>Yes (duplicates)</td>
<td>No</td>
<td>Open</td>
<td>General structuring of coherence and argumentation</td>
<td>Evaluated and widely used</td>
</tr>
<tr>
<td>De Waard</td>
<td>Rhetorical Blocks</td>
<td>Rhetorical element</td>
<td>Argumentative</td>
<td>Explicit (within the pairs of relations)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Open</td>
<td>Modularization of scientific publications</td>
<td>N/A</td>
</tr>
<tr>
<td>SWAN</td>
<td>No</td>
<td>Discourse element, Research statement</td>
<td>Argumentative and Cognitive Coherent</td>
<td>Implicit (within relations)</td>
<td>No</td>
<td>Yes (duplicates)</td>
<td>Yes (Gene, Protein)</td>
<td></td>
<td>Creation of a knowledge base</td>
<td>Evaluated and widely used</td>
</tr>
<tr>
<td>SALT</td>
<td>Rhetorical Blocks</td>
<td>Rhetorical element (Nucleus, Satellite, Claim, Support)</td>
<td>Rhetorical and argumentative</td>
<td>Implicit (within relations)</td>
<td>No</td>
<td>Yes (pointers)</td>
<td>Yes</td>
<td>Open</td>
<td>Structuring of the rhetoric and argumentation in scientific publications</td>
<td>Evaluated and infrequently used</td>
</tr>
</tbody>
</table>
Towards an unified discourse representation model

- Proper balance of currently existing features
- Emphasis on practicality for uptake maximization
- General structure
  - Layered – e.g. SWAN, SALT
- Coarse-grained structure
  - Rhetorical blocks – e.g. ABCDE, SALT
- Fine-grained structure
  - Discourse elements
- Relations
  - 2 layers
  - Argumentative
  - Rhetorical relations
Concrete (Web-oriented) Example
Outline

- Introduction
- Analysis features
- Models
- Analysis overview
- Conclusion
Conclusion

- Succinct overview of current discourse representation models
- Brief comparative analysis
- Next steps: ... open for discussion

Thank you!